Proceedings of the HERA - LHC workshop

Author List

S. Alekhin, G. Altarelli, N. Amapane, J. Andersen, V. Andreev, M. Arneodo, V. Avati, V. Avati, H. Böttcher, J. Baines, R.D. Ball, A. Banfi, S.P. Baranov, J. Bartels, O. Behnke, R. Bellan, J. Blümlein, S. Bolognesi, M. Boonekamp, D. Bourilkov, J. Braciník, G. Bruni, A. Bruni, A. Buckley, A. Bunyatyan, C.M. Buttar, J.M. Butterworth, S. Butterworth, M. Cacciari, T. Carli, G. Cerminara, S. Chekanov, M. Ciafaloni, D. Colferai, J. Collins, A. Cooper-Sarkar, G. Corcella, M. Corradi, B.E. Cox, R. Croft, Z. Czyczula, A. Dainese, M. Dasgupta, G. Davatz, L. Del Debbio, Y. Delenda, M. Diehl, S. Diglio, G. Dissertori, M. Dittmar, B. E. Cox, K.J. Eskola, T.O. Eynck, J. Feltesse, F. Ferro, R.D. Field, J. Forshaw, S. Forte, A. Geiser, S. Gieseke, A. Glazov, T. Gleisberg, P. Golonka, E. Gotsman, G. Grindhammer, M. Grothe, C. Group, M. Groys, A. Guffanti, G. Gustafson, C. Gwenlan, S. Höche, C. Hogg, J. Huston, G. Iacobucci, G. Ingelman, S. Jadach, H. Jung, J. Kalliopuska, M. Kapishin, B. Kersevan, V. Khoze, M. Klasen, M. Klein, B.A. Kniehl, V.J. Kolhinen, H. Kowalski, G. Kramer, F. Krauss, S. Kretzer, K. Kutak, J.W. Lämsä, L. Lönnblad G. Laštovička-Medin, T. Laštovička, E. Laenen, Th. Lagouri, J.I. Latorre, N. Lavesson, V. Lendermann, E. Levin, A. Levy, A.V. Lipatov, M. Lublinsky, L. Lytkin, T. Mäki, L. Magnea, F. Maltoni, M. Mangano, U. Maor, C. Mariotti, N. Marola, A.D. Martin, A. Meyer, S. Moch, J. Monk, A. Moraes, A. Morsch, L. Motyka, E. Naftali, P. Newman, A. Nikitenko, F. Oljemark, R. Orava, M. Ottela, K. Österberg, K. Peters, F. Petrucci, A. Piccione, A. Pilkington, K. Piotrzkowski, O.I. Piskounova, A. Proskuryakov, A. Prygarin, J. Pumplin, K. Rabbertz, R. Ranieri, V. Ravindran, B. Reisert, E. Richter-Was L. Rinaldi, P. Robbe, E. Rodrigues, J. Rojo, H. Ruiz, M. Ruspa, M.G. Ryskin, A. Sabio Vera, G.P. Salam, A. Schälicke, S. Schätzel, T. Schörner-Sadenius, I. Schienbein, F-P. Schilling, S. Schumann, M.H. Seymour, F. Siegert, T. Sjöstrand M. Skrzypek, J. Smith, M. Smizanska, H. Spiesberger, A. Staśto, H. Stenzel, W.J. Stirling, P. Szczypka, S. Tapprogge, C. Targett-Adams, M. Tasevsky, T. Teubner, R.S. Thorne, A. Tonazzo, A. Tricoli, N. Tuning, J. Turnau, U. Uwer, P. Van Mechelen, R. Venugopalan, M. Verducci, J.A.M. Vermaseren, R. Vogt, A. Vogt, B.F.L. Ward , Z. Was, G. Watt, B.M. Waugh, C. Weiser, M.R. Whalley, J. Winter, S.A. Yost, G. Zanderighi, N.P. Zotov

Contents

Plenary Presentations

Instanton-Induced Processes
An Overview

F. Schrempp
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Abstract
A first part of this review is devoted to a summary of our extensive studies of the discovery potential for instanton (I)-induced, deep-inelastic processes at HERA. Included are some key issues about I-perturbation theory, an exploitation of crucial lattice constraints and a status report about the recent I-search results by the HERA collaborations H1 and ZEUS in relation to our predictions. Next follows a brief outline of an ongoing project concerning a broad exploration of the discovery potential for hard instanton processes at the LHC. I then turn to an overview of our work on high-energy processes, involving larger-sized instantons. I shall mainly focus on the phenomenon of saturation at small Bjorken-x from an instanton perspective. In such a framework, the saturation scale is associated with the conspicuous average instanton size, $\langle\rho\rangle \sim 0.5$ fm, as known from lattice simulations. A further main result is the intriguing identification of the "Colour Glass Condensate" with the QCD *sphaleron* state.

1 Setting the stage

Instantons represent a basic non-perturbative aspect of non-abelian gauge theories like QCD. They were theoretically discovered and first studied by Belavin *et al.* [1] and 't Hooft [2], about 30 years ago.

Due to their rich vacuum structure, QCD and similar theories include topologically non-trivial fluctuations of the gauge fields, which in general carry a conserved, integer topological charge Q. Instantons ($Q = +1$) and anti-instantons ($Q = -1$) represent the simplest building blocks of topologically non-trivial vacuum structure. They are explicit solutions of the euclidean field equations in four dimensions [1]. They are known to play an important rôle in the transition region between a partonic and a hadronic description of strong interactions [3]. Yet, despite substantial theoretical evidence for the im-

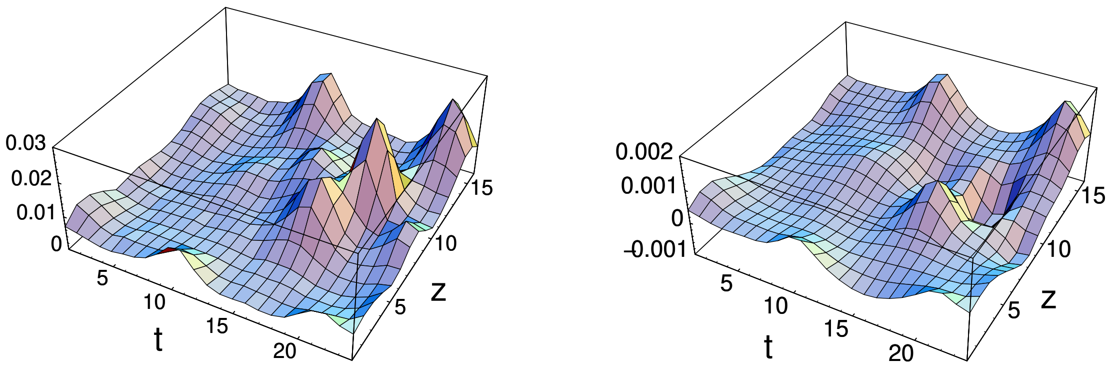


Fig. 1: Contribution from three instantons ($Q = +1$) and two anti-instantons ($Q = -1$) to the Lagrangian (left)) and the topological charge density (right) in a lattice simulation [4] (after cooling). The euclidean coordinates x and y are kept fixed while the dependence on z and t is displayed.

portance of instantons in chiral symmetry breaking and hadron spectroscopy, their direct experimental verification is lacking until now.

It turns out, however, that a characteristic *short distance* manifestation of instantons can be exploited [5] for an experimental search: Instantons induce certain (hard) processes that are forbidden in usual perturbative QCD. These involve all (light) quark flavours democratically along with a violation of chirality, in accord with the general chiral anomaly relation [2]. Based on this crucial observation, deep-inelastic scattering (DIS) at HERA has been shown to offer a unique opportunity [5] to discover such instanton-induced processes. It is of particular importance that a theoretical prediction of both the corresponding rate [6, 7, 8] and the characteristic event signature [5, 10, 11, 12] is possible in this hard scattering regime[1]. The instanton-induced cross section turns out to be in a measurable range [7, 10]. Crucial information on the region of validity for this important result, based on instanton-perturbation theory, comes from a high-quality lattice simulation [8, 13]. Another interesting possible spin-dependent signature of instantons in DIS, in form of a characteristic azimuthal spin asymmetry, has recently been discussed in Ref. [14].

In a first part (Sect. 2), I shall review our extensive investigations of deep-inelastic processes induced by small instantons. This includes a “flow-chart” of our calculations based on I-perturbation theory [6, 7], an exploitation of crucial lattice constraints [8, 13] and a confrontation [12] of the recent I-search results by the HERA collaborations H1 and ZEUS [15, 16] with our predictions. Next I shall briefly outline in Sect. 3 an ongoing project [17] to investigate theoretically and phenomenologically the discovery potential of hard instanton processes at the LHC. In Sect. 4, I then turn to an overview of our work [18, 19, 20, 21] on high-energy processes involving larger-sized instantons. I shall focus mainly on the important theoretical challenge of the phenomenon of saturation at small Bjorken-x from an instanton perspective. In such a framework we found [18, 19, 20, 21] that the conspicuous average instanton size scale, $\langle\rho\rangle \sim 0.5$ fm, as known from lattice simulations [13], plays the rôle of the saturation scale. As a further main and intriguing result, we were led to associate the “Colour Glass Condensate” [22] with the QCD *sphaleron* state [23]. For another more recent approach to small-x saturation in an instanton background with main emphasis on Wilson loop scattering and lacking direct lattice input, see Ref. [24]. The conclusions of this overview may be found in Sect. 5.

2 Small instantons in deep-inelastic scattering

2.1 Instanton-perturbation theory

Let us start by briefly summarizing the essence of our theoretical calculations [6, 7] based on so-called I-perturbation theory. As we shall see below, in an appropriate phase-space region of deep-inelastic scattering with generic hard scale $\mathcal{Q}$, the contributing I's and $\overline{I}$'s have *small size* $\rho \lesssim \mathcal{O}(\frac{1}{\alpha_s(\mathcal{Q})\mathcal{Q}})$ and may be self-consistently considered as a *dilute* gas, with the small QCD coupling $\alpha_s(\mathcal{Q})$ being the expansion parameter like in usual perturbative QCD (pQCD). Unlike the familiar expansion about the trivial vacuum $A_\mu^{(0)} = 0$ in pQCD, in I-perturbation theory the path integral for the generating functional of the Green's functions in Euclidean position space is then expanded about the known, classical one-instanton solution, $A_\mu = A_\mu^{(I)}(x) + \ldots$. After Fourier transformation to momentum space, LSZ amputation and careful analytic continuation to Minkowski space (where the actual on-shell limits are taken), one obtains a corresponding set of modified Feynman rules for calculating I-induced scattering amplitudes. As a further prerequisite, the masses m_q of the active quark flavours must be light on the scale of the inverse effective I-size $1/\rho_{\text{eff}}$, i.e. $\mathrm{m}_q \cdot \rho_{\text{eff}} \ll 1$. The leading, I-induced, chirality-violating process in the deep-inelastic regime of $e^{\pm}\mathrm{P}$ scattering is displayed in Fig. 2 (left) for $n_f = 3$ massless flavors. In the background of an I $(\overline{I}\,)$ (of topological charge $Q = +1\ (-1)$), all n_f massless quarks and anti-quarks are right (left)-handed such that the I-induced subprocess emphasized in the dotted box of Fig. 2 (left)

[1]For an exploratory calculation of the instanton contribution to the gluon-structure function, see Ref. [9].

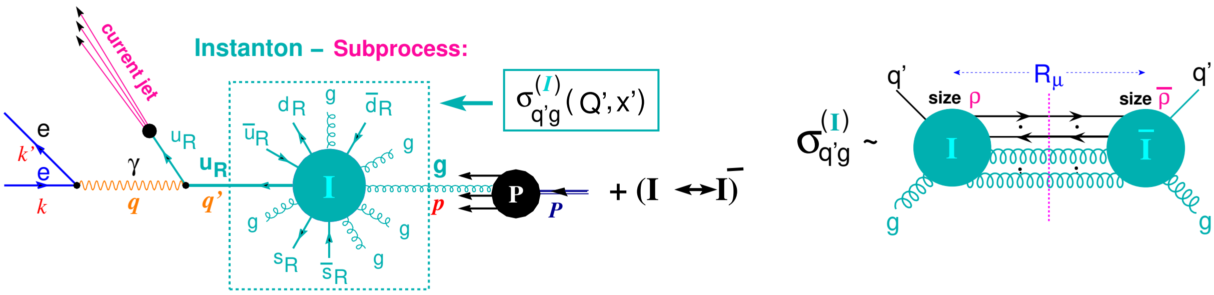


Fig. 2: (left): Leading, instanton-induced process in deep-inelastic $e^{\pm}$P scattering for $n_f = 3$ massless flavours. (right): Structure of the total cross section $\sigma^{(I)}_{q'\,g}$ for the chirality-violating "instanton-subprocess" $q'\,g \overset{(I)}{\Rightarrow} X$ according to the optical theorem. Note the illustration of the collective coordinates $\rho, \overline{\rho}$ and R_μ.

involves a violation of chirality $Q_5 = \#\,(q_{\rm R} + \overline{q}_{\rm R}) - \#\,(q_{\rm L} + \overline{q}_{\rm L})$ by an amount,

$$\Delta Q_5 = 2\,n_f\,Q, \tag{1}$$

in accord with the general chiral anomaly relation [2]. Within I-perturbation theory, one first of all derives the following factorized expression in the Bjorken limit of the I-subprocess variables $Q'^{\,2}$ and x' (c. f. Fig. 2 (left)):

$$\frac{{\rm d}\sigma^{(I)}_{\rm HERA}}{{\rm d}x'{\rm d}Q'^{\,2}} \simeq \frac{{\rm d}\mathcal{L}^{(I)}_{q'g}}{{\rm d}x'{\rm d}Q'^{\,2}} \cdot \sigma^{(I)}_{q'g}(Q',x') \quad \text{for} \; \left\{ \begin{array}{l} Q'^{\,2} = -q'^{\,2} > 0 \text{ large}, \\ 0 \le x' = \frac{Q'^{\,2}}{2p\cdot q'} \le 1 \text{ fixed}\,. \end{array} \right. \tag{2}$$

In Eq. (2), the differential luminosity, ${\rm d}\mathcal{L}^{(I)}_{q'\,g}$ counts the number of $q'\,g$ collisions per eP collisions. It is given in terms of integrals over the gluon density, the virtual photon flux, and the (known) flux of the virtual quark q' in the instanton background [7].

The essential instanton dynamics resides, however, in the total cross-section of the I-subprocess $q'\,g \overset{I}{\Rightarrow} X$ (dotted box of Fig. 2 (left) and Fig. 2 (right)). Being an observable, $\sigma^{(I)}_{q'g}(Q',x')$ involves integrations over all I and $\overline{I}$-"collective coordinates", i. e. the I $(\overline{I}\,)$ sizes ρ $(\overline{\rho}\,)$, the $I\overline{I}$ distance four-vector R_μ and the relative $I\overline{I}$ color orientation matrix U.

$$\sigma^{(I)}_{q'\,g} = \int d^4R\,{\rm e}^{{\rm i}\,(p+q')\cdot R} \int\limits_0^\infty d\rho \int\limits_0^\infty d\overline{\rho}\;{\rm e}^{-(\rho+\overline{\rho})Q'}\;D(\rho)\,D(\overline{\rho}\,) \int dU\,{\rm e}^{-\frac{4\pi}{\alpha_s}\,\Omega\left(U,\frac{R^2}{\rho\overline{\rho}},\frac{\overline{\rho}}{\rho}\right)}\,\{\ldots\} \tag{3}$$

Both instanton and anti-instanton degrees of freedom enter here, since the I-induced cross-section results from taking the modulus squared of an amplitude in the single I-background. Alternatively and more conveniently (c. f. Fig. 2 (right)), one may invoke the optical theorem to obtain the cross-section (3) in Minkowski space as a discontinuity of the $q'\,g$ forward elastic scattering amplitude in the $I\overline{I}$-background [7]. The $\{\ldots\}$ in Eq. (3) abreviates smooth contributions associated with the external partons etc. Let us concentrate on two crucial and strongly varying quantities of the I-calculus appearing in Eq. (3): $D(\rho)$, the (reduced) I-size distribution [2, 28], and $\Omega\left(U, \frac{R^2}{\rho\overline{\rho}}, \frac{\overline{\rho}}{\rho}\right)$, the $I\overline{I}$ interaction, associated with a resummation of final-state gluons. Both objects are *known* within I-perturbation theory, formally for $\alpha_s(\mu_r)\ln(\mu_r\,\rho) \ll 1$ and $\frac{R^2}{\rho\overline{\rho}} \gg 1$ (diluteness), respectively, with μ_r being the renormalization scale. In the $I\overline{I}$-valley approach [25], the functional form of $\Omega^{I\overline{I}}_{\rm valley}$ is analytically known [26, 27] (formally) for *all* values of $R^2/(\rho\bar{\rho})$. The *actual* region of validity of the valley approach is an important issue to be addressed again later.

Most importantly, the resulting power-law behaviour for the I-size distribution,

$$D(\rho) \propto \rho^{\beta_0 - 5 + \mathcal{O}(\alpha_s)}, \tag{4}$$

involving the leading QCD β-function coefficient, $\beta_0 = \frac{11}{3}\,N_c - \frac{2}{3}\,n_f,\ (N_c = 3)$, generically spoils the calculability of I-observables due to the bad IR-divergence of the integrations over the $I\ (\overline{I}\,)$-sizes for large $\rho\ (\overline{\rho}\,)$. Deep-inelastic scattering represents, however, a crucial exception: The *exponential* "form factor" $\exp(-Q^\prime(\rho+\overline{\rho}\,))$ that was shown [6] to arise in Eq. (3), insures convergence and *small* instantons for large enough $Q^\prime$, despite the strong power-law growth of $D(\rho)$. This is the key feature, warranting the calculability of I-predictions for DIS.

It turns out that for (large) $Q^\prime \neq 0$, all collective coordinate integrations in $\sigma^{(I)}_{q^\prime g}$ of Eq. (3) may be performed in terms of a *unique saddle point*:

$$\begin{aligned} U^* \quad &\Leftrightarrow \quad \text{most attractive relative } I\overline{I} \text{ orientation in color space,} \\ \rho^* \quad &= \quad \overline{\rho}^* \sim \frac{4\pi}{\alpha_s(\frac{1}{\rho^*})}\,\frac{1}{Q^\prime}; \quad \frac{R^{*\,2}}{\rho^{*\,2}} \overset{Q^\prime \text{ large}}{\sim} 4\frac{x^\prime}{1-x^\prime} \end{aligned} \tag{5}$$

This result underligns the self-consistency of the approach, since for large $Q^\prime$ and small $(1-x^\prime)$ the saddle point (5), indeed, corresponds to widely separated, small I's and $\overline{I}$'s.

2.2 Crucial impact of lattice results

The I-size distribution $D(\rho)$ and the $I\overline{I}$ interaction $\Omega\left(U, \frac{R^2}{\rho\overline{\rho}}, \frac{\overline{\rho}}{\rho}\right)$ form a crucial link between deep-inelastic scattering and lattice observables in the QCD vacuum [8].

Lattice simulations, on the other hand, provide independent, non-perturbative information on the *actual* range of validity of the form predicted from I-perturbation theory for these important functions of ρ and R/ρ, respectively. The one-to-one saddle-point correspondence (5) of the (effective) collective I-coordinates $(\rho^*, R^*/\rho^*)$ to $(Q^\prime, x^\prime)$ may then be exploited to arrive at a "fiducial" $(Q^\prime, x^\prime)$ region for our predictions in DIS. Let us briefly summarize the results of this strategy [8].

We have used the high-quality lattice data [13, 8] for quenched QCD ($n_f = 0$) by the UKQCD collaboration together with the careful, non-perturbative lattice determination of the respective QCD Λ-parameter, $\Lambda^{(n_f=0)}_{\overline{\rm MS}} = (238{\pm}19)$ MeV, by the ALPHA collaboration [29]. The results of an essentially parameter-free comparison of the continuum limit [8] for the simulated $(I+\overline{I})$-size and the $I\overline{I}$-distance distributions with I-perturbation theory versus ρ and R/ρ, respectively, is displayed in Fig. 3. The UKQCD data for the $I\overline{I}$-distance distribution provide the first direct test of the $I\overline{I}$ interaction $\Omega\left(U, \frac{R^2}{\rho\overline{\rho}}, \frac{\overline{\rho}}{\rho}\right)$ from the $I\overline{I}$-valley approach via [8]

$$\frac{{\rm d}\,n_{I\overline{I}}}{{\rm d}^4x\,{\rm d}^4R}_{|{\rm UKQCD}} \overset{?}{\simeq} \int\limits_0^\infty d\,\rho \int\limits_0^\infty d\,\overline{\rho}\, D(\rho)\, D(\overline{\rho}) \int d\,U\, {\rm e}^{-\frac{4\pi}{\alpha_s}\,\Omega\left(U, \frac{R^2}{\rho\overline{\rho}}, \frac{\overline{\rho}}{\rho}\right)}, \tag{6}$$

and the lattice measurements of $D(\rho)$.

From Fig. 3, I-perturbation theory appears to be quantitatively valid for

$$\left.\begin{aligned} \rho\cdot\Lambda^{(n_f=0)}_{\overline{\rm MS}} \quad &\lesssim \quad 0.42 \\ R/\rho \quad &\gtrsim \quad 1.05 \end{aligned}\right\} \overset{\textbf{saddle point}}{\Rightarrow} \left\{\begin{aligned} Q^\prime/\Lambda^{(n_f)}_{\overline{\rm MS}} \quad &\gtrsim \quad 30.8, \\ x^\prime \quad &\gtrsim \quad 0.35, \end{aligned}\right. \tag{7}$$

Beyond providing a quantitative estimate for the "fiducial" momentum space region in DIS, the good, parameter-free agreement of the lattice data with I-perturbation theory is very interesting in its own right. Uncertainties associated with the inequalities (7) are studied in detail in Ref. [12].

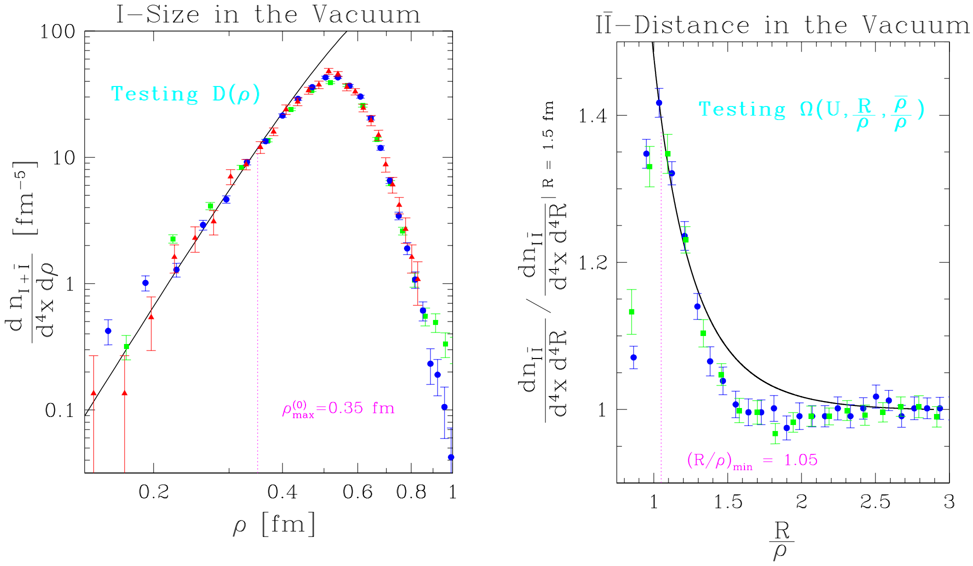


Fig. 3: Illustration of the agreement of recent high-quality lattice data [13, 8] for the instanton-size distribution (left) and the normalized $I\overline{I}$-distance distribution (right) with the predictions from instanton-perturbation theory [8] for $\rho \lesssim 0.35$ fm and $R/\rho \gtrsim 1.05$, respectively. $\alpha_{\overline{\rm MS}}^{3-\rm loop}$ with $\Lambda_{\overline{\rm MS}}^{(n_f=0)}$ from the ALPHA collaboration [29] was used.

2.3 Characteristic final-state signature

The qualitative origin of the characteristic final-state signature of I-induced events is intuitively explained and illustrated in Fig. 4. An indispensable tool for a quantitative investigation of the characteristic final-state signature and notably for actual experimental searches of I-induced events at HERA is our Monte-Carlo generator package QCDINS [10].

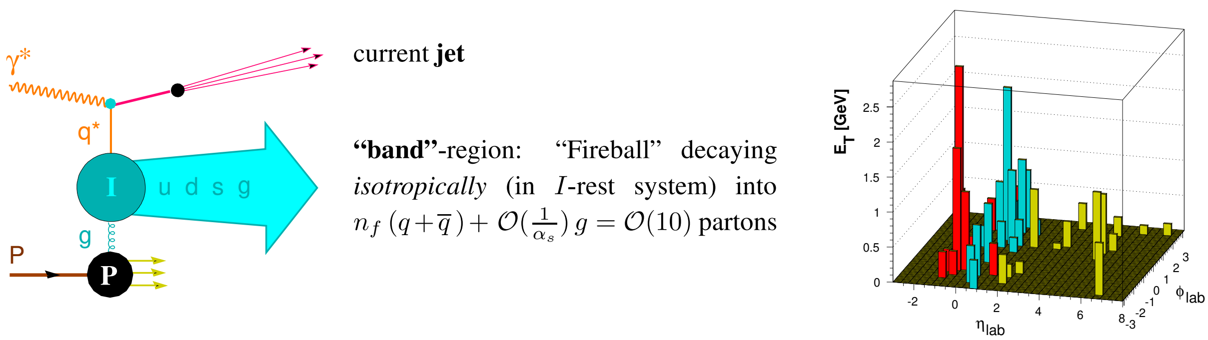


Fig. 4: Characteristic signature of I-induced events: *One* (current) *jet* along with a densely filled *band* of hadrons in the (η, ϕ) plane. Each event has large hadron multiplicity, large total E_t, u-d-s flavor democracy with 1 $s\overline{s}$-pair/event leading to $K's, \Lambda's \ldots$. An event from our QCDINS [10] generator (right) illustrates these features.

2.4 Status of searches at HERA

The results of dedicated searches for instanton-induced events by the H1 and ZEUS collaborations [15, 16], based on our theoretical work, have been finalized meanwhile. The H1 analysis was based on $\int \mathcal{L}dt \approx 21$ pb^{-1}, while ZEUS used $\int \mathcal{L}dt \approx 38$ pb^{-1}, with somewhat differing kinematical cuts. Since the upgraded HERA II machine is now performing very well, forthcoming searches based on a several times higher luminosity might turn out most interesting. Let me briefly summarize the present status from a theorist's perspective.

Fiducial Cuts		H1	ZEUS
Q^2	$\gtrsim$ 113 GeV^2 ?	**no**	**yes**
Q'^2	$\gtrsim$ 113 GeV^2 ?	**yes**	**yes**
x'	$\gtrsim$ 0.35 ?	**no**	**no**

Table 1: Comparison of implemented fiducial cuts that are required in principle to warrant the validity of I-perturbation theory.

While H1 indeed observed a statistically significant excess of events with instanton-like topology and in good agreement with the theoretical predictions, *physical* significance could not be claimed, due to remaining uncertainties in the standard DIS (sDIS) background simulation. The ZEUS collaboration obtained a conservative, background-independent upper limit on the instanton-induced HERA cross section of $26\ \mathrm{pb}@95\%\ \mathrm{CL}$, to be compared to our theoretical prediction of 8.9 pb for the given cuts. In both experiments it was demonstrated that a decisive experimental test of the theoretical predictions based on I-perturbation theory is well within reach in the near-future. In view of the present situation and the interesting prospects for HERA II, let me proceed with a number of comments.

A first important task consists in reconstructing the instanton-subprocess variables (Q'^2, x') from Eq. (2) and in implementing the theoretically required fiducial cuts (cf. Eq. (7)). The actual status is displayed in Table 1 for comparison. The implications of the lacking x'-cut both in the H1 and ZEUS data are presumably not too serious, since QCDINS –*with* its default x'-cut– models to some extent the sharp suppression of I-effects, apparent in the lattice data (cf. Fig. 3 (right)) for $R/\rho \lesssim 1.0-1.05$, i.e. $x' \lesssim 0.3-0.35$. Yet, this lacking, experimental cut introduces a substantial uncertainty in the predicted magnitude of the I-signal that hopefully may be eliminated soon. The lacking Q^2-cut in the H1 data is potentially more serious. As a brief reminder [6, 10], this cut assures in particular the dominance of "planar" handbag-type graphs in $\sigma^{(I)}_{\mathrm{HERA}}$ and all final-state observables. Because of computational complications, the non-planar contributions are *not* implemented in the QCDINS event generator, corresponding to unreliable QCDINS results for small Q^2.

The main remaining challenge resides in the fairly large sDIS background uncertainties. The essential reason is that the existing Monte Carlo generators have been typically designed and tested for kinematical regions different from where the instanton signal is expected! Although the residual problematics is not primarily related to lacking statistics, the near-future availability of many more events will allow to strengthen the cuts and thus hopefully to increase the gap between signal and background. A common search strategy consists in producing I-enriched data samples by cutting on several discriminating observables, each one being sensitive to *different* basic instanton characteristics. An optimized set may be found according to the highest possible

$$\text{instanton separation power} = \frac{\epsilon_I}{\epsilon_{\mathrm{sDIS}}}, \tag{8}$$

in terms of the sDIS and instanton efficiencies, with $\epsilon_I \gtrsim 5-10\%$. Substantial enhancements of the instanton sensitivity were obtained, by means of various multivariate discrimination methods, involving only a single cut on a suitable discriminant variable. In case of ZEUS, cuts on the Fisher discriminant have been used to obtain instanton-enhanced subsamples.

Let me summarize the results obtained so far in form of a theorist's "unified plot" of the H1 and ZEUS "excess" versus the I-separation power. Any visible correlation of a rising *experimental* "excess" with the (Monte-Carlo) *theoretical* I-separation power in Fig. 5 would be an intriguing first signature for a signal. The behaviour seen from the end of the ZEUS range into the H1 domain, might indeed suggest some increase of the excess towards rising I-sparation power. The comparatively low I-separation power of the ZEUS data (and thus perhaps also their *negative* excess?) is mainly due to the implementation of the fiducial cut in Q^2 that is lacking in case of H1.

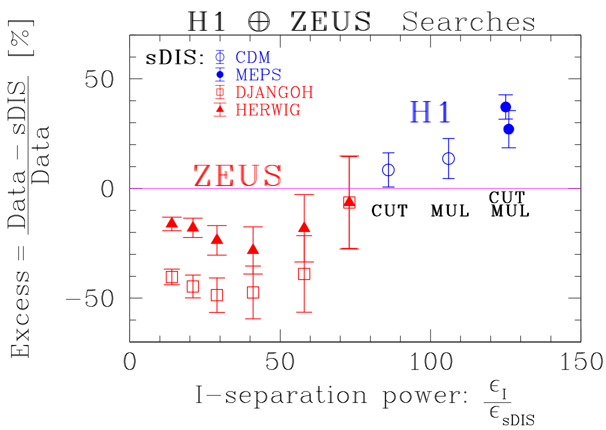


Fig. 5: A theorist's "unified plot" of the H1 and ZEUS "excess" versus I-separation power. The H1 and ZEUS data are seen to join smoothly. A first sign of a rising excess towards higher separation powers might be suspected.

3 Study of the discovery potential at the LHC

Given our extended experience with instanton physics both theoretically and experimentally at DESY, it is natural to ask about the discovery potential for instanton-induced processes at the forthcoming LHC. Indeed, a respective project has been set up around a theoretical PhD Thesis [17], but is still in a relatively early stage.

3.1 Outline of the project

We attempt to do a broad study, focussing both on theoretical and phenomenological issues. Let me just enumerate some interesting aspects that differ essentially from the familiar situation for spacelike hard scattering in DIS at HERA.

Theoretically: The first and foremost task is to identify and calculate the leading I-subprocess at the LHC within I-perturbation theory. Unlike HERA (Fig. 2 (left)), one starts from a $g\,g$-initial state at the LHC. Hence, the rate will be enhanced by a factor $\propto \frac{1}{\alpha_{\rm e.m.}\,\alpha_s}$ compared to $\gamma^*\,g$ scattering at HERA. Then, the next crucial question is how to enforce some parton virtuality in the respective instanton-induced $g\,g$-subprocess, such as to retain the applicability of I-perturbation theory.

An interesting possibility we are exploring is to enter the required virtuality through the *final state* in case of the LHC! One may consider the fragmentation of one or even two outgoing quarks from the $g\,g$-initiated I-instanton subprocess into a *large* $E_\perp$ photon or W-boson and other particles. The requirement of large $E_\perp$ then enforces a *timelike* virtuality onto the outgoing parent quark.

Experimentally: Crucial criteria will be a good signature paired with the lowest possible background, as well as a good trigger. At the experimental front we forsee the collaboration of T. Carli/CERN, who will be able to merge his actual knowledge of the LHC with many years of experience from searches for instantons at HERA. After the theoretical calculations are under control, the next task is to adapt our QCDINS event generator to the LHC, to work out characteristic event signatures, optimal observables, fiducial cuts etc.

4 Instanton-driven saturation at small x

One of the most important observations from HERA is the strong rise of the gluon distribution at small Bjorken-x [30]. On the one hand, this rise is predicted by the DGLAP evolution equations [31] at high Q^2 and thus supports QCD [32]. On the other hand, an undamped rise will eventually violate unitarity. The reason for the latter problem is known to be buried in the linear nature of the DGLAP- and the BFKL-equations [33]: For decreasing Bjorken-x, the number of partons in the proton rises, while their

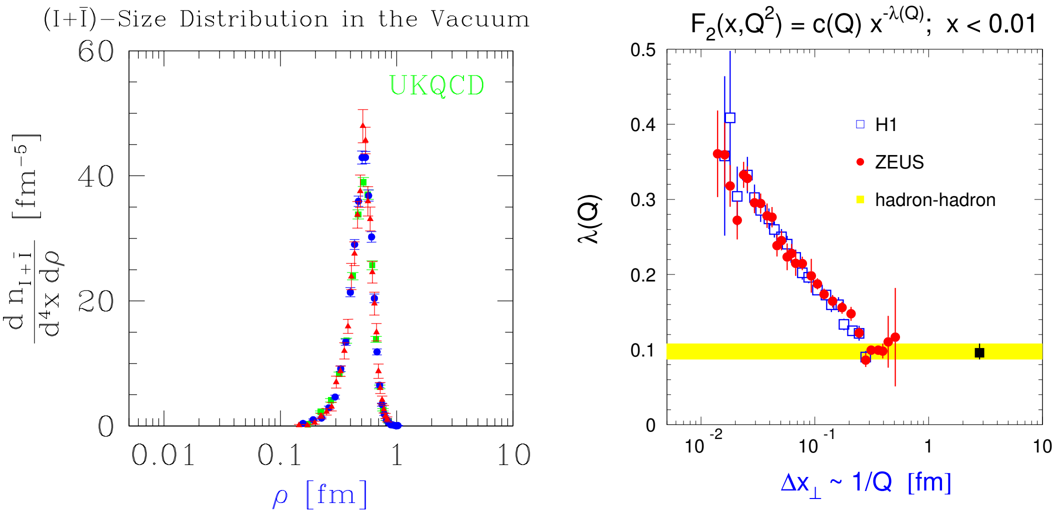


Fig. 6: The I-size scale $\langle\rho\rangle$ from lattice data [8, 13] (left) coincides surprisingly well with the transverse resolution $\Delta x_\perp \sim 1/Q$, where the small-x rise of the structure function $F_2(x, Q^2)$ *abruptly* starts to *increase* with falling $\Delta x_\perp$ (right).

effective size $\sim 1/Q$ increases with decreasing Q^2. At some characteristic scale $Q^2 \approx Q_s^2(x)$, the gluons in the proton start to overlap and so the linear approximation is no longer applicable; non-linear corrections to the linear evolution equations [34] arise and become significant, potentially taming the growth of the gluon distribution towards a "saturating" behaviour.

From a theoretical perspective, eP-scattering at small Bjorken-x and decreasing Q^2 uncovers a novel regime of QCD, where the coupling α_s is (still) small, but the parton densities are so large that conventional perturbation theory ceases to be applicable, eventually. Much interest has recently been generated through association of the saturation phenomenon with a multiparticle quantum state of high occupation numbers, the "Colour Glass Condensate" that correspondingly, can be viewed [22] as a strong *classical* colour field $\propto 1/\sqrt{\alpha_s}$.

4.1 Why instantons?

Being extended non-perturbative fluctuations of the gluon field, instantons come to mind naturally in the context of saturation, since

- classical *non-perturbative* colour fields are physically appropriate in this regime; I-interactions always involve many non-perturbative gluons with multiplicity $\langle n_g\rangle \propto \frac{1}{\alpha_s}$!
- the functional form of the instanton gauge field is explicitely known and its strength is $A_\mu^{(I)} \propto \frac{1}{\sqrt{\alpha_s}}$ as needed;
- an identification of the "Colour Glass Condensate" with the QCD-sphaleron state appears very suggestive [20, 21] (cf. below and Sec 4.4).
- At high energies ($x \to 0$), larger I-sizes ($\rho \gtrsim 0.35$ fm) are probed! Unlike DIS, now the sharply defined average I-size $\langle\rho\rangle \approx 0.5$ fm (known from lattice simulations [13]) comes into play and becomes a relevant and conspicuous length scale in this regime (cf. Fig. 6 (left)).
- An intriguing observation is that the I-size scale $\langle\rho\rangle$ coincides surprisingly well with the transverse resolution $\Delta x_\perp \sim 1/Q$, where the small-x rise of the structure function $F_2(x, Q^2)$ *abruptly* starts to *increase* with falling $\Delta x_\perp$! This striking feature[2] is illustrated in Fig. 6 (right), with the power $\lambda(Q)$ being defined via the ansatz $F_2(x, Q^2) = c(Q)\, x^{-\lambda(Q)}$ at small x. A suggestive interpretation is that instantons are getting resolved for $\Delta x_\perp \lesssim \langle\rho\rangle$.

[2] I wish to thank A. Levy for the experimental data in Fig. 6 (right)

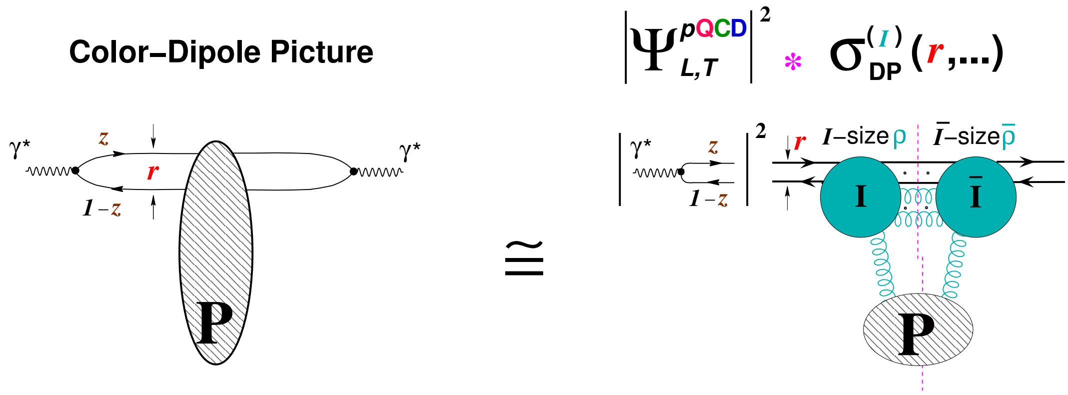


Fig. 7: Illustration of the color dipole picture, its associated variables, the factorization property and the structure of the dipole cross section in an instanton approach.

- We know already from I-perturbation theory that the instanton contribution tends to strongly increase towards the softer regime [5, 7, 10]. The mechanism for the decreasing instanton suppression with increasing energy is known since a long time [36, 35]: Feeding increasing energy into the scattering process makes the picture shift from one of tunneling between adjacent vacua ($E \approx 0$) to that of the actual creation of the sphaleron-like, coherent multi-gluon configuration [23] on top of the potential barrier of height [5, 37] $E = m_{\rm sph} \propto \frac{1}{\alpha_s \rho_{\rm eff.}}$.

4.2 From instanton-perturbation theory to saturation

The investigation of saturation becomes most transparent in the familiar colour-dipole picture [38] (cf. Fig. 7 (left)), notably if analyzed in the so-called dipole frame [39]. In this frame, most of the energy is still carried by the hadron, but the virtual photon is sufficiently energetic, to dissociate before scattering into a $q\overline{q}$-pair (a *colour dipole*), which then scatters off the hadron. Since the latter is Lorentz-contracted, the dipole sees it as a colour source of transverse extent, living (essentially) on the light cone. This colour field is created by the constituents of the well developed hadron wave function and – in view of its high intensity, i.e. large occupation numbers – can be considered as classical. Its strength near saturation is $\mathcal{O}(1/\sqrt{\alpha_s})$. At high energies, the lifetime of the $q\overline{q}$-dipole is much larger than the interaction time between this $q\overline{q}$-pair and the hadron and hence, at small $x_{\rm Bj}$, this gives rise to the familiar factorized expression of the inclusive photon-proton cross sections,

$$\sigma_{L,T}(x_{\rm Bj}, Q^2) = \int_0^1 dz \int d^2 r\, |\Psi_{L,T}(z,r)|^2\, \sigma_{\rm DP}(r,\ldots). \tag{9}$$

Here, $|\Psi_{L,T}(z,r)|^2$ denotes the modulus squared of the (light-cone) wave function of the virtual photon, calculable in pQCD, and $\sigma_{\rm DP}(r,\ldots)$ is the $q\overline{q}$-dipole - nucleon cross section. The variables in Eq. (9) are the transverse $(q\overline{q})$-size $\mathbf{r}$ and the photon's longitudinal momentum fraction z carried by the quark. The dipole cross section is expected to include in general the main non-perturbative contributions. For small r, one finds within pQCD [38, 40] that $\sigma_{\rm DP}$ vanishes with the area πr^2 of the $q\overline{q}$-dipole. Besides this phenomenon of "colour transparency" for small $r = |r|$, the dipole cross section is expected to saturate towards a constant, once the $q\overline{q}$-separation r exceeds a certain saturation scale r_s (cf. Fig. 7 (right)). While there is no direct proof of the saturation phenomenon, successful models incorporating saturation do exist [41] and describe the data efficiently.

Let us outline more precisely our underlying strategy:

- We start from the large Q^2 regime and appropriate cuts such that I-perturbation theory is strictly valid. The corresponding, known results on I-induced DIS processes [6] are then transformed into the colour-dipole picture.

- The guiding question is: Can background instantons of size $\sim \langle\rho\rangle$ give rise to a saturating, geometrical form for the dipole cross section,

$$\sigma_{\rm DP}^{(I)}(r,\ldots) \overset{r \gtrsim \langle\rho\rangle}{\sim} \pi\langle\rho\rangle^2. \tag{10}$$

- With the crucial help of lattice results, the $q\bar{q}$-dipole size r is next carefully increased towards hadronic dimensions. Thanks to the lattice input, IR divergencies are removed and the original cuts are no longer necessary.

4.3 The simplest process: $\gamma^* + g \overset{(I)}{\rightarrow} q_{\rm R} + \overline{q}_{\rm R}$

Let us briefly consider first the simplest I-induced process, $\gamma^*\, g \Rightarrow q_{\rm R}\overline{q}_{\rm R}$, with one flavour and no final-state gluons. More details may be found in Ref. [20]. Already this simplest case illustrates transparently that in the presence of a background instanton, the dipole cross section indeed saturates with a saturation scale of the order of the average I-size $\langle\rho\rangle$.

We start by recalling the results for the total $\gamma^* N$ cross section within I-perturbation theory from Ref. [6],

$$\sigma_{L,T}(x_{\rm Bj},Q^2) \;=\; \int\limits_{x_{\rm Bj}}^{1} \frac{dx}{x}\left(\frac{x_{\rm Bj}}{x}\right) G\left(\frac{x_{\rm Bj}}{x},\mu^2\right)\int dt \frac{d\hat{\sigma}^{\gamma^* g}_{L,T}(x,t,Q^2)}{dt}; \tag{11}$$

$$\frac{d\hat{\sigma}^{\gamma^* g}_{L}}{dt} \;=\; \frac{\pi^7}{2}\frac{e_q^2}{Q^2}\frac{\alpha_{\rm em}}{\alpha_{\rm s}}\left[x(1-x)\sqrt{tu}\,\frac{\mathcal{R}(\sqrt{-t})-\mathcal{R}(Q)}{t+Q^2} - (t\leftrightarrow u)\right]^2, \tag{12}$$

with a similar expression for $d\hat{\sigma}^{\gamma^*\,g}_{T}/dt$. Here, $G\left(x_{\rm Bj},\mu^2\right)$ denotes the gluon density and L,T refers to longitudinal and transverse photons, respectively.

Note that Eqs. (11), (12) involve the resolution dependent length scale

$$\mathcal{R}(\mathcal{Q}) = \int_0^\infty d\rho\; D(\rho)\rho^5(\mathcal{Q}\rho){\rm K}_1(\mathcal{Q}\rho). \tag{13}$$

which is of key importance for continuing towards $\mathcal{Q}\langle\rho\rangle \Rightarrow 0$! For sufficiently large $\mathcal{Q}\langle\rho\rangle$, the crucial factor $(\mathcal{Q}\rho)\,K_1(\mathcal{Q}\rho) \sim e^{-\mathcal{Q}\rho}$ in Eq.(13) exponentially suppresses large size instantons and I-perturbation theory holds, as shown first in Ref. [6]. In our continuation task towards smaller $\mathcal{Q}\langle\rho\rangle$, crucial lattice information enters. We recall that the I-size distribution $D_{\rm lattice}(\rho)$, as *measured* on the lattice [13, 8, 12], is strongly peaked around an average I-size $\langle\rho\rangle \approx 0.5$ fm, while being in excellent agreement with I-perturbation theory for $\rho \lesssim 0.35$ fm (cf. Sect. 2.2 and Fig. 3(left)). Our strategy is thus to generally identify $D(\rho) = D_{\rm lattice}(\rho)$ in Eq.(13), whence

$$\mathcal{R}(0) = \int_0^\infty d\rho\; D_{\rm lattice}(\rho)\rho^5 \approx 0.3 \text{ fm} \tag{14}$$

becomes finite and a $\mathcal{Q}^2$ cut is no longer necessary.

By means of an appropriate change of variables and a subsequent $2d$-Fourier transformation, Eqs. (11), (12) may indeed be cast [20] into a colour-dipole form (9), e.g. (with $\hat{Q} = \sqrt{z\,(1-z)}\,Q$)

$$\left(|\Psi_L|^2\,\sigma_{\rm DP}\right)^{(I)} \approx |\;\Psi_L^{\rm pQCD}(z,r)\;|^2\;\frac{1}{\alpha_{\rm s}}\,x_{\rm Bj}\,G(x_{\rm Bj},\mu^2)\,\frac{\pi^8}{12} \tag{15}$$

$$\times\left\{\int_0^\infty d\rho D(\rho)\,\rho^5\left(\frac{-\frac{d}{dr^2}\left(2r^2\frac{{\rm K}_1(\hat{Q}\sqrt{r^2+\rho^2/z})}{\hat{Q}\sqrt{r^2+\rho^2/z}}\right)}{{\rm K}_0(\hat{Q}r)} - (z\leftrightarrow 1-z)\right)\right\}^2.$$

The strong peaking of $D_{\rm lattice}(\rho)$ around $\rho \approx \langle\rho\rangle$, implies

$$\left(|\Psi_{L,T}|^2\sigma_{\rm DP}\right)^{(I)} \Rightarrow \begin{cases} \mathcal{O}(1) \text{ but exponentially small}; & r \to 0, \\ |\,\Psi^{\rm pQCD}_{L,T}\,|^2\ \frac{1}{\alpha_{\rm s}}\, x_{\rm Bj}\, G(x_{\rm Bj},\mu^2)\, \frac{\pi^8}{12}\, \mathcal{R}(0)^2; & \frac{r}{\langle\rho\rangle} \gtrsim 1. \end{cases} \tag{16}$$

Hence, the association of the intrinsic instanton scale $\langle\rho\rangle$ with the saturation scale r_s becomes apparent from Eqs. (15), (16): $\sigma^{(I)}_{\rm DP}(r,\ldots)$ rises strongly as function of r around $r_s \approx \langle\rho\rangle$, and indeed *saturates* for $r/\langle\rho\rangle > 1$ towards a *constant geometrical limit*, proportional to the area $\pi\,\mathcal{R}(0)^2 = \pi\left(\int_0^\infty d\rho\, D_{\rm lattice}(\rho)\,\rho^5\right)^2$, subtended by the instanton. Since $\mathcal{R}(0)$ would be divergent within I-perturbation theory, the information about $D(\rho)$ from the lattice (Fig. 6 (left)) is crucial for the finiteness of the result.

4.4 Identification of the color glass condensate with the QCD-sphaleron state

Next, let us consider the realistic process, $\gamma^* + g \stackrel{(I)}{\rightarrow} n_f\,(q_{\rm R} + \overline{q}_{\rm R})$ + gluons. On the one hand, the inclusion of final-state gluons and $n_f > 1$ causes a significant complication. On the other hand, it is due to the effect of those gluons that the identification of the QCD-sphaleron state with the colour glass condensate has emerged [20, 21], while the qualitative "saturation" features remain unchanged. Most of the I-dynamics resides in the I-induced $q^*\,g$-subprocess with an incoming off-mass-shell quark q^* originating from photon dissociation. The important kinematical variables are the I-subprocess energy $E = \sqrt{(q'+p)^2}$ and the quark virtuality $Q'^2 = -q'^2$, with the gluon 4-momentum denoted by p_μ.

It is most convenient to account for the final-state gluons by means of the $I\bar{I}$-valley method [25] (cf. also Sect. 2.1). It allows to achieve via the optical theorem, an elegant summation over the gluons. The result leads to an exponentiation of the final-state gluon effects, residing entirely in the $I\bar{I}$-valley interaction $-1 \le \Omega^{I\bar{I}}_{\rm valley}(\frac{R^2}{\rho\bar{\rho}} + \frac{\rho}{\bar{\rho}} + \frac{\bar{\rho}}{\rho}; U) \le 0$, introduced in Eq. (3) of Sect. 2.1. Due to the new gluon degrees of freedom, the additional integrations over the $I\bar{I}$-distance R_μ appear (cf. Fig. 2 (right)), while the matrix U characterizes the relative $I\bar{I}$ orientation in colour space. We remember from Sect. 2.1 that the functional form of $\Omega^{I\bar{I}}_{\rm valley}$ is analytically known [26, 27] (formally) for *all* values of $R^2/(\rho\bar{\rho})$. Our strategy here is identical to the one for the "simplest process" above: Starting point is the $\gamma^* N$ cross section, this time obtained by means of the $I\bar{I}$-valley method [7]. The next step is a variable and Fourier transformation into the colour-dipole picture. The dipole cross section $\tilde{\sigma}^{(I),\rm gluons}_{\rm DP}(\mathbf{l}^2, x_{\rm Bj},\ldots)$ before the final 2d-Fourier transformation of the quark transverse momentum $\mathbf{l}$ to the conjugate dipole size $\mathbf{r}$, arises simply as an energy integral over the I-induced total q^*g cross section in Eq. (3) from Ref. [7],

$$\tilde{\sigma}^{(I),\rm gluons}_{\rm DP} \approx \frac{x_{\rm Bj}}{2}\, G(x_{\rm Bj},\mu^2) \int_0^{E_{\rm max}} \frac{d\,E}{E}\left[\frac{E^4}{(E^2+Q'^2)\,Q'^2}\,\sigma^{(I)}_{q^*g}\left(E,\mathbf{l}^2,\ldots\right)\right], \tag{17}$$

involving in turn integrations over the $I\bar{I}$-collective coordinates $\rho, \bar{\rho}, U$ and R_μ.

In the softer regime of interest for saturation, we again substitute $D(\rho) = D_{\rm lattice}(\rho)$, which enforces $\rho \approx \bar{\rho} \approx \langle\rho\rangle$ in the respective $\rho,\bar{\rho}$-integrals, while the integral over the $I\bar{I}$-distance R is dominated by a *saddle point*,

$$\frac{R}{\langle\rho\rangle} \approx \text{function}\left(\frac{E}{m_{\rm sph}}\right); \quad m_{\rm sph} \approx \frac{3\pi}{4}\frac{1}{\alpha_{\rm s}\,\langle\rho\rangle} = \mathcal{O}(\text{ few GeV }). \tag{18}$$

At this point, the mass $m_{\rm sph}$ of the QCD-sphaleron [5, 37], i.e the barrier height separating neighbouring topologically inequivalent vacua, enters as the scale for the energy E. The saddle-point dominance implies a one-to-one relation,

$$\frac{R}{\langle\rho\rangle} \Leftrightarrow \frac{E}{m_{\rm sph}}; \quad \text{with } R = \langle\rho\rangle \Leftrightarrow E \approx m_{\rm sph}. \tag{19}$$

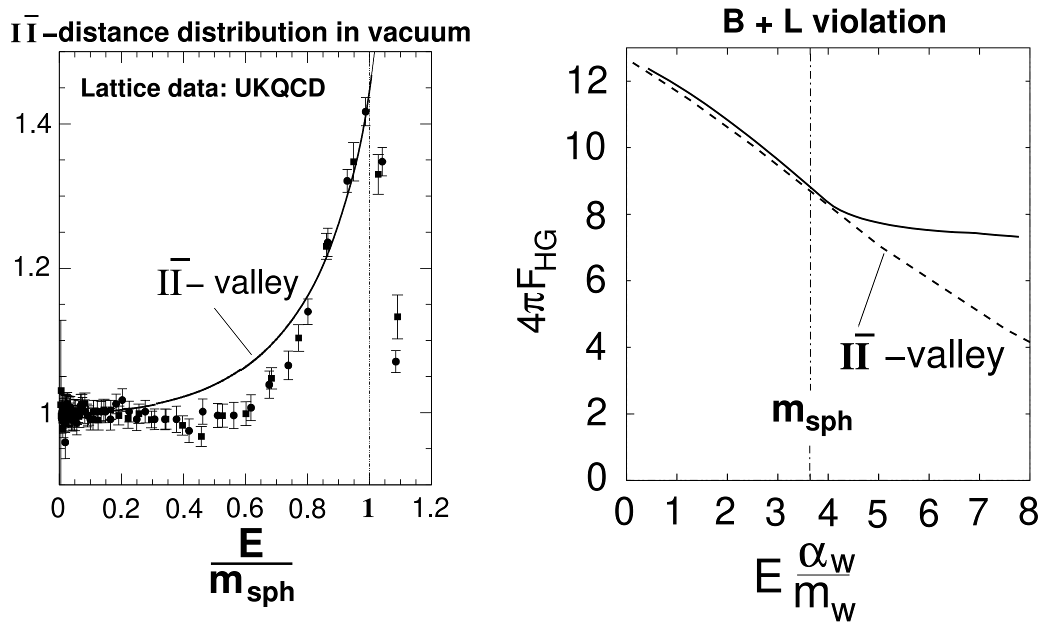


Fig. 8: (left) The UKQCD lattice data [13, 8] of the (normalized) $I\bar{I}$-distance distribution together with the corresponding $I\bar{I}$-valley prediction [20] from Fig. 3 (right) are re-displayed versus energy in units of the QCD sphaleron mass $m_{\rm sph}$.This illustrates the validity of the valley approach right until the sphaleron peak! (right) The same trend for electroweak $B + L$ -violation is apparent from an independent numerical simulation of the suppression exponent for two-particle collisions ('Holy Grail' function) $F_{\rm HG}(E)$ [42, 43]

Our continuation to the saturation regime now involves crucial lattice information about $\Omega^{I\bar{I}}$. The relevant lattice observable is the distribution of the $I\bar{I}$-distance [8, 20] R, providing information on $\left\langle \exp[-\frac{4\pi}{\alpha_{\rm s}}\Omega^{I\bar{I}}]\right\rangle_{U,\rho,\bar{\rho}}$ in euclidean space (cf. Fig. 3 (right)). Due to the crucial saddle-point relation Eqs. (18, 19), we may replace the original variable $R/\langle\rho\rangle$ by $E/m_{\rm sph}$. A comparison of the respective $I\bar{I}$-valley predictions with the UKQCD lattice data [13, 8, 20] versus $E/m_{\rm sph}$ is displayed in Fig. 8 (left). It reveals the important result that the $I\bar{I}$-valley approximation is quite reliable up to $E \approx m_{\rm sph}$. Beyond this point a marked disagreement rapidly develops: While the lattice data show a *sharp peak* at $E \approx m_{\rm sph}$, the valley prediction continues to rise indefinitely for $E \gtrsim m_{\rm sph}$! It is remarkable that an extensive recent and completely independent semiclassical numerical simulation [42] shows precisely the same trend for electroweak $B + L$-violation, as displayed in Fig. 8 (right).

It is again at hand to identify $\Omega^{I\bar{I}} = \Omega^{I\bar{I}}_{\rm lattice}$ for $E \gtrsim m_{\rm sph}$. Then the integral over the I-subprocess energy spectrum (17) in the dipole cross section appears to be dominated by the sphaleron configuration at $E \approx m_{\rm sph}$. The feature of saturation analogously to the "simplest process" in Sect. 4.3 then implies the announced identification of the colour glass condensate with the QCD-sphaleron state.

5 Conclusions

As non-perturbative, topological fluctuations of the gluon fields, *instantons* are a basic aspect of QCD. Hence their experimental discovery through hard instanton-induced processes would be of fundamental significance. A first purpose of this overview was to present a summary of our systematic theoretical and phenomenological investigations of the discovery potential in DIS at HERA, based on a calculable rate of measurable range and a characteristic "fireball"-like event signature. In a summary of the present status of experimental searches by H1 and ZEUS, the typical remaining challenges were particularly emphasized. In view of the good performance of the upgraded HERA II machine, one may expect further possibly decisive instanton search results in the near future.The existing H1 and ZEUS results have demonstrated already that the required sensitivity according to our theoretical predictions is within reach. Looking ahead, I have briefly discussed an ongoing project concerning a broad investigation of the discovery potential of instanton processes at the LHC. A final part of this review was devoted to our work on small-x saturation from an instanton perspective. After summarizing the considerable motivation for the relevance of instantons in this regime, the emerging intuitive, geometrical picture was illustrated with the simplest example, where indeed, saturation does occur. The form of the dipole cross section depends on the relation of two competing areas: the area $\pi\, r^2$, subtended by the $\bar{q}q$-dipole, and the area $\pi\,\langle\rho\rangle^2$ associated with the average size, $\langle\rho\rangle \approx 0.5$ fm, of the background instanton. For $r/\langle\rho\rangle \ll 1$, the dipole cross section is dominated by the dipole area, corresponding to 'color transparency'. For $r/\langle\rho\rangle \gtrsim 1$ it saturates towards a constant proportional to the background instanton area. Correspondingly, the average I-size scale $\langle\rho\rangle$ is associated with the saturation scale. A further central and intriguing result concerned the identification of the Color Glass Condensate with the QCD-sphaleron state. Throughout, the non-perturbative information from lattice simulations was instrumental.

References

[1] A. Belavin, A. Polyakov, A. Schwarz and Yu. Tyupkin, Phys. Lett. **B 59**, 85 (1975).

[2] G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. **D 14**, 3432 (1976); Phys. Rev. **D 18**, 2199 (1978) (Erratum).

[3] T. Schäfer and E. Shuryak, Rev. Mod. Phys. **70**, 323 (1998) and references cited therein.

[4] M.-C. Chu et al.,(**Phys. Rev. D**,) (49)60391994.

[5] A. Ringwald and F. Schrempp, *Proc. Quarks '94*, ed. D.Yu. Grigoriev et al. (Singapore: World Scientific) p. 170.

[6] S. Moch, A. Ringwald and F. Schrempp, Nucl. Phys. **B 507**, 134 (1997).

[7] A. Ringwald and F. Schrempp, Phys. Lett. **B 438**, 217 (1998).

[8] A. Ringwald and F. Schrempp, Phys. Lett. **B 459**, 249 (1999).

[9] I. Balitsky and V. Braun, Phys. Lett. **B 314**, 237 (1993).

[10] A. Ringwald and F. Schrempp, Comput. Phys. Commun. **132**, 267 (2000).

[11] T. Carli, J. Gerigk, A. Ringwald and F. Schrempp, *Proc. Monte Carlo Generators for HERA Physics (DESY-PROC-1999-02)* ed. A.T. Doyle et al. (Hamburg: DESY) p. 329.

[12] A. Ringwald and F. Schrempp, Phys. Lett. **B 503**, 331 (2001).

[13] D.A. Smith and M.J. Teper (UKQCD Collab.), Phys. Rev. **D 58**, 014505 (1998).

[14] D. Ostrovsky and E. Shuryak, Phys. Rev. **D 71**, 014037 (2005).

[15] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 25**, 495 (2002).

[16] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 34**, 255 (2004).

[17] M. Petermann, PhD Thesis in preparation; M. Petermann and F. Schrempp, in preparation.

[18] F. Schrempp, J. Phys. **G 28**, 915 (2002).

[19] F. Schrempp and A. Utermann, Acta Phys. Polon. **B 33**, 3633 (2002).

[20] F. Schrempp and A. Utermann, Phys. Lett. **B 543**, 197 (2002).

[21] F. Schrempp and A. Utermann, *Proc. Strong and Electroweak Matter 2002*, Heidelberg, Oct. 2002, ed. M.G. Schmidt, p. 477 [arXiv:hep-ph/0301177].

[22] E. Iancu, A. Leonidov and L. D. McLerran, Nucl. Phys. **A 692**, 583 (2001);
E. Ferreiro *et al.*, Nucl. Phys. **A 703**, 489 (2002).

[23] F. R. Klinkhamer and N. S. Manton, Phys. Rev. **D 30**, 2212 (1984).

[24] E. V. Shuryak and I. Zahed, Phys. Rev. **D 69**, 014011 (2004).

[25] A. Yung, Nucl. Phys. **B 297**, 47 (1988).

[26] V.V. Khoze and A. Ringwald, Phys. Lett. **B 259**, 106 (1991).

[27] J. Verbaarschot, Nucl. Phys. **B 362**, 33 (1991).

[28] C. Bernard, Phys. Rev. **D 19**, 3013 (1979).

[29] S. Capitani, M. Lüscher, R. Sommer and H. Wittig, Nucl. Phys. **B 544**, 669 (1999).

[30] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 21**, 33 (2001);
ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 21**, 443 (2001).

[31] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972); L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975); G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977); Y.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[32] A. De Rujula et al., Phys. Rev. **D 10**, 1649 (1974).

[33] L.N. Lipatov, Sov. J. Nucl. Phys. **23**, 338 (1976); V.S. Fadin, E.A. Kuraev and L.N. Lipatov, Phys. Lett. **B 60**, 50 (1975), Sov. Phys. JETP **44**, 443 (1976), Sov. Phys. JETP **45**, 199 (1977);
I.I. Balitsky and L.N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[34] L.V. Gribov, E.M. Levin and M.G. Ryskin, Nucl. Phys. **B 188**, 555 (1981); L.V. Gribov, E.M. Levin and M.G. Ryskin, Phys. Rept. **100**, 1 (1983).

[35] D.M. Ostrovsky et al., Phys. Rev. **D 66**, 036004 (2002).

[36] H. Aoyama and H. Goldberg, Phys. Lett. **B 188**, 506 (1987); A. Ringwald, Nucl. Phys. **B 330**, 1 (1990); O. Espinosa, Nucl. Phys. **B 343**, 310 (1990).

[37] D. Diakonov and V. Petrov, Phys. Rev. **D 50**, 266 (1994).

[38] N. Nikolaev and B.G. Zakharov, Z. Phys. **C 49**, 607 (1990); Z. Phys. **C 53**, 331 (1992); A.H. Mueller, Nucl. Phys. **B 415**, 373 (1994).

[39] A.H. Mueller, *Parton Saturation - An Overview*, hep-ph/0111244, in *QCD Perspectives on Hot and Dense Matter*, NATO Science Series, Kluwer, 2002.

[40] F. E. Low, Phys. Rev. **D 12**, 163 (1975);
L. Frankfurt, G.A. Miller and M. Strikman, Phys. Lett. **B 304**, 1 (1993).

[41] K. Golec-Biernat and M. Wüsthoff, Phys. Rev. **D 59**, 014017 (1999); Phys. Rev. **D 60**, 114023 (1999).

[42] F. Bezrukov, D. Levkov, C. Rebbi, V. Rubakov and P. Tinyakov, Phys. Rev. **D 68**, 036005 (2003)

[43] A. Ringwald, Phys. Lett. **B 555**, 227 (2003); arXiv:hep-ph/0302112.

Heavy quark production at HERA and the LHC

Matthew Wing
University College London and DESY

Abstract
Measurements of heavy quark production, particularly from HERA, their theoretical understanding and their relevance for the LHC are reviewed[1]. The status of beauty and charm production is discussed in the context of the different components of the production process: the parton density function of the colliding hadrons; the hard scatter; and the fragmentation of the quarks into hadrons. The theory of QCD at next-to-leading order generally describes well the hadronic structure and the production of heavy quarks although sometimes fails in details which are highlighted. The fragmentation of heavy quarks measured at HERA is consistent with that at LEP and hence supports the notion of universality.

1 Why study heavy quark production?

The measurement of heavy quarks can give insights into many physical phenomena such as: new particles which are expected to decay predominantly to beauty (and charm); precise measurements of electroweak parameters; and, the subject of this paper, a deeper understanding of the strong force of nature. The strong force as described within perturbative Quantum Chromodynamics (QCD) should be able to give a precise description of heavy quark production. This postulate is described and tested here. The measurement of heavy quark production also yields valuable information on the structure of colliding hadrons. The production of a pair of heavy quarks in a generic hadron collision is shown in Fig. 1 where it can be seen that the process is directly sensitive to the gluon content of the hadron. Most information on the structure of a hadron comes from inclusive deep inelastic scattering where the gluon content is determined in the evolution of the QCD equations. Therefore measurement of such a process in Fig. 1 provides complimentary information to that from inclusive measurements.

As well as understanding for its own sake, knowledge of the structure of hadrons will be important at future colliders such as the LHC and International Linear Collider where hadronic photons will have large cross section in both e^+e^- and $\gamma\gamma$ modes. Heavy quarks will be copiously produced at future colliders as a background to the more exotic processes expected. Therefore a precise description of their production properties within QCD will aid in the discovery of physics beyond the Standard Model. An example of this was studied by the ATLAS collaboration using Monte Carlo to simulate the production at the LHC of a $b\bar{b}$ pair along with a supersymmetric Higgs particle (H/A) which subsequently decays to a $b\bar{b}$ pair [1]. For an assumed mass $m_A = 500$ GeV, even requiring four beauty jets, a signal-to-background ratio of only a few percent would be achieved. The irreducible background arises from QCD processes where the dominant processes are gg and gb with a gluon splitting to a $b\bar{b}$ pair. A discovery in this channel would therefore only be possible with precise knowledge these QCD background processes.

2 Theoretical and phenomenological overview

For a generic collision between two hadrons, H_a and H_b, in which a heavy quark pair is produced (see Fig. 1),

$$H_a + H_b \to Q\bar{Q} + X,$$

[1]Since the presentation, some results have been updated; these are used in what follows.

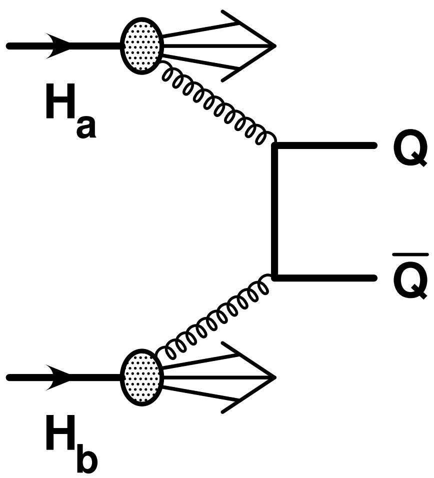


Fig. 1: Example of the production of a heavy quark pair in the collision of two hadrons.

the production cross section, $\sigma(S)$, for such a reaction at a centre-of-mass energy, S, can be written as:

$$\sigma(S) = \sum_{i,j} \int dx_1 \int dx_2 \; \hat{\sigma}_{ij}(x_1 x_2 S, m^2, \mu^2) f_i^{H_a}(x_1, \mu) f_j^{H_b}(x_2, \mu),$$

where the right-hand side is a convolution of the parton densities in the colliding hadrons, $f_i^{H_a}$ and $f_j^{H_b}$, and the short-distance cross section, $\hat{\sigma}_{ij}$. These are evaluated at a renormalisation and factorisation scale, μ, and momentum fractions of the colliding partons, x_1 and x_2. The parton densities are extracted from QCD fits to inclusive deep inelastic scattering and other data. The short-distance cross section is calculable in QCD and is a perturbative expansion in the mass of the heavy quark, m:

$$\hat{\sigma}_{ij}(s, m^2, \mu^2) = \frac{\alpha_s^2(\mu^2)}{m^2} \left[f_{ij}^{(0)}(\rho) + 4\pi\alpha_s(\mu^2) \left[f_{ij}^{(1)}(\rho) + \bar{f}_{ij}^{(1)}(\rho) \log(\mu^2/m^2) \right] + \mathcal{O}(\alpha_s^2) \right], \quad \rho = 4m^2/s.$$

The expansion demonstrates that the larger the mass the faster the convergence. Hence predictions for beauty production should be more accurate than those for charm.

The treatment of the mass of the heavy quark is an important consideration for the implementation of the perturbative formalism in calculations. There are three schemes used: the fixed-order (FO) or “massive” scheme, the resummed to next-to-leading logarithms (NLL), or “massless” scheme and more recently a scheme matching the two, known as FONLL [2]. In the FO scheme, the predictions should be valid for transverse momenta of the order of the mass of the heavy quark. In this scheme, the heavy quarks are not active flavours in the parton distributions of the incoming hadron(s); they are produced in the hard scatter through processes such as $gg \to Q\bar{Q}$ shown in Fig. 1. The resummed scheme is valid for transverse momenta much larger than the heavy quark mass. The heavy quarks are active flavours in the parton distributions of the incoming hadron(s), so can be produced by reactions such as $gQ \to gQ$. The FONLL calculations match the two schemes and are valid for all transverse momenta. The validity of the different calculations is investigated in comparison with data, particularly as a function the energy scale.

The fixed-order calculations used are from Frixione et al. (FMNR) [3] for photoproduction processes and HVQDIS from Harris and Smith [4] for deep inelastic scattering. Resummed calculations are only available for photoproduction at HERA from two groups of authors, Cacciari et al. [5] and Kniehl et al. [6]. The FONLL calculation is also only available in photoproduction. A calculation which is already available for some processes in pp collisions, MC@NLO [7], combines a fixed-order calculation with the parton showering and hadronisation from the HERWIG Monte Carlo generator [8]. Processes at HERA are not yet included, but it is hoped they will be done in the future and thereby provide a new level of detail in comparison with experimental data.

The advantages of a programme such as MC@NLO are its simulation of higher orders and also its sophisticated approach to hadronisation which attempts to describe the whole of the final state. The other programmes produce partons in the final state and fragment the outgoing quark to a hadron usually via the Peterson function [9]. Therefore these calculations may not be able to describe the full hadronic final state of an event. The validity of the fragmentation functions used also needs to be tested; they are usually extracted from fits to e^+e^- data and their applicability to ep or pp needs to be demonstrated. Therefore the fragmentation function should be measured at HERA, and is discussed later, or measurements need to be made at high transverse energy or using jets where the effects of fragmentation are reduced.

Hadron-hadron collisions producing heavy quarks pairs can be simplified to and provide information on: the parton densities and in particular the gluon and heavy quark content of the hadron; the hard scatter and the dynamics of QCD as implemented into programmes; fragmentation or description of the parton to hadron transition. All of these aspects are discussed in this write-up.

3 Information needed by the LHC experiments

The information needed by the LHC which can be provided by the HERA experiments is the following:

- the state of the description of heavy quark production data by theoretical predictions. The production of heavy quarks in the hard scattering process is discussed here in detail. Information on heavy quarks produced in the splitting of a gluon outgoing from the hard sub-process is also important for the LHC, but the information from HERA is currently limited;
- the gluon and heavy quark content of the proton parton density functions;
- details of fragmentation in a hadronic environment;
- the effect of the underlying event in heavy quark processes. This information is limited at HERA but may be studied in the future;
- HERA results can provide general information on event and jet topologies which will be useful for designing algorithms or triggers at the LHC experiments.

The designing of effective triggers for b physics is particularly acute for the LHCb experiment [10]. Large backgrounds are expected although event topologies should be different to the signal b physics. For example minimum bias events will have a smaller track multiplicity and a lower transverse momentum for the highest p_T track. Therefore using Monte Carlo simulation, cuts can be found to be able to reduce the rate of minimum bias whilst triggering efficiently on b events. Such simulations require reliable Monte Carlo simulation of the event topologies of both classes of events.

Measurements of the proton structure function at HERA will constrain the parton densities in a large region of the kinematic plane where B mesons will be produced within the acceptance of the LHCb detector. According to Monte Carlo simulations, these events are produced predominantly with a b quark in the proton. However, this is just a model (PYTHIA [11]) and at NLO some of the events will be summed into the gluon distribution of the proton. Nevertheless, measuring all flavours in the proton at HERA is one of the goals of the experiments and recent results on the beauty contribution to the proton structure function [12] shed some light on the issue.

4 Open beauty production

The production of open beauty and its description by QCD has been of great interest in the last 10–15 years.The difference between the rates observed by the Tevatron experiments [13] and NLO QCD predictions led to a mini crisis with many explanations put forward. Several measurements were performed in different decay channels and then extrapolated to the quark level to facilitate a comparison with QCD and between themselves. The NLO QCD prediction was found to be a factor of 2–3 below the data for all measurements as shown in Fig. 2a. As mentioned, these results were extrapolated to the b-quark level using Monte Carlo models which may or may not give a good estimate of this extrapolation. To facilitate a particular comparison, an extrapolation can be useful, but should always be treated with caution and the procedure clearly stated and values of extrapolation factors given. Initial measurements in terms of measured quantities should also always be given.

The CDF collaboration also published measurements of B meson cross sections. They were also found to be significantly above NLO calculations, but allowed for phenomenological study. Work on the fragmentation function was performed by Cacciari and Nason [14] which in combination with updated parton density functions and the FONLL calculation gave an increased prediction. New measurements at Run II have also been made by the CDF collaboration which probe down to very low transverse momenta. In combination with a measured cross section lower (but consistent) than the Run I data, and the above theoretical improvements, the data and theory now agree very well as shown in Fig. 2b. The programme MC@NLO also gives a good description of the data.

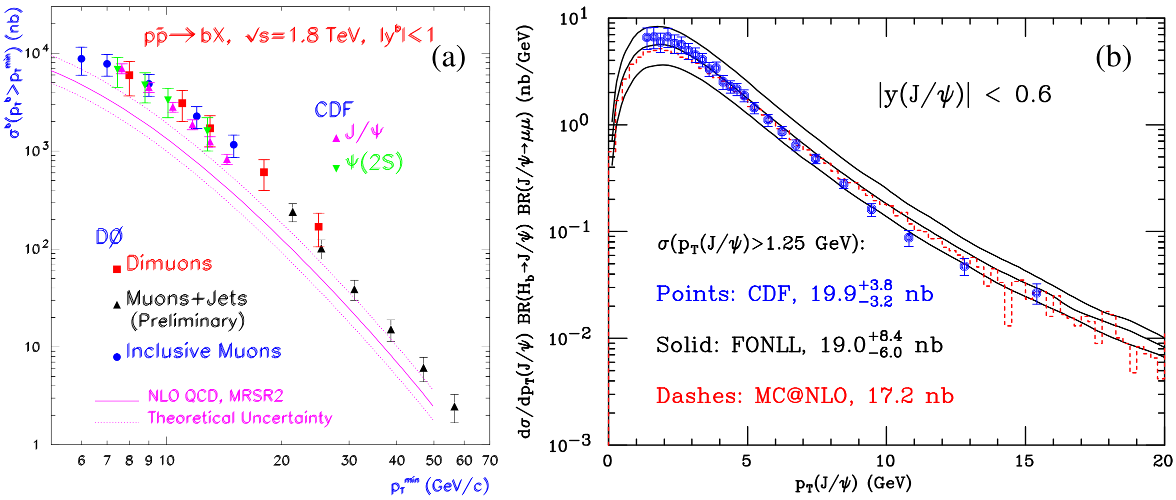


Fig. 2: (a) Tevatron Run I data extrapolated and compared to NLO QCD predictions and (b) Run II data presented in terms of the measured quantities and compared to improved QCD theory.

The first result from HERA [15] also revealed a large discrepancy with NLO QCD predictions. This analysis also presented an extrapolated quantity, whereas later measurements [16, 17, 18] also presented measured quantities. The most recent and precise measurements [17] of beauty production with accompanying jets are shown in Fig. 3 compared with predictions from NLO QCD. The measurements in photoproduction (Fig. 3a) are shown to be very well described by the prediction and the data from the two collaborations also agree well. The H1 data is somewhat higher than that from ZEUS; the difference is concentrated at low p_T^μ where the H1 data is also above the NLO calculation. The measurements in deep inelastic scattering are also generally described by NLO QCD although some differences at forward η^μ and low p_T^μ are observed by both collaborations. However, inclusive measurements which lead to a measurement of the beauty contribution to the proton structure function [12] are well described by QCD (see next Section).

The situation for the QCD description of b production has recently changed significantly. In general, QCD provides a good description of the data with some hints at differences in specific regions. Certainly, there is no longer a difference of a factor of 2–3 independent of p_T. The HERA experiments will produce several new measurements in the next few years of higher precision and covering a larger kinematic region at both low and high p_T and forward η. Allied with expected calculational and phe-

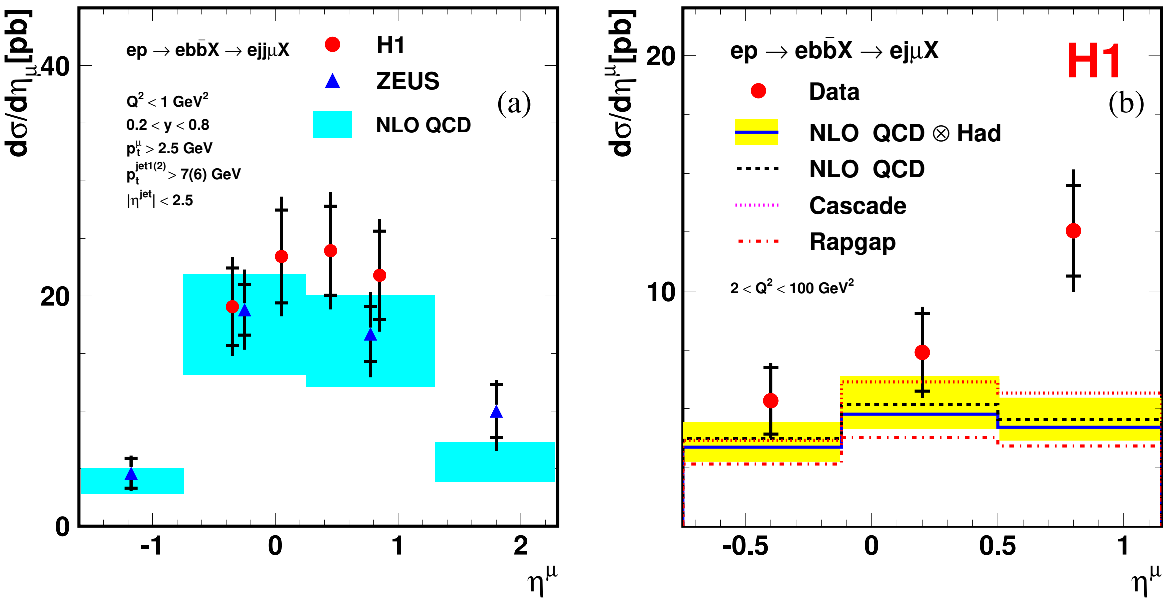


Fig. 3: Measurement of open beauty production as a function of the pseudorapidity of the decay muon for (a) dijet photoproduction from the H1 and ZEUS experiments and (b) inclusive jet deep inelastic scattering from the H1 experiment. (The measurement from ZEUS experiment for (b) is in a different kinematic region but reveals the same physics message and so for brevity, is not shown)

nomenological improvements, a deep understanding of beauty production should be achieved by the turn-on of the LHC.

5 Open charm production

Due to its smaller mass, predictions for charm production are less accurate than for beauty. However large data samples allow detailed comparisons with theory. An example of a measured D^* cross section in deep inelastic scattering is shown in Fig. 4a; data from the two experiments agree with each other and are well described by the prediction of QCD. Similar measurements have been made in photoproduction in which the data is less well described. Due to the larger cross section, the photoproduction data could prove valuable in constraining the photon as well as the proton structure. However, as can be seen from Fig. 4b, the theoretical precision is lagging well behind that of the data. Therefore more exclusive quantities and regions, with smaller theoretical uncertainties, are measured.

Measurements of charm photoproduction accompanied with jets pose a challenge for theory due the extra scale of the jet transverse energy. Such complicated final states will be copious at the LHC, so the verification of theory to HERA data will aid in the understanding of these high-rate QCD events. Dijet correlations in photoproduction have recently been measured [19] and compared with available calculations. Events were selected in two regions: one enriched in direct photon events where the photon acts as a pointlike object and one enriched in resolved photon events where the photon acts as a source of partons. The cross section of the difference in the azimuthal angle, $\Delta\phi^{\rm jj}$, of the two highest E_T jets has been measured. For the LO $2 \to 2$ process, the two jets are back-to-back. The data exhibit a significant cross section at low $\Delta\phi^{\rm jj}$ and for the direct photon events are reasonably well described by NLO QCD (not shown). However, the description for resolved photon events is poor as shown in Fig. 5a. This region is particularly sensitive to higher orders not present in the QCD calculation. Monte Carlo models are compared to the data in Fig. 5b; although the normalisation is poor, the shape of the distribution is very well described by the HERWIG simulation. This indicates that for the precise description of such

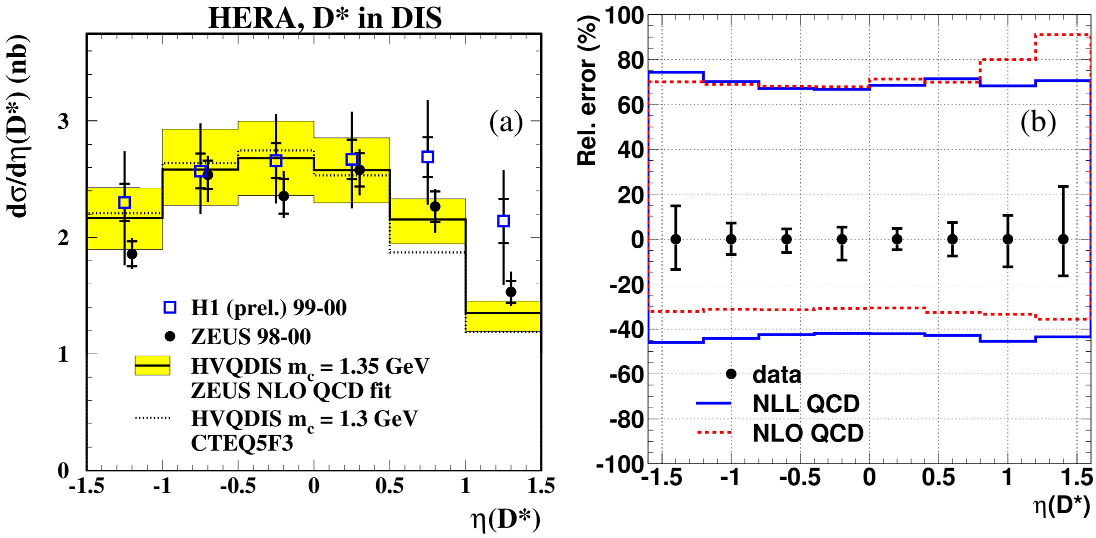


Fig. 4: Measurement of D^* production compared with NLO QCD predictions: (a) the differential cross section in deep inelastic scattering and (b) the relative uncertainty in data and theory in photoproduction.

processes, higher-order calculations or the implementation of additional parton showers in current NLO calculations are needed.

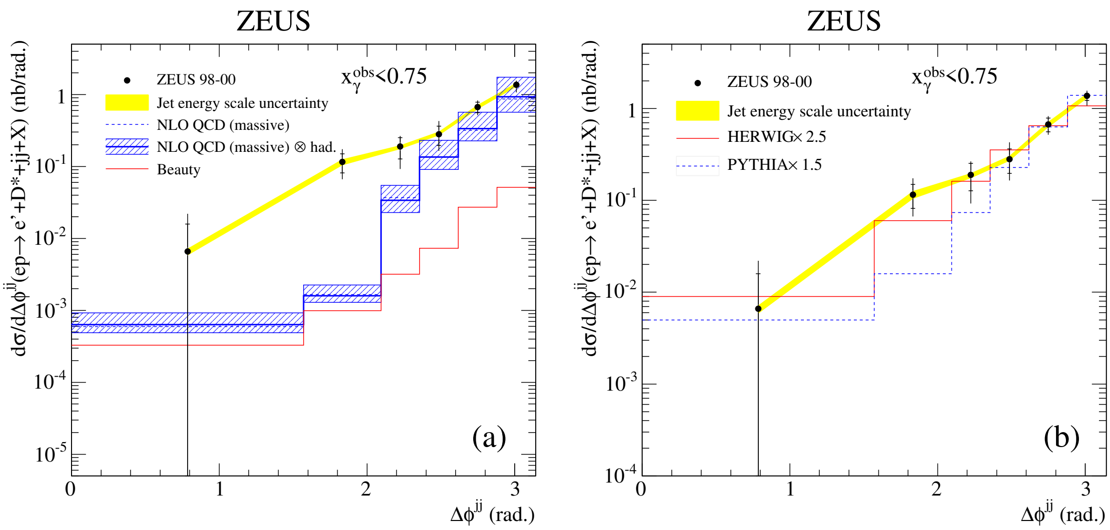


Fig. 5: Difference in the azimuthal angle of the two highest E_T jets in charm photoproduction for a sample enriched in resolved photon events compared to (a) a NLO QCD calculation and (b) Monte Carlo models.

6 The structure of the proton

Open charm (and more recently beauty) production in deep inelastic scattering acts as a powerful probe of the structure of the proton, particularly the gluon and heavy quark densities. Such a direct measurement of the gluon density complements its extraction in QCD fits to inclusive data. The cross section for the production of a heavy quark pair can be written in terms of the heavy quark contribution to the proton structure functions:

$$\frac{d^2\sigma^{Q\bar{Q}}\left(x,Q^2\right)}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left\{\left[1+(1-y)^2\right]F_2^{Q\bar{Q}}\left(x,Q^2\right) - y^2 F_L^{Q\bar{Q}}\left(x,Q^2\right)\right\}$$

The value of the charm contribution, $F_2^{c\bar{c}}$, has traditionally been extracted by measuring D^* mesons within the acceptance of the detector and extrapolating to the full phase space.

The values of $F_2^{c\bar{c}}$ extracted from the measured D^* cross sections [20, 21, 22] are shown in Fig. 6a compared with NLO QCD. New measurements of $F_2^{c\bar{c}}$ have been recently performed using an inclusive sample of high p_T tracks [12]. This data is more inclusive than the D^* measurements probing much lower p_T and thereby having much reduced extrapolation factors (a factor of 1.2 rather than 2–3 as for the D^* measurements). These results confirm the previous data and add extra information on $F_2^{c\bar{c}}$. The results on $F_2^{c\bar{c}}$ demonstrate a large gluon density in the proton as exhibited by the scaling violations versus Q^2 and are well described by such a parton density function. At high Q^2, charm contributes up to about 30% of the inclusive cross section. It is hoped with higher statistics and a better control over the systematics that the charm cross section data can be used in QCD fits to constrain the gluon (or heavy quark) density in the proton.

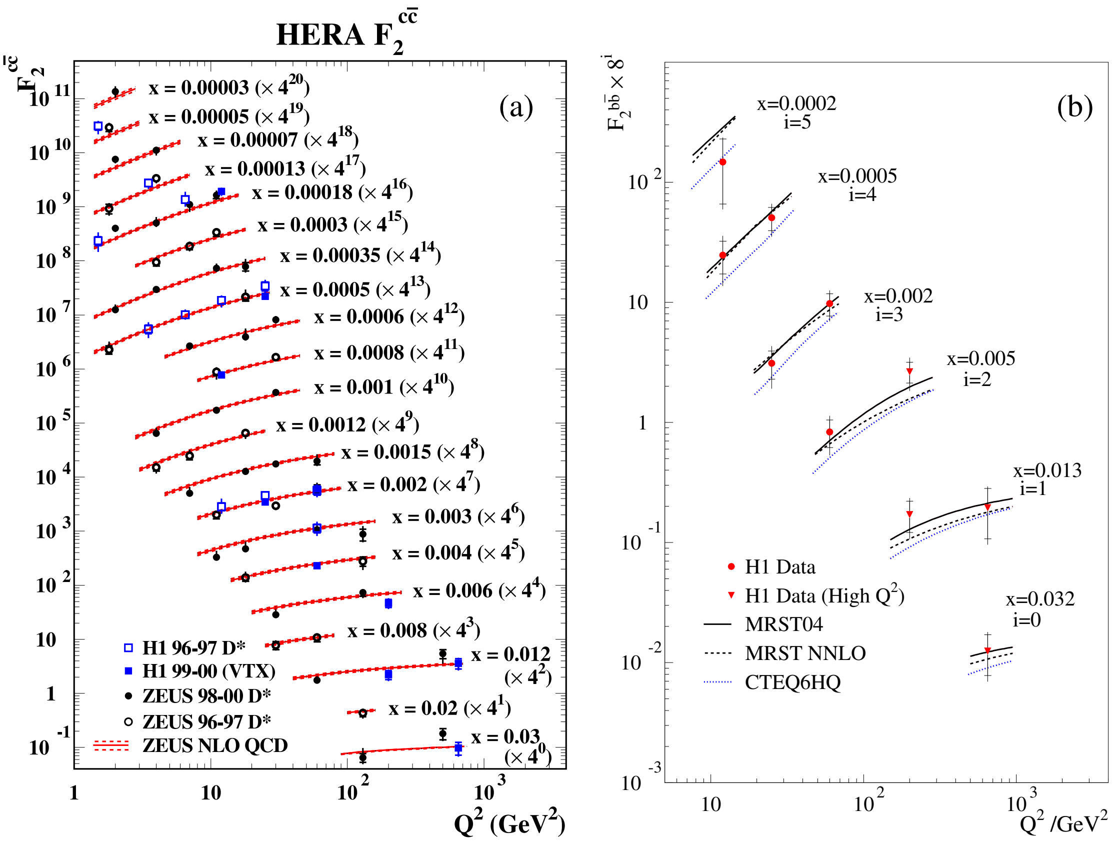


Fig. 6: (a) Charm contribution, $F_2^{c\bar{c}}$, and (b) beauty contribution, $F_2^{b\bar{b}}$, to the proton structure function, F_2, versus Q^2 for fixed x.

Applying the same technique of using high p_T tracks, the H1 collaboration have made measurements of $F_2^{b\bar{b}}$ which are shown in Fig. 6. The results are consistent with scaling violations and are well described by new parton density functions. The differences between the different parametrisations are not insignificant and future measurements should be able to discriminate between them. For the Q^2 range measured, beauty production contributes up to 3% of the inclusive cross section.

7 Universality of charm fragmentation

Heavy quark fragmentation has been extensively studied in e^+e^- collisions. The clean environment, control over the centre-of-mass energy and back-to-back dijet system provide an ideal laboratory for accurate measurement of fragmentation parameters. The measured parameters, e.g. fragmentation function and fraction of charm quarks hadronising to a particular meson, are used as inputs to models and NLO QCD calculations of ep collisions. Therefore, the validity of using fragmentation parameters extracted from e^+e^- data in ep data needs to be verified. The strangeness suppression factor, γ_s, the ratio of neutral and charged D-meson production rates, $R_{u/d}$, the fraction of charged D mesons produced in a vector state, P_v^d and the fragmentation fractions, $f(c \to D, \Lambda)$, have been measured in deep inelastic scattering [23] and in photoproduction [24]. The results are shown in Fig. 7 compared with values obtained in e^+e^- collisions. The data obtained in different processes are consistent with each other and thereby consistent with the concept of universal fragmentation. The measurements in photoproduction also have precision competitive with the combined e^+e^- data. The data therefore provide extra constraints and demonstrate that the fragmentation at a hadron collider in the central part of the detector looks like that in an e^+e^- collision. This will provide useful input for future models to be used at the LHC.

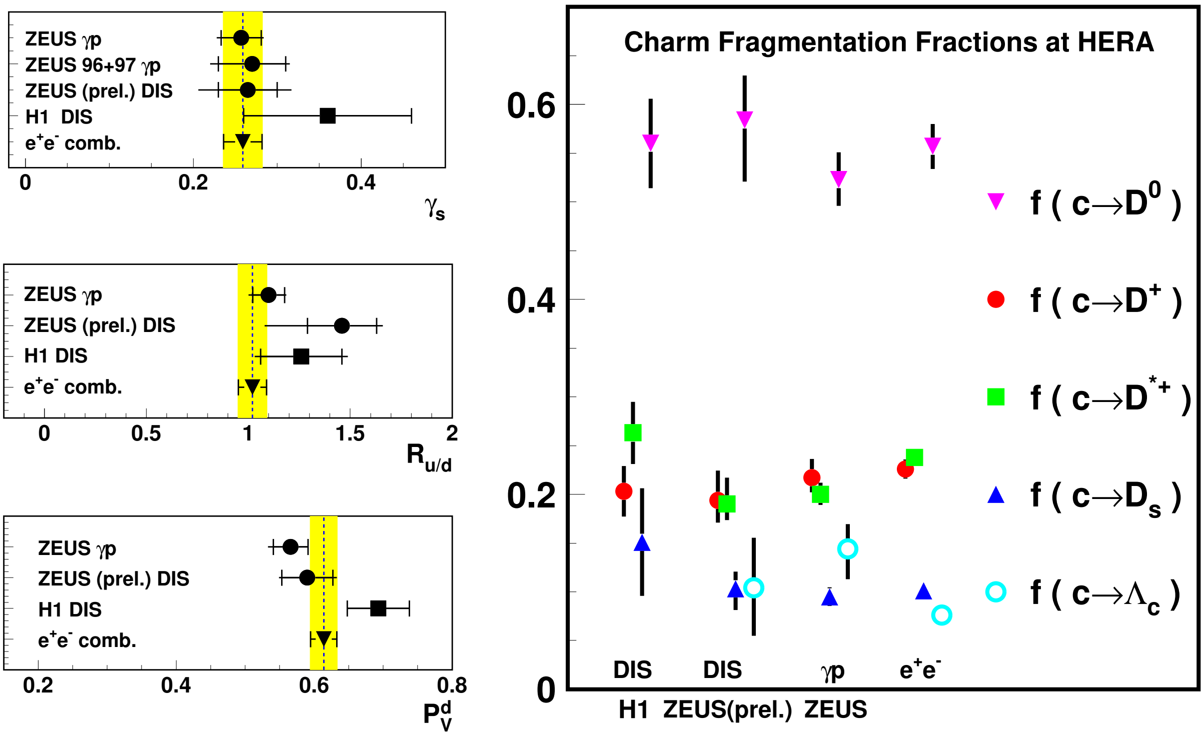


Fig. 7: Comparisons of fragmentation parameters, γ_s, $R_{u/d}$ and P_v^d and $f(c \to D, \Lambda)$ in photoproduction, deep inelastic scattering and e^+e^- collisions.

The charm fragmentation function to D^* mesons has been measured by both the H1 and ZEUS collaborations [25] and compared to e^+e^- data. Although the definitions of the fragmentation function and the energies are different, the general trends are the same. However, a consistent fit to all data within a given Monte Carlo or NLO calculation is needed to clarify this situation. Measurements at the Tevatron would also contribute significantly to this area.

8 Conclusions

An increasing number of high precision measurements of heavy quark production from HERA have recently become available. They are providing valuable information on the parton densities, the overall production rates and the concept of the universality of fragmentation. Precise and well-defined measurements have allowed phenomenological improvements to be made. Generally QCD describes the production of heavy quarks; in particular, due in part to the advances made in the HERA measurements, the prediction for the production of beauty quarks is no longer well below the data. There are some details still lacking which await to be confronted with higher order calculations or NLO calculations interfaced with parton showers and hadronisation. There is also ongoing work in tuning Monte Carlo predictions to all known data which demonstrates the need to have global calculations which can predict all processes under study. In the next few years in the run up to the LHC, HERA will produce a lot more data and more will be known about heavy quark production.

References

[1] ATLAS Coll., *ATLAS detector and physics performance: Technical Design Report*, Technical Report CERN/LHCC 99-14, 1999.

[2] M. Cacciari, M. Greco and P. Nason **9805**, 007 (1998).

[3] S. Frixione et al., Phys. Lett. **B 348**, 633 (1995);
S. Frixione et al., Nucl. Phys. **B 454**, 3 (1995).

[4] B. W. Harris and J. Smith, Phys. Rev. **D 57**, 2806 (1998).

[5] M. Cacciari et al., Phys. Rev. **D 55**, 2736 (1997);
M. Cacciari and M. Greco, Phys. Rev. **D 55**, 7134 (1997).

[6] J. Binnewies, B.A. Kniehl and G. Kramer, Z. Phys. **C 76**, 677 (1997);
B.A. Kniehl, G. Kramer and M. Spira, Z. Phys. **C 76**, 689 (1997);
J. Binnewies, B.A. Kniehl and G. Kramer, Phys. Rev. **D 48**, 014014 (1998);
G. Heinrich and B.A. Kniehl, Phys. Rev. **D 70**, 094035 (2004).

[7] S. Frixione and B.R. Webber, JHEP **0206**, 029 (2002);
S. Frixione, P. Nason and B.R. Webber, JHEP **0308**, 007 (2003).

[8] HERWIG 6.5: G. Corcella et al., JHEP **0101**, 010 (2001).

[9] C. Peterson et al., Phys. Rev. **D 27**, 105 (1983).

[10] N. Brook. Private communication.

[11] PYTHIA 6.154: T. Sjöstrand et al., Comp. Phys. Comm. **135**, 238 (2001).

[12] H1 Coll., A. Aktas et al., Eur. Phys. J. **C 40**, 349 (2005);
H1 Coll., A. Aktas et al. Preprint DESY-05-110, (2005), submitted to Eur. Phys. J.

[13] CDF Coll., F. Abe et al., Phys. Rev. Lett. **71**, 500 (1993);
CDF Coll., F. Abe et al., Phys. Rev. Lett. **71**, 2396 (1993);
CDF Coll., F. Abe et al., Phys. Rev. Lett. **75**, 1451 (1995);
CDF Coll., F. Abe et al., Phys. Rev. **D 53**, 1051 (1996);
CDF Coll., P. Acosta et al., Phys. Rev. **D 65**, 052005 (2002);
D0 Coll., S. Abachi et al., Phys. Rev. Lett. **74**, 3548 (1995);
D0 Coll., B. Abbott et al., Phys. Lett. **B 487**, 264 (2000);
D0 Coll., B. Abbott et al., Phys. Rev. Lett. **84**, 5478 (2000);
D0 Coll., B. Abbott et al., Phys. Rev. Lett. **85**, 5068 (2000).

[14] M. Cacciari and P. Nason, Phys. Rev. Lett. **89**, 122003 (2002).

[15] H1 Coll., C. Adloff et al., Phys. Lett. **B 467**, 156 (1999). Erratum-ibid. **B 518**, 331 (2001).

[16] ZEUS Coll., J. Breitweg et al., Eur. Phys. J. **C 18**, 625 (2001).

[17] ZEUS Coll., S. Chekanov et al., Phys. Rev. **D 70**, 012008 (2004);
H1 Coll., A. Aktas et al., Eur. Phys. J. **C 41**, 453 (2005).

[18] ZEUS Coll., S. Chekanov et al., Phys. Lett. **B 599**, 173 (2004).

[19] ZEUS Coll., S. Chekanov et al. Preprint DESY-05-132, (2005), submitted to Nucl. Phys. B.

[20] ZEUS Coll., J. Breitweg et al., Eur. Phys. J. **C 12**, 35 (2000).

[21] H1 Coll., C. Adloff et al., Phys. Lett. **B 528**, 199 (2002).

[22] ZEUS Coll., S. Chekanov et al., Phys. Rev. **D 69**, 012004 (2004).

[23] H1 Coll., A. Aktas et al., Eur. Phys. J. **C 38**, 447 (2005);
ZEUS Coll. Abstract 266, XXII International Symposium on Lepton-Photon Interactions at High Energy, Uppsala, Sweden, June-July, 2005.

[24] ZEUS Coll., S. Chekanov et al. Preprint DESY-05-147, (2005), submitted to Eur. Phys. C.

[25] ZEUS Coll. Abstract 778, XXX1 International Conference on High Energy Physics Amsterdam, The Netherlands, July, 2002;
H1 Coll. Abstract 407, XXII International Symposium on Lepton-Photon Interactions at High Energy, Uppsala, Sweden, June-July, 2005.

From HERA to the LHC

John Ellis
CERN, Geneva

CERN-PH-TH/2005-243

Abstract
Some personal comments are given on some of the exciting interfaces between the physics of HERA and the LHC. These include the quantitative understanding of perturbative QCD, the possible emergence of saturation phenomena and the Colour-Glass Condensate at small x and large Q^2, the link between forward physics and ultra-high-energy cosmic rays, and new LHC opportunities opened up by the discovery of rapidity-gap events at HERA, including the search for new physics such as Higgs bosons in double-diffraction events.

1 Preview

There are many exciting interfaces between physics at HERA and the LHC, and I cannot do justice to all of them in this talk. Therefore, in this talk I focus on a few specific subjects that interest me personally, starting with the LHC's 'core business', namely the search for new physics at the TeV scale, notably the Higgs boson(s) and supersymmetry [1]. Identifying any signals for such new physics will require understanding of the Standard Model backgrounds, and QCD in particular. I then continue by discussing some other topics of specific interest to the DESY community.

• The understanding of QCD will be important for making accurate studies of any such new physics. Perturbative QCD at moderate x and large p_T is quite well understood, with dramatic further progress now being promised by novel calculational techniques based on string theory [2].

• Novel experimental phenomena are now emerging at RHIC at small x, following harbingers at HERA. The parton density saturates, and a powerful organizational framework is provided by the Colour-Glass Condensate (CGC) [3]. Forward measurements at the LHC will provide unique opportunities for following up on this HERA/RHIC physics.

• Forward physics at the LHC will also provide valuable insight into the interpretation of ultra-high-energy cosmic rays (UHECRs) [4]. One of the principal uncertainties in determining their energy scale is the modeling of the hadronic showers they induce, and the LHC will be the closest laboratory approximation to UHECR energies.

• Looking further forward, there is increasing interest in exploring at the LHC the new vistas in hard and soft diffraction opened up by the discovery of rapidity-gap events at HERA [5]. One particularly interesting possibility is quasi-exclusive diffractive production of Higgs bosons or other new particles at the LHC [6]. This is particularly interesting in supersymmetric extensions of the Standard Model, notably those in which CP is violated [7].

2 Prospects in Higgs Physics

Many studies have given confidence that the Standard Model Higgs boson will be found at the LHC, if it exists [8]. Moreover, there are some chances that it might be found quite quickly, in particular if its mass is between about 160 GeV and 600 GeV. However, discovering the Higgs boson will take rather longer if its mass is below about 130 GeV, as suggested in the minimal supersymmetric extension of the Standard Model (MSSM) [9]. In this case, the Higgs signal would be composed of contributions from several different production and decay channels, notably including $gg \to H \to \gamma\gamma$.

Understanding the gluon distribution at $x \sim 10^{-2}$ is therefore a high priority, and one to which HERA measurements of processes involving gluons have been playing key roles [10]. Perturbative corrections to the $gg \to H$ production process need to be understood theoretically, as do the corrections to $H \to \gamma\gamma$ decay. Resummation of the next-to-next-to-leading logarithms has by now reduced these uncertainties to the 10% level, and further improvements may be possible with the string-inspired calculational techniques now being introduced [11].

Fig. 1 shows estimates of the accuracy with which various Higgs couplings may be determined at the LHC, also if the luminosity may be increased by an order of magnitude (SLHC) [12] [see also [13]]. There are interesting prospects for measuring the couplings to $\tau\tau$, $\bar{b}b$, WW, ZZ and $\bar{t}t$ as well as the total Higgs decay width, though not with great accuracy. Measurements at the ILC would clearly be much more powerful for this purpose [13].

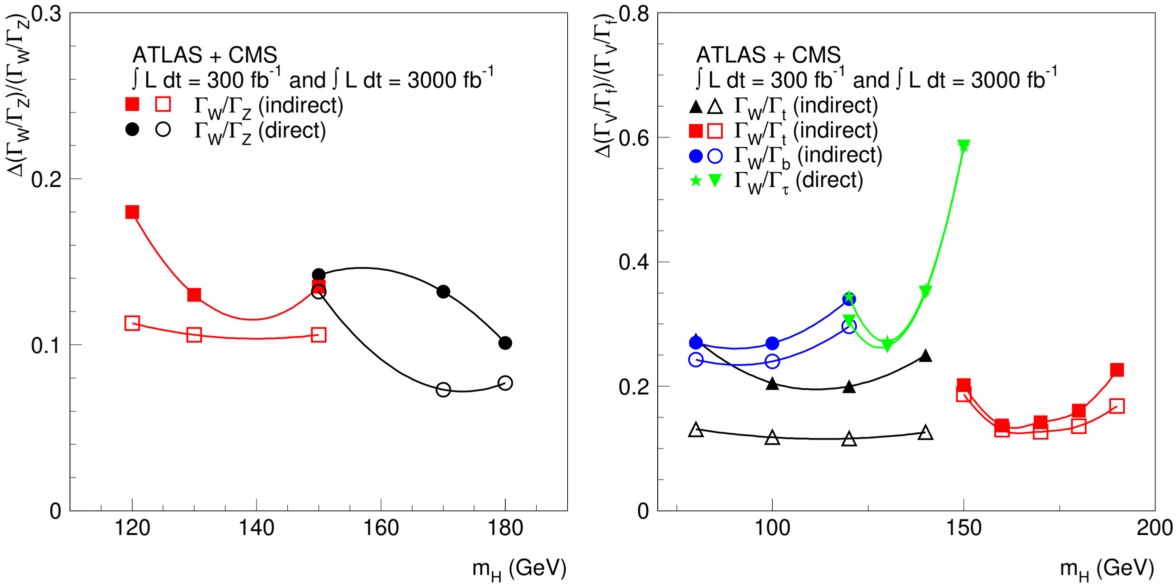


Fig. 1: Illustrations of the accuracy with which Higgs couplings could be measured at the LHC with the planned luminosity and with a possible upgrade by a factor of ten (SLHC) [12].

3 Theorists are Hedging their Bets

The prospect of imminent Higgs discovery is leading theorists to place their last bets on the LHC roulette wheel, and many are hedging their bets by proposing and discussing alternatives to the Standard Model or the MSSM. Composite Higgs models are not greatly favoured, since they have a strong tendency to conflict with the precision electroweak data [14]. This problem has led some theorists to question the interpretation of the electroweak data, which are normally taken to favour $m_H < 300$ GeV, debating their consistency and even arguing that some data should perhaps be discounted [15]. Personally, I see no strong reason to doubt the hints from the electroweak data. An alternative corridor leading towards higher Higgs masses is provided by including higher-dimensional operators in the electroweak data analysis [16]: this would require some fine-tuning, but cannot be excluded. An even more extreme alternative that has been re-explored recently is that of Higgsless models [17]. However, these lead to strong WW scattering and conflict with the available electroweak data. These problems are alleviated, but not solved, by postulating extra dimensions at the TeV scale [18].

One of the least unappetizing alternatives to the supersymmetric Higgs paradigm is offered by little Higgs models [19]. Their key idea is to embed the Standard Model in a larger gauge group, from which the Higgs boson emerges as a relatively light pseudo-Goldstone boson. The one-loop quadratic divergence due to the top quark:

$$\delta m^2_{H,top}(SM) \sim (115 \text{ GeV})^2 \left(\frac{\Lambda}{400 \text{ GeV}}\right)^2$$

is cancelled by the contribution of a new heavy T quark:

$$\delta m^2_{H,top}(LH) \sim \frac{6G_F m_t^2}{\sqrt{2}\pi^2} m_T^2 \log \frac{\Lambda}{m_T}$$

Additionally, there are new gauge bosons and exotic Higgs representations. The Standard-Model-like Higgs boson is expected to be relatively light, possibly below ~ 150 GeV, whereas the other new particles are expected to be heavier:

$$\begin{array}{lcl} M_T & < & 2\,\text{TeV}(m_h/200\text{GeV})^2 \\ M'_W & < & 6\,\text{TeV}(m_h/200\text{GeV})^2 \\ M_{H++} & < & 10\,\text{TeV} \end{array}$$

Certainly the new T quark, probably the W' boson and possibly even the doubly-charged Higgs boson will be accessible to the LHC. Thus little Higgs models have quite rich phenomenology, as well being decently motivated. However, they are not as complete as supersymmetry, and would require more new physics at energies > 10 TeV.

Depending on the mass scale of this new physics, there may be some possibility for distinguishing a little Higgs model from the Standard Model by measurements of the $gg \to H \to \gamma\gamma$ process at the LHC. However, the ILC would clearly have better prospects in this regard [13].

4 Supersymmetry

No apologies for repeating the supersymmetric mantra: it resolves the naturalness aspect of the hierarchy problem by cancelling systematically the quadratic divergences in all loop corrections to the Higgs mass and hence stabilizes the electroweak scale [20], it enables the gauge couplings to unify [21], it predicts $m_H < 150$ GeV [9] as suggested by the precision electroweak data [14], it stabilizes the Higgs potential for low Higgs masses [22], and it provides a plausible candidate [23] for the dark matter that astrophysicists and cosmologists claim clutters up the Universe.

However, all we have from accelerators at the moment are lower limits on the possible supersymmetric particle masses, most notably from the absence of sparticles at LEP: $m_{\tilde{\ell}}, m_{\chi^\pm} > 100$ GeV and the Tevatron collider: $m_{\tilde{g}}, m_{\tilde{q}} > 300$ GeV, the LEP lower limit $m_H > 114.4$ GeV, and the consistency of $b \to s\gamma$ decay with the Standard Model. However, if we assume that the astrophysical cold dark matter is largely composed of the lightest supersymmetric particle (LSP), and require its density to lie within the range allowed by WMAP et al [24]:

$$0.094 < \Omega_\chi h^2 < 0.129,$$

we obtain upper as well as lower limits on the possible sparticle masses. The anomalous magnetic moment of the muon, $g_\mu - 2$, provides intermittent hints on the supersymmetric mass scale [25]: these are lower limits if you do not believe there is any significant discrepancy with the Standard Model prediction, but also an upper limit if you do not believe that the Standard Model can fit the data, as is indicated by the current interpretation of the e^+e^- data used to calculate the Standard Model prediction.

If one compares the production of the lightest neutral Higgs boson in the constrained MSSM (CMSSM) in which all the soft supersymmetry-breaking scalar masses m_0 and gaugino masses $m_{1/2}$ are assumed to be universal, the *good news* is that the rate for $gg \to h \to \gamma\gamma$ is expected to be within

10% of the Standard Model value, as seen in Fig. 2(a) [26]. On the other hand, the bad news is the rates are so similar that it will be difficult to distinguish a CMSSM Higgs boson from its Standard Model counterpart. This would be much easier at the ILC, as seen in Fig. 2(b) [27].

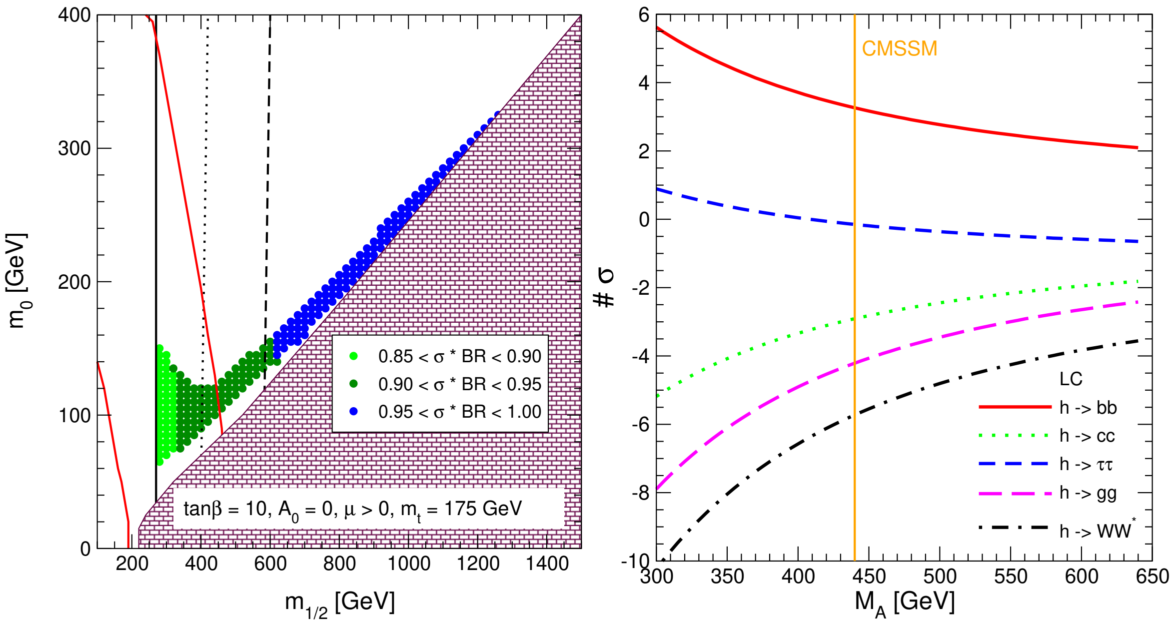


Fig. 2: Left panel: The cross section for production of the lightest CP-even MSSM Higgs boson in gluon fusion and its decay into a photon pair, $\sigma(gg \to h) \times \mathcal{B}(h \to \gamma\gamma)$, normalized to the Standard Model value with the same Higgs mass, is given in the $(m_{1/2}, m_0)$ plane for $\mu > 0$, $\tan\beta = 10$, assuming $A_0 = 0$ and $m_t = 175$ GeV [26]. The diagonal (red) solid lines are the $\pm 2 - \sigma$ contours for $g_\mu - 2$. The near-vertical solid, dotted and dashed (black) lines are the $m_h = 113, 115, 117$ GeV contours. The (brown) bricked regions are excluded since in these regions the LSP is the charged $\tilde{\tau}_1$. Right panel: The numbers of standard deviations by which the predictions of the MSSM with non-universal Higgs masses may be distinguished from those of the Standard Model in different channels by measurements at the ILC [27]. The predictions with the CMSSM values of M_A and μ are indicated by light vertical (orange) lines. The other parameters have been chosen as $m_{1/2} = 300$ GeV, $m_0 = 100$ GeV, $\tan\beta = 10$ and $A_0 = 0$.

One of the distinctive possibilities opened up by the MSSM is the possibility of CP violation in the Higgs sector, induced radiatively by phases in the gaugino masses and the soft supersymmetry-breaking trilinear couplings. Fig. 3 displays CP-violating asymmetries that might be observable in the $gg, \bar{b}b \to \tau^+\tau^-$ and $W^+W^- \to \tau^+\tau^-$ processes at the LHC, in one particular CP-violating scenario with large three-way mixing between all three of the neutral MSSM Higgs bosons [28].

A typical supersymmetric event at the LHC is expected to contain high-p_T jets and leptons, as well as considerable missing transverse energy. Studies show that the LHC should be able to observe squarks and gluinos weighing up to about 2.5 TeV [8], covering most of the possibilities for astrophysical dark matter. As seen in Fig. 4(a) [1], the dark matter constraint restricts $m_{1/2}$ and m_0 to narrow strips extending to an upper limit $m_{1/2} \sim 1$ TeV. As seen in Fig. 4(b), whatever the value of $m_{1/2}$ along one of these strips, the LHC should be able to observe several distinct species of sparticle [1]. In a favourable case, such as the benchmark point B in Fig. 4(a) (also known as SPS Point 1a), experiments at the LHC should be able to measure the CMSSM parameters with sufficient accuracy to calculate the supersymmetric relic density $\Omega_\chi h^2$ (blue histogram) with errors comparable to the present astrophysical error (yellow band) as seen in Fig. 4(c) [1]. Fig. 4(d) summarizes the scapabilities of the LHC and other accelerators to detect various numbers of sparticle species. We see that the LHC is almost guaranteed to discover supersymmetry if it is relevant to the naturalness of the mass hierarchy. However, there are

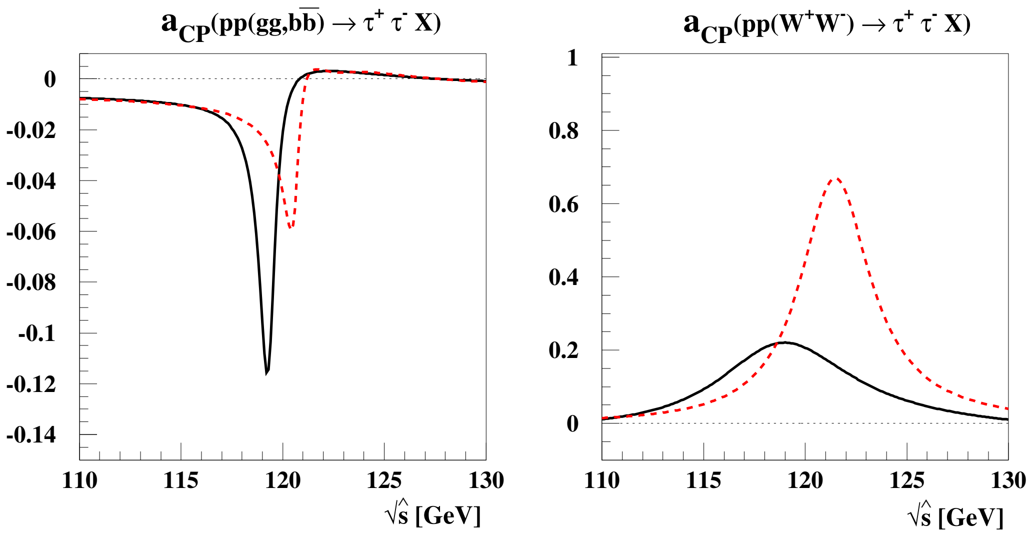


Fig. 3: Numerical estimates of differential CP asymmetries as functions of the effective reduced centre-of-mass energy $\sqrt{\hat{s}}$ in a CP-violating three-Higgs mixing scenario with gaugino phase $\Phi_3 = -90^\circ$ (solid lines) and $\Phi_3 = -10^\circ$ (dashed lines) [28].

some variants of the CMSSM, in particular at the tips of the WMAP strips for large $\tan\beta$, that might escape detection at the LHC.

As we also see in Fig. 4(d), linear colliders would be able to observe a complementary subset of sparticles, particularly sleptons, charginos and neutralinos [1]. A linear collider with a centre-of-mass energy of 1 TeV would have comparable physics reach to the LHC, but a higher centre-of-mass energy, such as the 3 TeV option offered by CLIC [29], would be needed to complete the detection and accurate measurement of all the sparticles in most variants of the CMSSM.

We have recently evaluated whether precision low-energy observables currently offer any hint about the mass scale of supersymmetric particles, by exploring their sensitivities to $m_{1/2}$ along WMAP lines for different values of the trilinear supersymmetry-breaking parameter A_0 and the ratio of Higgs v.e.v's, $\tan\beta$ [31]. The measurements of m_W and $\sin^2\theta_W$ each currently favour $m_{1/2} \sim 300$ GeV for $\tan\beta = 10$ and $m_{1/2} \sim 600$ GeV for $\tan\beta = 50$. The agreement of $b \to s\gamma$ decay with the Standard Model is compatible with a low value of $m_{1/2}$ for $\tan\beta = 10$ but prefers a larger value for $\tan\beta = 50$, whereas $B_s \to \mu^+\mu^-$ decay currently offers no useful information on the scale of supersymmetry breaking [30]. The current disagreement of the measured value of the anomalous magnetic moment of the muon, $g_\mu - 2$, also favours independently $m_{1/2} \sim 300$ GeV for $\tan\beta = 10$ and $m_{1/2} \sim 600$ GeV for $\tan\beta = 50$. Putting all these indications together, as seen in Fig. 5, we see a preference for $m_{1/2} \sim 300$ GeV when $\tan\beta = 10$, and a weaker preference for $m_{1/2} \sim 600$ GeV when $\tan\beta = 50$ [31]. At the moment, this preference is far from definitive, and $m_{1/2} \to \infty$ is excluded at lass than 3 σ, but it nevertheless offers some hope that supersymmetry might lurk not far away.

As seen in Fig. 6, the likelihood function for $m_{1/2}$ can be converted into the corresponding likelihood functions for the masses of various species of sparticles. The preferred squark and gluino masses lie below 1000 GeV for $\tan\beta = 10$, with somewhat heavier values for $\tan\beta = 50$, though still well within the reach of the LHC [31].

5 Gravitino Dark Matter

The above analysis assumed that the lightest supersymmetric particle (LSP) is the lightest neutralino χ, assuming implicitly that the gravitino is sufficiently heavy and/or rare to have been neglected. This

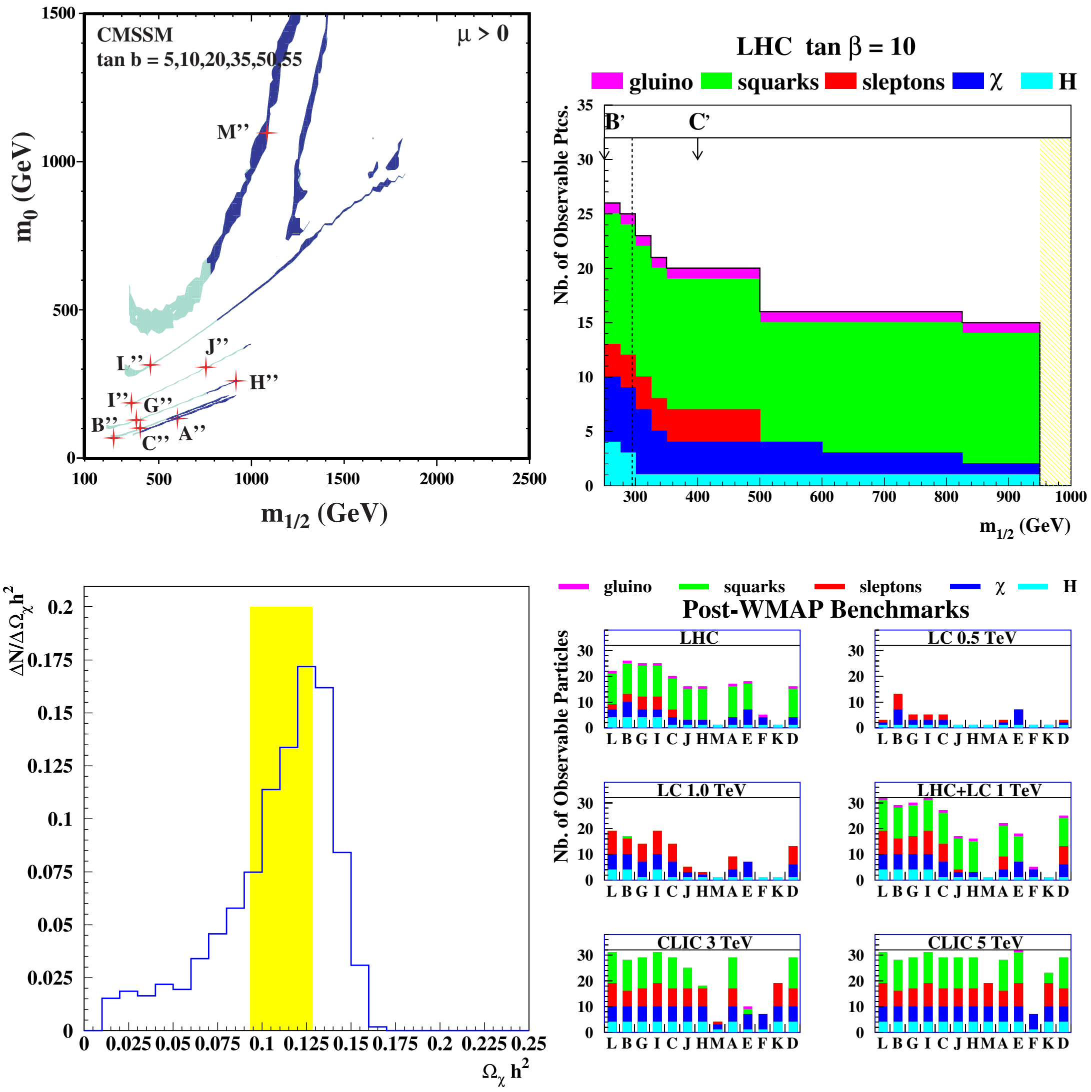


Fig. 4: Top left panel: The strips of CMSSM parameter space allowed by WMAP and other constraints, with specific benchmark scenarios indicated by (red) crosses. Top right panel: The numbers of MSSM particle species observable at the LHC as a function of $m_{1/2}$ along the WMAP strip for $\tan\beta = 10$ [26]. Bottom left panel: The accuracy with which the relic dark matter density could be calculated using LHC measurements at benchmark point B, compared with the uncertainty provided by WMAP and other astrophysical data. Bottom right panel: The numbers of MSSM particle species observable in the benchmark scenarios at the LHC and e^+e^- colliders with different centre-of-mass energies [27].

implicit assumption may or may not be true in a minimal supergravity model, where the gravitino mass $m_{3/2} = m_0$, as seen in Fig. 7 [1] [32]. In this model, the gravitino mass is fixed throughout the $(m_{1/2}, m_0)$ plane: there is a familiar WMAP strip where the χ is the LSP, but there is also a wedge of parameter space where the LSP is the gravitino. There is no way known to detect such astrophysical gravitino dark matter (GDM), since the gravitino has very weak interactions.

However, the LHC may have prospects for detecting GDM indirectly [33, 34, 35]. In the GDM region, the lighter stau, $\tilde{\tau}_1$, is expected to be the next-to-lightest sparticle (NLSP), and may be metastable with a lifetime measurable in hours, days, weeks, months or even years. The $\tilde{\tau}_1$ would be detectable in CMS or ATLAS as a slow-moving charged particle. Staus that are sufficiently slow-moving might be stopped in the detector itself, in some external detection volume designed to observe and measure their

[1]Minimal supergravity also relates the trilinear and bilinear supersymmetry-breaking parameters: $A_0 = B_0 + 1$, thereby fixing $\tan\beta$ as a function of $m_{1/2}, m_0$ and A_0, see the contours in Fig. 7(b).

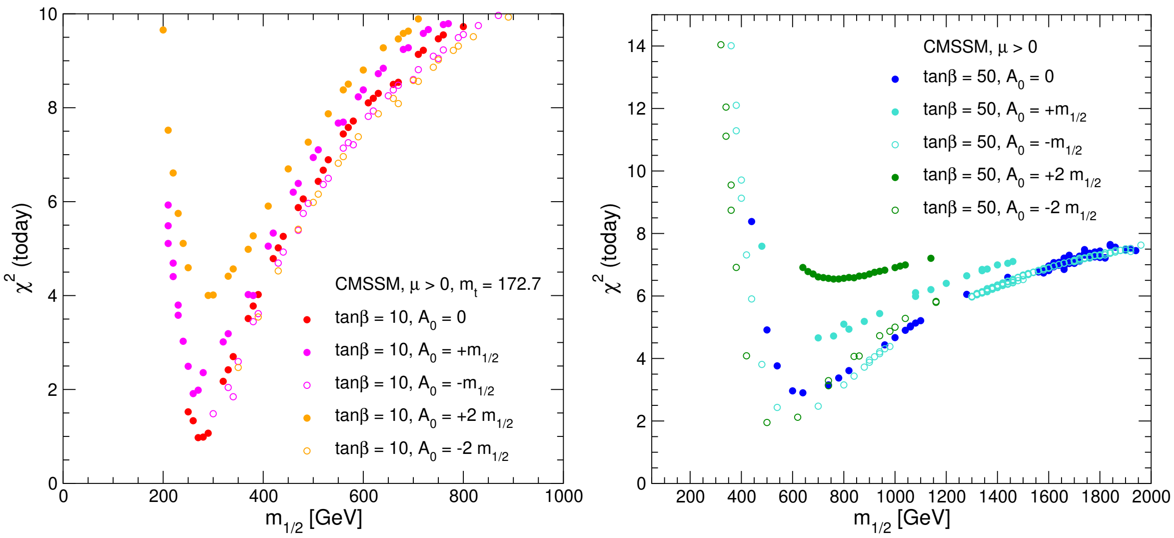


Fig. 5: The results of χ^2 fits based on the current experimental results for the precision observables M_W, $\sin^2\theta_{weff}$, $(g-2)_\mu$ and $b \to s\gamma$ are shown as functions of $m_{1/2}$ in the CMSSM parameter space with WMAP constraints for different values of A_0 and (left) $m_t = 172.7 \pm 2.9$ GeV and $\tan\beta = 10$ and (right) $m_t = 178.0 \pm 4.3$ GeV and $\tan\beta = 50$ [31].

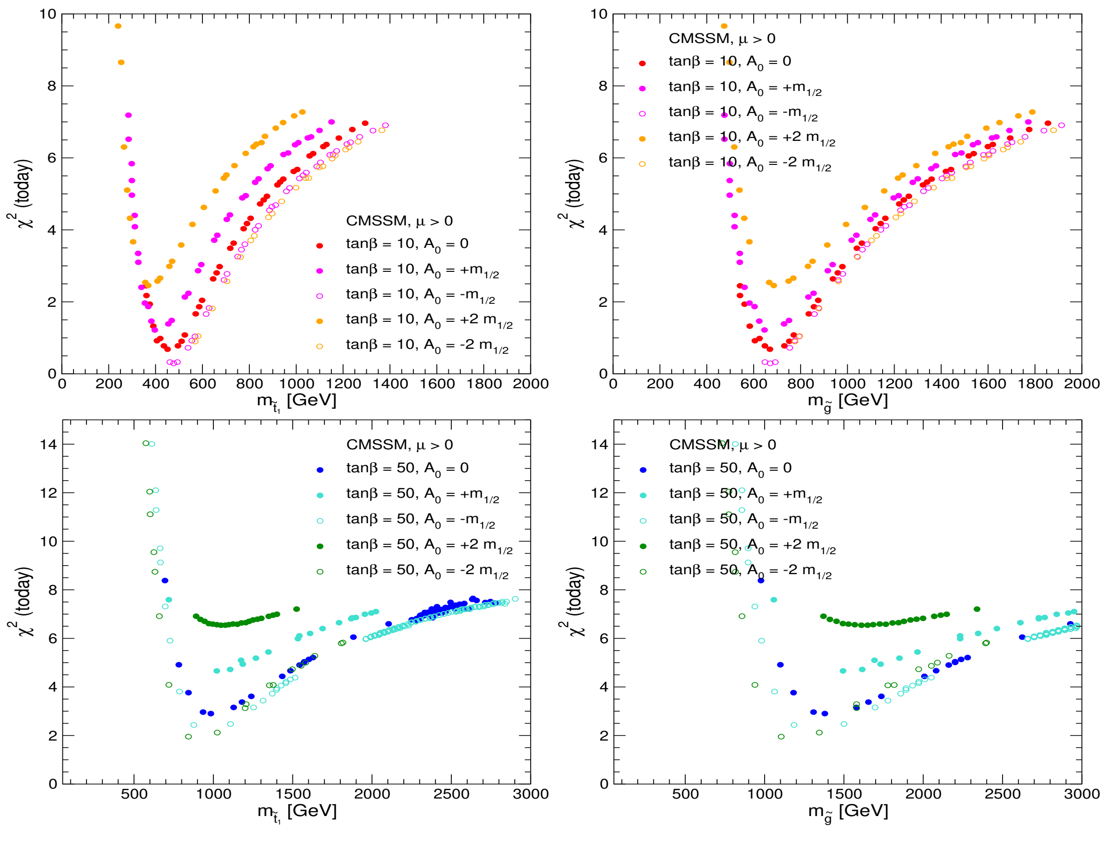


Fig. 6: The χ^2 contours in the CMSSM with $\tan\beta = 10$ for the lighter stop (left) and gluino (right) masses, assuming $\tan\beta = 10$ (top) and $\tan\beta = 50$ (bottom) [31].

late decays into GDM [33, 34], or in the walls of the caverns surrounding the detectors [35].

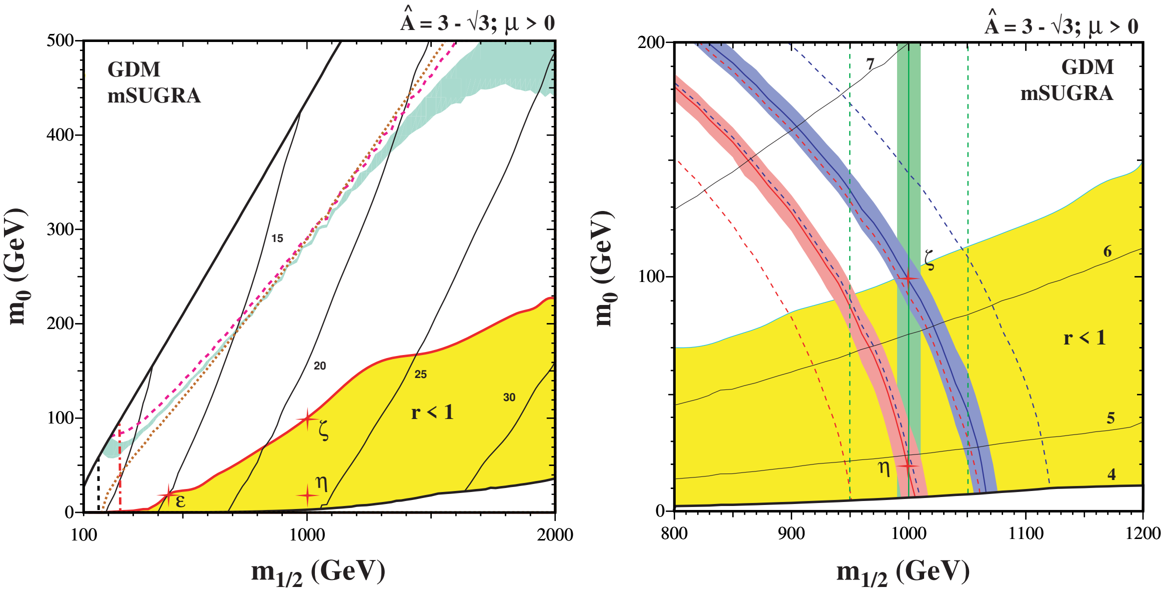


Fig. 7: Left panel: The allowed region in the $(m_{1/2}, m_0)$ planes for minimal supergravity (mSUGRA) with a gravitino LSP [light (yellow) shaded regions labelled $r < 1$], for $A \equiv \hat{A} m_0 : \hat{A} = 3 - \sqrt{3}$. The red crosses denote the benchmark GDM models ϵ, ζ and η [35]. Right panel: The potential impacts on the determination of GDM parameters in the mSUGRA framework of prospective measurement errors of 1 % and 5 % for $m_{\tilde{\tau}_1}$ (diagonal bands and lines) and $m_{1/2}$ (vertical bands and lines), shown as constraints in the $(m_{1/2}, m_0)$ plane [35]. The smaller errors would enable the benchmark scenarios ζ and η to be distinguished, and the possible NLSP lifetime to be estimated. The near-horizontal thin solid lines are labelled by the logarithm of the NLSP lifetime in seconds.

6 The LHC and Ultra-High-Energy Cosmic Rays

Historically, the two experiments with (until recently) the largest statistics for ultra-high-energy cosmic rays (UHECRs), AGASA [36] and HiRes [37], have not agreed on their energy spectra above about 10^{19} eV and, specifically, whether there is a significant number of events beyond the GZK cutoff due to interactions of primary UHECRs with the cosmic microwave background radiation. The Auger experiment now has the second-largest statistics but does not yet have sufficient data to settle the issue [38], though these should soon be forthcoming. If there are super-GZK events, they might be due either to nearby astrophysical sources that have not yet been identified, or (more speculatively) to the decays of metastable superheavy particles [39]. Normalizing the energies of UHECRs requires understanding of the development of extensive air showers. At the moment, this is not very well known, and models of shower development are not even able to tell us the composition of cosmic rays with lower energies between 10^{15} and 10^{19} eV [4].

The LHC is the accelerator that comes closest to reproducing the UHECR energy range, with a centre-of-mass energy corresponding to 4×10^{17} eV, in the range where the cosmic-ray composition is still uncertain. This uncertainty would be reduced by better modelling of hadronic showers, which would in turn benefit from measurements in the forward direction [4].

Unfortunately, the LHC is currently not equipped to make good measurements in this kinematic region, where most of the centre-of-mass energy is deposited. More instrumentation in the forward direction would be most welcome in both CMS and ATLAS. This region is also of fundamental importance for our understanding of QCD, as I now explain.

7 Back to Forward QCD

We discussed earlier the success of perturbative QCD, and the accuracy with which it could be used to calculate high-p_T physics, thanks to the structure functions provided by HERA data [10], in particular. The simple parton description is expected, however, to break down at 'small' x and 'large' Q^2, due to saturation effects. At small x, there is a large probability to emit an extra gluon $\sim \alpha_s \ln(1/x)$, and the number of gluons grows in a limited transverse area. When the transverse density becomes large, partons of size $1/Q$ may start to overlap, and non-linear effects may appear, such as the annihilation of low-x partons. The Malthusian growth in the number of gluons seen at HERA is eventually curbed by these annihilation effects when $\ln(1/x)$ exceeds some critical x-dependent saturation value of Q^2. At larger values of x, the parton evolution with Q^2 is described by the usual DGLAP equations, and the evolution with $\ln(1/x)$ is described by the BFKL equation. However, at lower values of x and large Q^2, a new description is need for the saturated configuration, for which the most convincing proposal is the Colour-Glass Condensate (CGC) [3].

According to the CGC proposal, the proton wave function participating in interactions at low x and Q^2 is to be regarded as a classical colour field that fluctuates more slowly than the collision time-scale. This possibility may be probed in Gold-Gold collisions at RHIC and proton-proton collisions at the LHC: the higher beam energy of LHC compensates approximately for the higher initial parton density in Gold-Gold collisions at RHIC. At central rapidities $y \sim 0$, effects of the CGC are expected to appear only when the parton transverse momentum < 1 GeV. However, CGC effects are expected to appear at larger parton transverse momenta in the forward direction when $y \sim 3$. Lead-Lead collisions at the LHC should reveal even more important saturation effects [40].

What is the experimental evidence for parton saturation? First evidence came from HERA, and Fig. 8(a) displays an extraction of the saturation scale from HERA data [41]. At RHIC, in proton-nucleus collisions one expects the suppression of hard particles at large rapidity and small angle compared to proton-proton collisions, whereas one expects an enhancement at small rapidity, the nuclear 'Cronin effect'. The data [42] from the BRAHMS collaboration at RHIC shown in Fig. 8(b) are quite consistent with CGC expectations [43], but it remains to be seen whether this approach can be made more quantitative than older nuclear shadowing ideas.

8 New Physics in Diffraction?

HERA has revealed a menagerie of different diffractive phenomena, opening up a Pandora's box of possible new physics at the LHC. Classically one had soft diffraction dissociation in peripheral proton-proton collisions, in which one (or both) of the colliding protons would dissociate into a low-mass system (or systems). HERA discovered an additional class of diffractive events [5], which may be interpreted [44] as a small colour dipole produced by an incoming virtual photon penetrates the proton and produces a high-mass system. Additionally, one expects at the LHC soft double diffraction, in which a peripheral proton-proton collision produces a low-mass central system separated from each beam by a large rapidity gap. Events with mixed hard and soft diffraction are also possible at the LHC, as are events with multiple large rapidity gaps. The LHC will certainly provide good prospects for deepening our understanding of diffraction, building upon the insights being gained from HERA.

Double diffraction also offers the possibility of searching for new physics in a relatively clean experimental environment containing, in addition to Higgs boson or other new particle, only a couple of protons or their low-mass diffraction-dissociation products [2]. The leading-order cross-section formula is [6]:

[2]New physics might also be produced in other classes of diffractive events, but with less distinctive signatures.

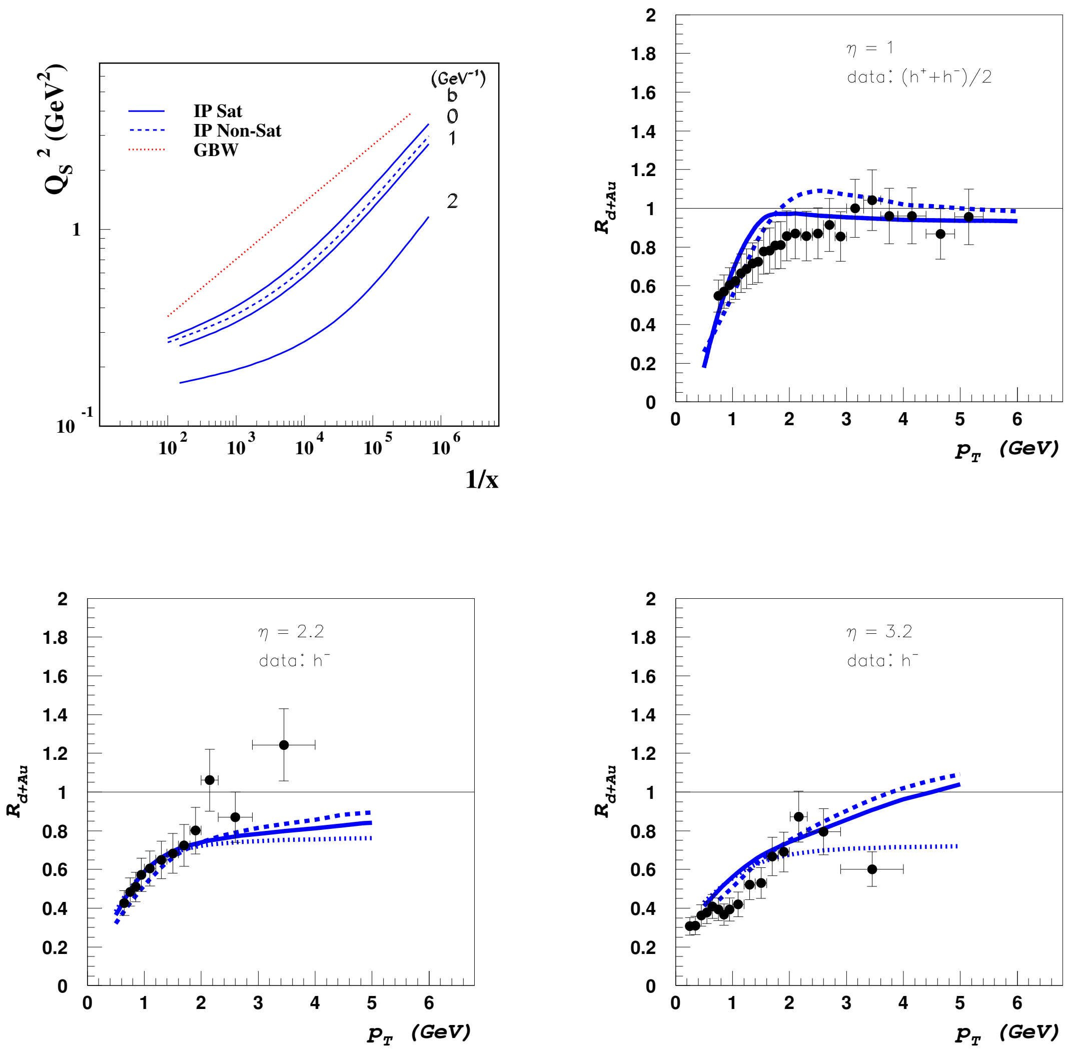


Fig. 8: Top left panel: The parton saturation scale as a function of Bjorken x, extracted from HERA data in [41]. Other three panels: Nuclear modification factor R_{dAu} of charged particles for rapidities $\eta = 1, 2.2, 3.2$ [42], compared with calculations from [43].

$$M^2 \frac{\partial^2 \mathcal{L}}{\partial y \partial M^2} = 4.0 \times 10^{-4} \left[\frac{\int_{lnQ_{min}}^{ln\mu} F_g(x_1, x_2, Q_T, \mu) dlnQ_T}{\mathrm{GeV}^{-}2} \right]^2 \left(\frac{\hat{S}^2}{0.02} \right) \left(\frac{4}{b\mathrm{GeV}^2} \right)^2 \left(\frac{R_g}{1.2} \right)^4$$

where nominal values of the diffractive parameters are quoted in the brackets. The gluon collision factor is currently inferred from HERA data via different parameterizations of the integrated gluon distribution function, and has an uncertainty of a factor of about two [6]. Further analyses of HERA data, as well as future LHC data, would enable the determination to be refined.

The observation of diffractive Higgs production at the LHC would be a challenge in the Standard Model, but the cross section is expected to be considerably larger in the MSSM, particularly at large $\tan\beta$. One of the enticing possibilities offered by supersymmetry is a set of novel mechanisms for CP violation induced by phases in the soft supersymmetry-breaking parameters [7]. These would show up in the MSSM Higgs sector, generating three-way mixing among the neutral MSSM Higgs bosons. This might be observable in inclusive Higgs production at the LHC [7], but could be far more dramatic in double diffraction. Fig. 9(a) displays the mass spectrum expected in double diffraction in one particular three-way mixing scenario [45]: it may exhibit one or more peaks that do not coincide with the Higgs masses. Analogous structures may also be seen in CP-violating asymmetries in $H_i \to \tau^+\tau^-$ decay, as seen in Fig. 9(b). These structures could not be resolved in conventional inclusive Higgs production at the LHC, but may be distinguished in exclusive double diffraction by exploiting the excellent missing-mass resolution ~ 2 GeV that could be provided by suitable forward spectrometers [46].

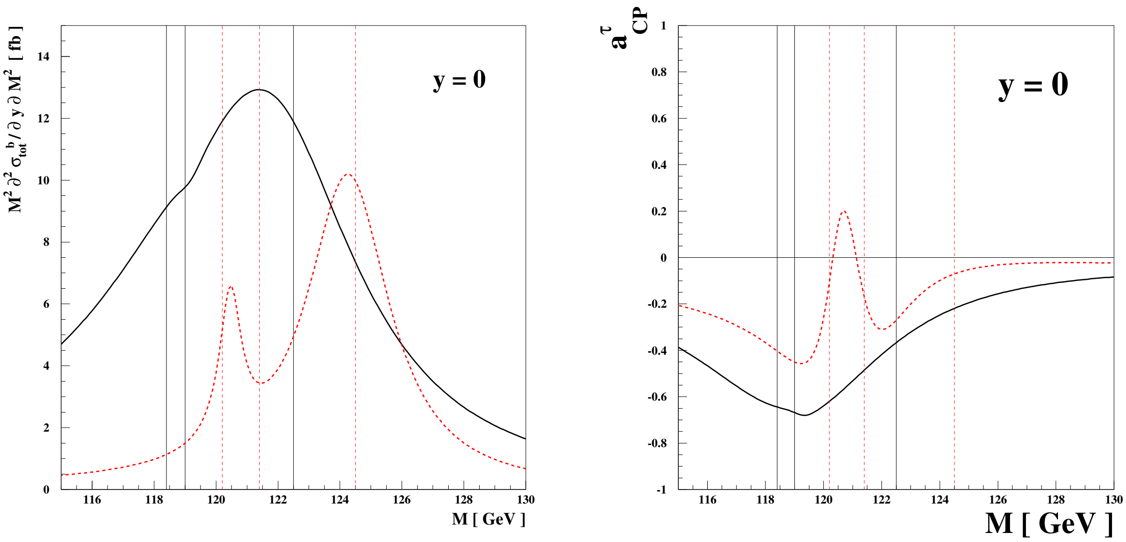


Fig. 9: Left panel: The hadron-level cross section for the double-diffractive producion of Higgs bosons decaying into b quarks. CP-violating three-way mixing scenarios have been taken, with the gluino phase $\Phi_3 = -90^\circ$ (solid lines) and $\Phi_3 = -10^\circ$ (dotted line). The vertical lines indicate the three Higgs-boson pole-mass positions. Right panel: The CP-violating asymmetry $a^\tau_{\rm CP}$ observable in three-way mixing scenarios when Higgs bosons decay into τ leptons, using the same line styles [45].

9 Summary

We do not know what the LHC will find - maybe there will be no supersymmetry and we will observe mini-black-hole production instead! However, whatever the physics scenario, HERA physics will provide crucial inputs, for example via measuring the parton distributions that will be crucial for searches for new physics such as the Higgs boson, or via the observation of saturation effects that will be important for forward physics, or via measurements of diffraction.

Forward physics is a potentially exciting area of LHC physics that is not covered by the present detectors. HERA and RHIC suggest that parton saturation and the Colour Glass Condensate may be observable here, understanding of forward physics is essential for the modelling of cosmic-ray air showers and hence determining the spectrum of ultra-high-energy cosmic rays, and diffractive events related to those observed by HERA may be a valuable tool for discovering new physics such as Higgs production. There is still plenty of room at the LHC for novel experimental contributions [46].

References

[1] See, for example: M. Battaglia *et al.*, Eur. Phys. J. C **22** (2001) 535 [arXiv:hep-ph/0106204]; B. C. Allanach *et al.*, in *Proc. of the APS/DPF/DPB Summer Study on the Future of Particle*

Physics (Snowmass 2001) ed. N. Graf, Eur. Phys. J. C **25** (2002) 113 [eConf **C010630** (2001) P125] [arXiv:hep-ph/0202233]; M. Battaglia, A. De Roeck, J. R. Ellis, F. Gianotti, K. A. Olive and L. Pape, Eur. Phys. J. C **33** (2004) 273 [arXiv:hep-ph/0306219].

[2] F. Cachazo, P. Svrcek and E. Witten, JHEP **0409** (2004) 006 [arXiv:hep-th/0403047]; for a review, see: V. V. Khoze, *Gauge theory amplitudes, scalar graphs and twistor space*, arXiv:hep-th/0408233.

[3] See, for example: E. Iancu, A. Leonidov and L. McLerran, *The colour glass condensate: An introduction*, arXiv:hep-ph/0202270.

[4] See, for example: A. A. Watson, *Observations of Ultra-High Energy Cosmic Rays*, arXiv:astro-ph/0511800.

[5] M. Derrick *et al.* [ZEUS Collaboration], Phys. Lett. B **315** (1993) 481; T. Ahmed *et al.* [H1 Collaboration], Nucl. Phys. B **429** (1994) 477.

[6] V. A. Khoze, A. B. Kaidalov, A. D. Martin, M. G. Ryskin and W. J. Stirling, *Diffractive processes as a tool for searching for new physics*, arXiv:hep-ph/0507040.

[7] See, for example: J. S. Lee, A. Pilaftsis, M. Carena, S. Y. Choi, M. Drees, J. R. Ellis and C. E. M. Wagner, *CPsuperH: A computational tool for Higgs phenomenology in the minimal supersymmetric standard model with explicit CP violation*, Comput. Phys. Commun. **156** (2004) 283 [arXiv:hep-ph/0307377].

[8] ATLAS Collaboration, *ATLAS detector and physics performance Technical Design Report*, CERN/LHCC 99-14/15 (1999); CMS Collaboration, *Technical Proposal*, CERN/LHCC 94-38 (1994).

[9] Y. Okada, M. Yamaguchi and T. Yanagida, Prog. Theor. Phys. **85** (1991) 1; Phys. Lett. B **262** (1991) 54; H. E. Haber and R. Hempfling, Phys. Rev. Lett. **66** (1991) 1815; J. R. Ellis, G. Ridolfi and F. Zwirner, Phys. Lett. B **257** (1991) 83; Phys. Lett. B **262** (1991) 477; A. Yamada, Phys. Lett. B **263**, 233 (1991); M. Drees and M. M. Nojiri, Phys. Rev. D **45** (1992) 2482; P. H. Chankowski, S. Pokorski and J. Rosiek, Phys. Lett. B **274** (1992) 191; Phys. Lett. B **286** (1992) 307.

[10] See, for example: R. S. Thorne, A. D. Martin, R. G. Roberts and W. J. Stirling, *Recent progress in parton distributions and implications for LHC physics*, arXiv:hep-ph/0507015.

[11] See, for example: L. J. Dixon, E. W. N. Glover and V. V. Khoze, JHEP **0412** (2004) 015 [arXiv:hep-th/0411092]. S. D. Badger, E. W. N. Glover and V. V. Khoze, JHEP **0503** (2005) 023 [arXiv:hep-th/0412275].

[12] See, for example: F. Gianotti *et al.*, Eur. Phys. J. C **39** (2005) 293 [arXiv:hep-ph/0204087].

[13] G. Weiglein *et al.* [LHC/LC Study Group], *Physics interplay of the LHC and the ILC*, arXiv:hep-ph/0410364.

[14] See, for example: M. W. Grunewald, *Precision tests of the standard model*, arXiv:hep-ex/0511018.

[15] M. S. Chanowitz, Phys. Rev. D **66** (2002) 073002 [arXiv:hep-ph/0207123].

[16] R. Barbieri and A. Strumia, arXiv:hep-ph/0007265; R. Barbieri, A. Pomarol, R. Rattazzi and A. Strumia, arXiv:hep-ph/0405040.

[17] See, for example: T. Rizzo, arXiv:hep-ph/0409042;

[18] See, for example: J. L. Hewett, B. Lillie and T. G. Rizzo, arXiv:hep-ph/0407059; H. Davoudiasl, J. L. Hewett, B. Lillie and T. G. Rizzo, JHEP **0405** (2004) 015 [arXiv:hep-ph/0403300].

[19] See, for example: M. Schmaltz, *Physics beyond the standard model (Theory): Introducing the little Higgs*, Nucl. Phys. Proc. Suppl. **117** (2003) 40 [arXiv:hep-ph/0210415]; JHEP **0408** (2004) 056 [arXiv:hep-ph/0407143].

[20] L. Maiani, *Proceedings of the 1979 Gif-sur-Yvette Summer School On Particle Physics*, 1; G. 't Hooft, in *Recent Developments in Gauge Theories, Proceedings of the Nato Advanced Study Institute, Cargese, 1979*, eds. G. 't Hooft *et al.*, (Plenum Press, NY, 1980); E. Witten, Phys. Lett. B **105**, 267 (1981).

[21] J. Ellis, S. Kelley and D. V. Nanopoulos, Phys. Lett. **B260** (1991) 131; U. Amaldi, W. de Boer and H. Furstenau, Phys. Lett. **B260** (1991) 447; C. Giunti, C. W. Kim and U. W. Lee, Mod. Phys. Lett. **A6** (1991) 1745.

[22] J. R. Ellis and D. Ross, Phys. Lett. B **506** (2001) 331 [arXiv:hep-ph/0012067].

[23] J. R. Ellis, J. S. Hagelin, D. V. Nanopoulos, K. A. Olive and M. Srednicki, Nucl. Phys. B **238** (1984) 453; see also H. Goldberg, Phys. Rev. Lett. **50**, 1419 (1983).

[24] D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148** (2003) 175 [arXiv:astro-ph/0302209].

[25] G. W. Bennett *et al.* [Muon g-2 Collaboration], Phys. Rev. Lett. **92** (2004) 161802 [arXiv:hep-ex/0401008]; M. Davier, S. Eidelman, A. Hocker and Z. Zhang, Eur. Phys. J. C **31** (2003) 503 [arXiv:hep-ph/0308213].

[26] J. R. Ellis, S. Heinemeyer, K. A. Olive and G. Weiglein, Phys. Lett. B **515** (2001) 348 [arXiv:hep-ph/0105061].

[27] J. R. Ellis, S. Heinemeyer, K. A. Olive and G. Weiglein, JHEP **0301** (2003) 006 [arXiv:hep-ph/0211206].

[28] J. R. Ellis, J. S. Lee and A. Pilaftsis, Phys. Rev. D **70** (2004) 075010 [arXiv:hep-ph/0404167].

[29] E. Accomando *et al.* [CLIC Physics Working Group], arXiv:hep-ph/0412251.

[30] J. R. Ellis, K. A. Olive and V. C. Spanos, Phys. Lett. B **624** (2005) 47 [arXiv:hep-ph/0504196].

[31] J. R. Ellis, S. Heinemeyer, K. A. Olive and G. Weiglein, JHEP **0502** (2005) 013 [arXiv:hep-ph/0411216] and arXiv:hep-ph/0508169.

[32] See, for example: J. R. Ellis, K. A. Olive, Y. Santoso and V. C. Spanos, Phys. Lett. B **588** (2004) 7 [arXiv:hep-ph/0312262] and Phys. Rev. D **70** (2004) 055005 [arXiv:hep-ph/0405110].

[33] J. L. Feng and B. T. Smith, Phys. Rev. D **71** (2005) 015004 [Erratum-ibid. D **71** (2005) 0109904] [arXiv:hep-ph/0409278].

[34] K. Hamaguchi, Y. Kuno, T. Nakaya and M. M. Nojiri, Phys. Rev. D **70** (2004) 115007 [arXiv:hep-ph/0409248]. A. Brandenburg, L. Covi, K. Hamaguchi, L. Roszkowski and F. D. Steffen, Phys. Lett. B **617** (2005) 99 [arXiv:hep-ph/0501287]; F. D. Steffen, arXiv:hep-ph/0507003.

[35] A. De Roeck, J. R. Ellis, F. Gianotti, F. Moortgat, K. A. Olive and L. Pape, arXiv:hep-ph/0508198.

[36] M. Takeda *et al.*, Astropart. Phys. **19** (2003) 447 [arXiv:astro-ph/0209422].

[37] R. U. Abbasi *et al.* [High Resolution Fly's Eye Collaboration], Phys. Rev. Lett. **92** (2004) 151101 [arXiv:astro-ph/0208243].

[38] P. Sommers [Pierre Auger Collaboration], arXiv:astro-ph/0507150.

[39] See, for example: S. Sarkar, Acta Phys. Polon. B **35** (2004) 351 [arXiv:hep-ph/0312223]; J. R. Ellis, V. E. Mayes and D. V. Nanopoulos, Phys. Rev. D **70** (2004) 075015 [arXiv:hep-ph/0403144] and in preparation.

[40] D. Kharzeev, E. Levin and M. Nardi, Nucl. Phys. A **747** (2005) 609 [arXiv:hep-ph/0408050].

[41] H. Kowalski and D. Teaney, Phys. Rev. D **68** (2003) 114005 [arXiv:hep-ph/0304189].

[42] I. Arsene *et al.* [BRAHMS Collaboration], Nucl. Phys. A **757** (2005) 1 [arXiv:nucl-ex/0410020].

[43] D. Kharzeev, Y. V. Kovchegov and K. Tuchin, Phys. Lett. B **599** (2004) 23 [arXiv:hep-ph/0405045].

[44] K. Golec-Biernat and M. Wusthoff, Phys. Rev. D **59** (1999) 014017 [arXiv:hep-ph/9807513] and Phys. Rev. D **60** (1999) 114023 [arXiv:hep-ph/9903358].

[45] J. R. Ellis, J. S. Lee and A. Pilaftsis, Phys. Rev. D **71** (2005) 075007 [arXiv:hep-ph/0502251].

[46] M. G. Albrow *et al.*, *FP420: An R and D proposal to investigate the feasibility of installing proton tagging detectors in the 420-m region at LHC*, CERN-LHCC-2005-025.

1 Working group 1: Parton Density Functions

List of participants in working group:

S. Forte, A. Glazov, S. Moch, S. Alekhin, G. Altarelli, J. Andersen, R. D. Ball, J. Blümlein, H. Böttcher, T. Carli, M. Ciafaloni, D. Colferai, A. Cooper-Sarkar, G. Corcella, L. Del Debbio, G. Dissertori, J. Feltesse, A. Guffanti, C. Gwenlan, J. Huston, G. Ingelman, M. Klein, J. I. Latorre, T. Laštovička, G. Laštovička-Medin, L. Magnea, A. Piccione, J. Pumplin, V. Ravindran, B. Reisert, J. Rojo, A. Sabio Vera, G. P. Salam, F. Siegert, A. Staśto, H. Stenzel, C. Targett-Adams, R.S. Thorne, A. Tricoli, J.A.M. Vermaseren, A. Vogt
C. Anastasiou, M. Cacciari, G. Chachamis, V. Chekelian, J. Cole, T. Falter, S. Ferrag, E. Gardi, M. Grazzini, V. Guzey, B. Heinemann, H. Jung, V. Kolhinen, K. Kutak, E. Lobodzinska, M. Lublinsky, F. Maltoni, S. Munier, K. Peters, C. Salgado, T. Schorner, M. Skrzypek, G. Steinbrueck, J. Stirling, A. Tonazzo, H. Weigert, M. Whalley

Introduction to Parton Density Functions

CONVENERS:
M. Dittmar [1], *S. Forte* [2], *A. Glazov* [3], *S. Moch* [4]
CONTRIBUTING AUTHORS:
S. Alekhin [5], *G. Altarelli* [6,7], *J. Andersen* [8], *R. D. Ball* [9], *J. Blümlein* [4], *H. Böttcher* [4], *T. Carli* [10], *M. Ciafaloni* [11], *D. Colferai* [11], *A. Cooper-Sarkar* [12], *G. Corcella* [6], *L. Del Debbio* [6,9], *G. Dissertori* [1], *J. Feltesse*[13], *A. Guffanti* [4], *C. Gwenlan* [12], *J. Huston* [14], *G. Ingelman* [15], *M. Klein* [4], *J. I. Latorre* [16], *T. Laštovička* [10], *G. Laštovička-Medin* [17], *L. Magnea* [18], *A. Piccione* [18], *J. Pumplin* [14], *V. Ravindran* [19], *B. Reisert* [20], *J. Rojo* [16], *A. Sabio Vera* [21], *G. P. Salam* [22], *F. Siegert* [10], *A. Staśto* [23], *H. Stenzel* [24], *C. Targett-Adams* [25], *R.S. Thorne* [8], *A. Tricoli*[12], *J.A.M. Vermaseren* [26], *A. Vogt* [27]

[1] Institute for Particle Physics, ETH-Zürich Hönggerberg, CH 8093 Zürich, Switzerland
[2] Dipartimento di Fisica, Universitá di Milano, INFN Sezione di Milano, Via Celoria 16, I 20133 Milan, Italy
[3] DESY, Notkestrasse 85, D 22603 Hamburg, Germany
[4] DESY, Platanenallee 6, D 15738 Zeuthen, Germany
[5] Institute for High Energy Physics, 142284 Protvino, Russia
[6] CERN, Department of Physics, Theory Division, CH 1211 Geneva 23, Switzerland
[7] Dipartimento di Fisica "E.Amaldi", Università Roma Tre and INFN, Sezione di Roma Tre, via della Vasca Navale 84, I 00146 Roma, Italy
[8] Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge, CB3 0HE, UK
[9] School of Physics, University of Edinburgh, Edinburgh EH9 3JZ, UK
[10] CERN, Department of Physics, CH 1211 Geneva 23, Switzerland
[11] Dipartimento di Fisica, Università di Firenze and INFN, Sezione di Firenze, I 50019 Sesto Fiorentino, Italy
[12] Department of Physics, Nuclear and Astrophysics Lab., Keble Road, Oxford, OX1 3RH, UK
[13] DSM/DAPNIA, CEA, Centre d'Etudes de Saclay, F 91191 Gif-sur-Yvette, France
[14] Department of Physics and Astronomy, Michigan State University, E. Lansing, MI 48824, USA
[15] High Energy Physics, Uppsala University, Box 535, SE 75121 Uppsala, Sweden
[16] Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona, Diagonal 647, E 08028 Barcelona, Spain
[17] University of Podgorica, Cetinjski put bb, CNG 81000 Podgorica, Serbia and Montenegro
[18] Dipartimento di Fisica Teorica, Università di Torino and INFN Sezione di Torino, Via P. Giuria 1, I 10125 Torino, Italy
[19] Harish-Chandra Research Institute, Chhatnag Road, Jhunsi, Allahabad, India
[20] FNAL, Fermi National Accelerator Laboratory, Batavia, IL 60126, USA
[21] II. Institut für Theoretische Physik, Universität Hamburg, Luruper Chaussee 149, D 22761 Hamburg, Germany
[22] LPTHE, Universities of Paris VI and VII and CNRS, F 75005, Paris, France
[23] H. Niewodniczański Institute of Nuclear Physics, PL 31-342 Kraków, Poland
[24] II. Physikalisches Institut, Universität Giessen, Heinrich-Buff-Ring 16, D 35392 Giessen, Germany
[25] Department of Physics and Astronomy, UC London, Gower Street, London, WC1E 6BT, UK
[26] NIKHEF Theory Group, Kruislaan 409, NL 1098 SJ Amsterdam, The Netherlands
[27] IPPP, Department of Physics, Durham University, Durham DH1 3LE, UK

Abstract
We provide an assessment of the impact of parton distributions on the determination of LHC processes, and of the accuracy with which parton distributions (PDFs) can be extracted from data, in particular from current and forthcoming

HERA experiments. We give an overview of reference LHC processes and their associated PDF uncertainties, and study in detail W and Z production at the LHC. We discuss the precision which may be obtained from the analysis of existing HERA data, tests of consistency of HERA data from different experiments, and the combination of these data. We determine further improvements on PDFs which may be obtained from future HERA data (including measurements of F_L), and from combining present and future HERA data with present and future hadron collider data. We review the current status of knowledge of higher (NNLO) QCD corrections to perturbative evolution and deep-inelastic scattering, and provide reference results for their impact on parton evolution, and we briefly examine non-perturbative models for parton distributions. We discuss the state-of-the art in global parton fits, we assess the impact on them of various kinds of data and of theoretical corrections, by providing benchmarks of Alekhin and MRST parton distributions and a CTEQ analysis of parton fit stability, and we briefly presents proposals for alternative approaches to parton fitting. We summarize the status of large and small x resummation, by providing estimates of the impact of large x resummation on parton fits, and a comparison of different approaches to small x resummation, for which we also discuss numerical techniques.

The physics of parton distributions, especially within the context of deep-inelastic scattering (DIS), has been an active subject of detailed theoretical and experimental investigations since the origins of perturbative quantum chromodynamics (QCD), which, thanks to asymptotic freedom, allows one to determine perturbatively their scale dependence [1, 2, 3, 4, 5].

Since the advent of HERA, much progress has been made in determining the Parton Distribution Functions (PDFs) of the proton. A good knowledge of the PDFs is vital in order to make predictions for both Standard Model and beyond the Standard Model processes at hadronic colliders, specifically the LHC. Furthermore, PDFs must be known as precisely as possible in order to maximize the discovery potential for new physics at the LHC. Conversely, LHC data will lead to an improvement in the knowledge of PDFs.

The main aim of this document is to provide a state-of-the art assessment of the impact of parton distributions on the determination of LHC processes, and of the accuracy with which parton distributions can be extracted from data, in particular current and forthcoming HERA data.

In [6] we will set the stage by providing an overview of relevant LHC processes and a discussion of their experimental and theoretical accuracy. In [7] we will turn to the experimental determination of PDFs, and in particular examine the improvements to be expected from forthcoming measurements at HERA, as well as from analysis methods which allow one to combine HERA data with each other, and also with data from existing (Tevatron) and forthcoming (LHC) hadron colliders. In [8] we will discuss the state of the art in the extraction of parton distributions of the data, by first reviewing recent progress in higher-order QCD corrections and their impact on the extraction of PDFs, and then discussing and comparing the determination of PDFs from global fits. Finally, in [9] we will summarize the current status of resummed QCD computations which are not yet used in parton fits, but could lead to an improvement in the theoretical precision of PDF determinations.

Whereas we will aim at summarizing the state of the art, we will also provide several new results, benchmarks and predictions, which have been obtained within the framework of the HERA-LHC workshop.

Acknowledgments

A. Glazov thanks E. Rizvi, M. Klein, M. Cooper-Sarkar and C. Pascaud for help and many fruitful discussions. C. Gwenlan, A. Cooper-Sarka and C. Targett-Adams thank M. Klein and R. Thorne for providing the F_L predictions, as well as for useful discussions. T. Carli, G. Salam and F. Siegert thank Z. Nagy, M. H. Seymour, T. Schörner-Sadenius, P. Uwer and M. Wobisch for useful discussions on the grid technique, A. Vogt for discussion on moment-space techniques and Z. Nagy for help and support with NLOJET++. R. Thorne thanks S. Alekhin for collaboration on the project of obtaining the benchmark parton distributions, for providing his benchmark partons and for many useful exchanges. J. Huston and J. Pumplin thank W.K. Tung and D. Stump for collaboration on the presented research work.

S. Moch acknowledges partial support by the Helmholtz Gemeinschaft under contract VH-NG-105 and by DFG Sonderforschungsbereich Transregio 9, Computergestützte Theoretische Physik. J. Blümlein acknowledges partial support by DFG Sonderforschungsbereich Transregio 9, Computergestützte Theoretische Physik. C. Gwenlan acknowledges support by PPARC. F. Siegert acknowledges support by the CERN Summer Student Program. R.S. Thorne acknowledges support by the Royal Society as University Research Fellow. J. Huston and J. Pumplin acknowledge support by the National Science Foundation. J.R. Andersen acknowledges support from PPARC under contract PPA/P/S/2003/00281. A. Sabio Vera acknowledges suppoprt from the Alexander von Humboldt foundation.

References

[1] Gross, D. J. and Wilczek, F., Phys. Rev. Lett. **30**, 1343 (1973).

[2] Politzer, H. D., Phys. Rev. Lett. **30**, 1346 (1973).

[3] Gross, D. J. and Wilczek, F., Phys. Rev. **D8**, 3633 (1973).

[4] Georgi, H. and Politzer, H. D., Phys. Rev. **D9**, 416 (1974).

[5] Gross, D. J. and Wilczek, F., Phys. Rev. **D9**, 980 (1974).

[6] M. Dittmar et al., *Lhc final states and their potential experimental and theoretical accuracies*. These proceedings.

[7] A. Glazov et al., *Experimental determination of parton distributions*. These proceedings.

[8] A. Glazov et al., *Glap evolution and parton fits*. These proceedings.

[9] G. Altarelli et al., *Resummation*. These proceedings.

LHC final states and their potential experimental and theoretical accuracies

Amanda Cooper-Sarkar, Michael Dittmar, Günther Dissertori Claire Gwenlan, Hasko Stenzel, Alessandro Tricoli

1 LHC final states and their potential experimental and theoretical accuracies [1]

1.1 Introduction

Cross section calculations and experimental simulations for many LHC reactions, within the Standard Model and for many new physics scenarios have been performed during the last 20 years. These studies demonstrate how various final states might eventually be selected above Standard Model backgrounds and indicate the potential statistical significance of such measurements. In general, these studies assumed that the uncertainties from various sources, like the PDF uncertainties, the experimental uncertainties and the various signal and background Monte Carlo simulations will eventually be controlled with uncertainties small compared to the expected statistical significance. This is the obvious approach for many so called discovery channels with clean and easy signatures and relatively small cross sections.

However, during the last years many new and more complicated signatures, which require more sophisticated selection criteria, have been discussed. These studies indicate the possibility to perform more ambitious searches for new physics and for precise Standard Model tests, which would increase the physics potential of the LHC experiments. Most of these studies concentrate on the statistical significance only and potential systematic limitations are rarely discussed.

In order to close this gap from previous LHC studies, questions related to the systematic limits of cross section measurements from PDF uncertainties, from imperfect Standard Model Monte Carlo simulations, from various QCD uncertainties and from the efficiency and luminosity uncertainties were discussed within the PDF working group of this first HERA-LHC workshop. The goal of the studies presented during the subgroup meetings during the 2004/5 HERA LHC workshop provide some answers to questions related to these systematic limitations. In particular, we have discussed potential experimental and theoretical uncertainties for various Standard Model signal cross sections at the LHC. Some results on the experimental systematics, on experimental and theoretical uncertainties for the inclusive W, Z and for diboson production, especially related to uncertainties from PDF's and from higher order QCD calculations are described in the following sections.

While it was not possible to investigate the consequences for various aspects of the LHC physics potential in detail, it is important to keep in mind that many of these Standard Model reactions are also important backgrounds in the search for the Higgs and other exotic phenomena. Obviously, the consequences from these unavoidable systematic uncertainties need to be investigated in more detail during the coming years.

1.2 Measuring and interpreting cross sections at the LHC [2]

The LHC is often called a machine to make discoveries. However, after many years of detailed LHC simulations, it seems clear that relatively few signatures exist, which do not involve cross section measurements for signals and the various backgrounds. Thus, one expects that cross section measurements for a large variety of well defined reactions and their interpretation within or perhaps beyond the Standard Model will be one of the main task of the LHC physics program.

[1] Subsection coordinator: Michael Dittmar
[2] Contributing author: Michael Dittmar

While it is relatively easy to estimate the statistical precision of a particular measurement as a function of the luminosity, estimates of potential systematic errors are much more complicated. Furthmore, as almost nobody wants to know about systematic limitations of future experiments, detailed studies are not rewarding. Nevertheless, realistic estimates of such systematic errors are relevant, as they might allow the LHC community to concentrate their efforts on the areas where current systematic errors, like the ones which are related to uncertainties from Parton Distribution Functions (PDF) or the ones from missing higher order QCD calculations, can still be improved during the next years.

In order to address the question of systematics, it is useful to start with the basics of cross section measurements. Using some clever criteria a particular signature is separated from the data sample and the surviving $\mathrm{N_{observed}}$ events can be counted. Backgrounds, $\mathrm{N_{background}}$, from various sources have to be estimated either using the data or some Monte Carlo estimates. The number of signal events, $\mathrm{N_{signal}}$, is then obtained from the difference. In order to turn this experimental number of signal events into a measurement one has to apply a correction for the efficiency. This experimental number can now be compared with the product of the theoretical production cross section for the considered process and the corresponding Luminosity. For a measurement at a hadron collider, like the LHC, processes are calculated on the basis of quark and gluon luminosities which are obtained from the proton-proton luminosity "folded" with the PDF's.

In order to estimate potential systematic errors one needs to examine carefully the various ingredients to the cross section measurement and their interpretation. First, a measurement can only be as good as the impact from of the background uncertainties, which depend on the optimized signal to background ratio. Next, the experimental efficiency uncertainty depends on many subdetectors and their actual real time performance. While this can only be known exactly from real data, one can use the systematic error estimates from previous experiments in order to guess the size of similar error sources for the future LHC experiments. We are furthermore confronted with uncertainties from the PDF's and from the proton-proton luminosity. If one considers all these areas as essentially experimental, then one should assign uncertainties originating from imperfect knowledge of signal and background cross sections as theoretical.

Before we try to estimate the various systematic errors in the following subsections, we believe that it is important to keep in mind that particular studies need not to be much more detailed than the largest and limiting uncertainty, coming from either the experimental or the theoretical area. Thus, one should not waste too much time, in order to achieve the smalled possible uncertainty in one particular area. Instead, one should try first to reduce the most important error sources and if one accepts the "work division" between experimental and theoretical contributions, then one should simply try to be just a little more accurate than either the theoretical or the experimental colleagues.

1.2.1 Guessing experimental systematics for ATLAS and CMS

In order to guess experimental uncertainties, without doing lengthy and detailed Monte Carlo studies, it seems useful to start with some simple and optimistic assumptions about ATLAS and CMS[3].

First of all, one should assume that both experiments can actually operate as planned in their proposals. As the expected performance goals are rather similar for both detectors the following list of measurement capabilities looks as a reasonable first guess.

- Isolated electrons, muons and photons with a transverse momentum above 20 GeV and a pseudorapidity η with $|\eta| \leq 2.5$ are measured with excellent accuracy and high (perhaps as large as 95% for some reactions) "homogeneous" efficiency. Within the pseudo rapidity coverage one can assume that experimentalists will perhaps be able, using the large statistics from leptonic W and Z decays, to control the efficiency for electrons and muons with a 1% accuracy. For simplicity, one

[3] Up to date performance of the ATLAS and CMS detectors and further detailed references can be found on the corresponding homepages `http://atlas.web.cern.ch/Atlas/` and `http://cmsinfo.cern.ch/Welcome.html/`

can also assume that these events will allow to control measurements with high energy photons to a similar accuracy. For theoretical studies one might thus assume that high p_t electrons, muons and photons and $|\eta| \leq 2.5$ are measured with a systematic uncertainty of $\pm$ 1% for each lepton (photon).

- Jets are much more difficult to measure. Optimistically one could assume that jets can be seen with good efficiency and angular accuracy if the jet transverse momentum is larger than 30 GeV and if their pseudo rapidity fulfills $|\eta| \leq 4.5$. The jet energy resolution is not easy to quantify, but numbers could be given using some "reasonable" assumptions like $\Delta E/E \approx 100 - 150\%/\sqrt{E}$. For various measurements one want to know the uncertainty of the absolute jet energy scale. Various tools, like the decays of $W \to q\bar{q}$ in $t\bar{t}$ events or the photon-jet final state, might be used to calibrate either the mean value or the maximum to reasonably good accuracy. We believe that only detailed studies of the particular signature will allow a quantitative estimate of the uncertainties related to the jet energy scale measurements.
- The tagging of b–flavoured jets can be done, but the efficiency depends strongly on the potential backgrounds. Systematic efficiency uncertainties for the b–tagging are difficult to quantify but it seems that, in the absence of a new method, relative b-tagging uncertainties below $\pm$ 5% are almost impossible to achieve.

With this baseline LHC detector capabilities, it seems useful to divide the various high q^2 LHC reactions into essentially five different non overlapping categories. Such a devision can be used to make some reasonable accurate estimates of the different systematics.

- Drell–Yan type lepton pair final states. This includes on– and off–shell W and Z decays.
- γ–jet and $\gamma\gamma X$ final states.
- Diboson events of the type WW, WZ, ZZ, $W\gamma$ with leptonic decays of the W and Z bosons. One might consider to include the Standard Model Higgs signatures into this group of signatures.
- Events with top quarks in the final state, identified with at least one isolated lepton.
- Hadronic final states with up to n(=2,3 ..) Jets and different p_t and mass.

With this "grouping" of experimental final states, one can now start to analyze the different potential error sources. Where possible, one can try to define and use relative measurements of various reactions such that some systematic errors will simply cancel.

Starting with the resonant W and Z production with leptonic decays, several million of clean events will be collected quickly, resulting in relative statistical errors well below $\pm 1\%$. Theoretical calculations for these reactions are well advanced and these reactions are among the best understood LHC final states allowing to build the most accurate LHC Monte Carlo generators. Furthermore, some of the experimental uncertainties can be reduced considerably if ratio measurements of cross section, such as W^+/W^- and Z/W, are performed. The similarities in the production mechanism should also allow to reduce theoretical uncertainties for such ratios. The experimental counting accuracy of W and Z events, which includes background and efficiency corrections, might achieve eventually uncertainties of 1% or slightly better for cross section ratios.

Furthermore, it seems that the shape of the p_t distribution of the Z, using the decay into electron pairs ($pp \to ZX \to e{+}e^-X$), can be determined with relative accuracies of much less than 1%. This distribution, shown in figure 1, can be used to tune the Monte Carlo description of this particular process. This tuning of the Monte Carlo can than be used almost directly to predict theoretically also the W p_t spectrum, and the p_t spectrum for high mass Drell-Yan lepton pair events. Once an accurate model description of these Standard Model reactions is achieved one might use these insights also to predict the p_t spectrum of other well defined final states.

From all the various high q^2 reactions, the inclusive production of W and Z events is known to be the theoretically best understood and best experimentally measurable LHC reaction. Consequently, the idea to use these simple well defined final states as the LHC cross section normalisation tool, or standard

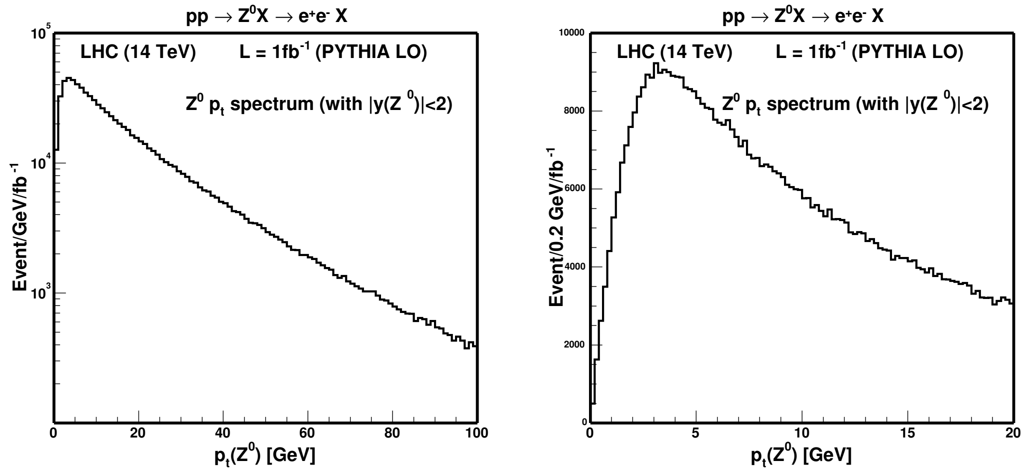


Fig. 1: Simple simulation of a potential measurement of the Z p_t spectrum, possible with a luminosity of only 1 fb^{-1}. Who will be able to predict this p_t spectrum in all its beauty and with similar accuracy?

candle was described first in reference [1]. This study indicated that the W and Z production might result in a precise and simple parton luminosity monitor. In addition, these reactions can also be used to improve the relative knowledge of the PDF's. In fact, if one gives up on the idea to measure absolute cross sections, the relative parton luminosity can in principle be determined with relative uncertainties well below $\pm 5\%$, the previously expected possible limit for any absolut proton-proton luminosity normalisation procdure.

In summary, one can estimate that it should be possible to reduce experimental uncertainties for Drell-Yan processes to systematic uncertainties below $\pm 5\%$, optimistically one might envisage an event counting accuracy of perhaps $\pm 1\%$, limited mainly from the lepton identification efficiency.

The next class of final states, which can be measured exclusively with leptons, are the diboson pair events with subsequent leptonic decays. Starting with the ZZ final state, we expect that the statistical accuracy will dominate the measurement for several years. Nevertheless, the systematic uncertainties of the measurement, based on four leptons, should in principle be possible with relative errors of a few % only.

The production of WZ and WW involves unmeasurable neutrinos. Thus, experimentally only an indirect and incomplete determination of the kinematics of the final states is possible and very detailed simulations with precise Monte Carlo generators are required for the interpretation of these final state. It seems that a measurement of the event counting with an accuracy below $\pm 5\%$, due to efficiency uncertainties from the selection alone, to be highly non trivial. Nevertheless, if the measurements and the interpretations can be done relative to the W and Z resonance production, some uncertainties from the lepton identification efficiency, from the PDF and from the theoretical calculation can perhaps be reduced. Without going into detailed studies for each channel, one could try to assume that a systematic uncertainty of $\pm 5\%$ might be defined as a goal. Similar characteristics and thus limitations can be expected for other diboson signatures.

The production cross section of top antitop quark pairs is large and several million of semileptonic tagged and relatively clean events ($pp \rightarrow t\bar{t} \rightarrow WbWb$ identified with one leptonic W decay) can be expected. However, the signature involves several jets, some perhaps tagged as b–flavoured, and missing transverse momentum from the neutrino(s). The correct association of the various jets to the corresponding top quark is known to be extremely difficult, leading to large combinatorial backgrounds. Thus, it seems that, even if precise Monte Carlo generators will become eventually available, that systematic uncertainties smaller than 5-10% should not be expected. Consequently,we assume that top antitop backgrounds for a wide class of signals can not be determined with uncertainties smaller than 5-10%.

Measurements of so called “single” top quarks are even more difficult, as the cross section is smaller and larger backgrounds exist. Systematic errors will therefore always be larger than the one guessed for top-antitop pair production.

Finally, we can address the QCD jet production. Traditionally one measures and interprets the so called jet cross section as a function of p_t jet and the mass of the multi jet system using various rapidity intervals. With the steeply falling p_t jet spectrum and essentially no background, one will determine the differential spectrum such that only the slope has to be measured with good relative accuracy. If one is especially interested into the super high mass or high p_t events, then we expect that migrations due to jet mis-measurements and non Gaussian tails in the jet energy measurements will limit any measurement. A good guess might be that the LHC experiments can expect absolut normalisation uncertainties similar to the ones achieved with CDF and D0, corresponding to uncertainties of about $\pm$ 10-20%.

Are the above estimated systematic limits for the various measurements pessimistic, optimistic or simply realistic? Of course, only real experiments will tell during the coming LHC years. However, while some of these estimates will need perhaps some small modification, they could be used as a limit waiting to be improved during the coming years. Thus, some people full of ideas might take these numbers as a challenge, and discover and develop new methods that will improve these estimates. This guess of systematic limitations for LHC experiments could thus be considered as a “provocation”, which will stimulate activities to prove them wrong. In fact, if the experimental and theoretical communities could demonstrate why some of these “pessimistic” numbers are wrong the future real LHC measurements will obviously benefit from the required efforts to develope better Monte Carlo programs and better experimental methods.

The following summary from a variety of experimental results from previous high energy collider experiments might help to quantify particular areas of concern for the LHC measurements. These previous measurements can thus be used as a starting point for an LHC experimenter, who can study and explain why the corresponding errors at LHC will be smaller or larger.

1.2.2 Learning from previous collider experiments

It is broadly accepted, due to the huge hadronic interaction rate and the short bunch crossing time, that the experimental conditions at the LHC will be similar or worse than the ones at the Tevatron collider. One experimental answer was to improve the granularity, speed and accuracy of the different detector elements accordingly. Still, no matter how well an experiment can be realized, the LHC conditions to do experiments will be much more difficult than at LEP or any hypothetical future high energy e^+e^- collider. One important reason is the large theoretical uncertainty, which prevents to make signal and background Monte Carlos with accuracies similar to the ones which were used at LEP.

Thus, we can safely expect that systematic errors at LHC experiments will be larger than the corresponding ones from LEP and that the Tevatron experience can be used as a first guess.

- Measurement of $\sigma\times$BR for W and Z production from CDF [2] and D0 [3]:
  The CDF collaboration has presented a high statistics measurement with electrons and muons. Similar systematic errors of about $\pm$ 2% were achieved for efficiency and thus the event counting with electrons and muons. The error was reduced to $\pm$ 1.4% for the ratio measurement where some lepton identification efficiencies cancel. Similar errors about $\times$ 1.5-2 larger have been obtained by the corresponding measurements from the D0 experiment.
- Measurement of the cross section for $p\bar{p} \to Z\gamma(\gamma)$ from D0 [4]:
  A total of 138 $ee\gamma$ and 152 $\mu\mu\gamma$ candidate events were selected. The background was estimated to be about 10% with a systematic uncertainty of $\pm$ 10-15%, mainly from γ-jet misidentification. Using Monte Carlo and a large sample of inclusive Z events, the efficiency uncertainty has

been estimated to be $\approx 5\%$ and when the data were used in comparison with the Standard Model prediction another uncertainty of 3.3% originating from PDF's was added.

- Measurement of the $p\bar{p} \to t\bar{t}$ production cross section from CDF [5]
  A recent CDF measurement, using 197 pb^{-1}, obtained a cross section (in pb) of 7.0 +2.4 (-2.1) from statistics. This should be comapred with +1.7 (-1.2) from systematics, which includes ± 0.4 from the luminosity measurement. Thus, uncertainties from efficiency and background are roughly $\pm 20\%$. It is expected that some of the uncertainties can be reduced with the expected 10 fold luminosity increase such that the systematic error will eventually decrease to about $\pm$ 10%, sufficient to be better than the expected theoretical error of $\pm$ 15%.
- A search for Supersymmetry with b-tagged jets from CDF [6]:
  This study, using single and double b-tagged events was consistent with background only. However, it was claimed that the background uncertainty was dominated by the systematic error, which probably originated mostly from the b tagging efficiency and the misidentification of b-flavoured jets. The numbers given were 16.4$\pm$ 3.7 events (3.15 from systematics) for the single b-tagged events and 2.6$\pm$0.7 events (0.66 from systematics) for the double b-tagged events. These errors originate mainly from the b-tagging efficiency uncertainties, which are found to roughly $\pm$ 20-25% for this study of rare events.
- Some "random" selection of recent e^+e^- measurements:
  A recent measurement from ALEPH (LEP) of the W branching ratio to $q\bar{q}$ estimated a systematic uncertainty of about $\pm$ 0.2% [7]. This small uncertainty was possible because many additional constraints could be used.
  OPAL has reported a measurement of R_b at LEP II energies, with a systematic uncertainty of $\pm$ 3.7%. Even though this uncertainty could in principle be reduced with higher statistics, one can use it as an indication on how large efficiency uncertainties from b-tagging are already with clean experimental conditions [8]
  Recently, ALEPH and DELPHI have presented cross section measurements for $e^+e^- \to \gamma\gamma$ with systematic errors between 2.2% (ALEPH) [9] and 1.1% (DELPHI) [10]. In both cases, the efficiency uncertainty, mainly from conversions, for this in principle easy signal was estimated to be roughly 1%. In the case of ALEPH an uncertainty of about $\pm$0.8% was found for the background correction.

Obviously, these measurements can only be used, in absence of anything better, as a most optimistic guess for possible systematic limitations at a hadron collider. One might conclude that the systematics from LEP experiments give (1) an optimistic limit for comparable signatures at the LHC and (2) that the results from CDF and D0 should indicate systematics which might be obtained realistically during the early LHC years.

Thus, in summary the following list might be used as a first order guess on achievable LHC systematics[4].

- "Isolated" muons, electrons and photons can be measured with a small momentum (energy) uncertainty and with an almost perfect angular resolution. The efficiency for $p_t \geq 20$ GeV and $|\eta| \leq 2.5$ will be "high" and can be controlled optimistically to $\pm$ 1%. Some straight forward selection criteria should reduce jet background to small or negligible levels.
- "Isolated" jets with a $p_t \geq 30$ GeV and $|\eta| \leq 4.5$ can be seen with high (veto) efficiency and a small uncertainty from the jet direction measurement. However, it will be very difficulty to measure the absolute jet energy scale and Non-Gaussian tails will limit the systematics if the jet energy scale is important.
- Measurements of the missing transverse momentum depend on the final state but will in general be a sum of the errors from the lepton and the jet accuracies.

[4]Reality will hopefully show new brilliant ideas, which combined with hard work will allow to obtain even smaller uncertainties.

Using these assumptions, the following "optimistic" experimental systematic errors can be used as a guideline:

1. Efficiency uncertainties for isolated leptons and photons with a p_t above 20 GeV can be estimated with a $\pm 1\%$ accuracy.
2. Efficiencies for tagging jets will be accurate to a few percent and the efficiency to tag b-flavoured jets will be known at best within $\pm 5\%$.
3. Backgrounds will be known, combining theoretical uncertainties and some experimental determinations, at best with a $\pm$5-10% accuracy. Thus, discovery signatures without narrow peaks require signal to background ratios larger than 0.25-0.5, if 5 σ discoveries are claimed. Obviously, for accurate cross section measurements, the signal to background ratio should be much larger.
4. In case of ratio measurements with isolated leptons, like $pp \to W^+/pp \to W^-$, relative errors between 0.5-1% should be possible. Furthermore, it seems that the measurement of the shape of Z p_t spectrum, using Z$\to e^+e^-$, will be possible with a systematic error much smaller than 1%. As the Z cross section is huge and clean we expect that this signature will become the best measurable final state and should allow to test a variety of production models with errors below $\pm$ 1%, thus challenging future QCD calculations for a long time.

1.3 Uncertainties on W and Z production at the LHC[5]

1.3.1 Introduction

At leading order (LO), W and Z production occur by the process, $q\bar{q} \to W/Z$, and the momentum fractions of the partons participating in this subprocess are given by $x_{1,2} = \frac{M}{\sqrt{s}} \exp(\pm y)$, where M is the centre of mass energy of the subprocess, $M = M_W$ or M_Z, $\sqrt{s}$ is the centre of mass energy of the reaction ($\sqrt{s} = 14$ TeV at the LHC) and $y = \frac{1}{2} \ln \frac{(E+p_l)}{(E-p_l)}$ gives the parton rapidity. The kinematic plane for LHC parton kinematics is shown in Fig. 2. Thus, at central rapidity, the participating partons have small momentum fractions, $x \sim 0.005$. Moving away from central rapidity sends one parton to lower x and one to higher x, but over the measurable rapidity range, $|y| < 2.5$, x values remain in the range, $10^{-4} < x < 0.1$. Thus, in contrast to the situation at the Tevatron, valence quarks are not involved, the scattering is happening between sea quarks. Furthermore, the high scale of the process $Q^2 = M^2 \sim 10,000$ GeV^2 ensures that the gluon is the dominant parton, see Fig. 2, so that these sea quarks have mostly been generated by the flavour blind $g \to q\bar{q}$ splitting process. Thus the precision of our knowledge of W and Z cross-sections at the LHC is crucially dependent on the uncertainty on the momentum distribution of the gluon.

HERA data have dramatically improved our knowledge of the gluon, as illustrated in Fig. 3, which shows W and Z rapidity spectra predicted from a global PDF fit which does not include the HERA data, compared to a fit including HERA data. The latter fit is the ZEUS-S global fit [11], whereas the former is a fit using the same fitting analysis but leaving out the ZEUS data. The full PDF uncertainties for both fits are calculated from the error PDF sets of the ZEUS-S analysis using LHAPDF [12] (see the contribution of M.Whalley to these proceedings). The predictions for the W/Z cross-sections, decaying to the lepton decay mode, are summarised in Table 1. The uncertainties in the predictions for these cross-sections have decreased from $\sim 16\%$ pre-HERA to $\sim 3.5\%$ post-HERA. The reason for this can be seen clearly in Fig. 4, where the sea and gluon distributions for the pre- and post-HERA fits are shown for several different Q^2 bins, together with their uncertainty bands. It is the dramatically increased precision in the low-x gluon PDF, feeding into increased precision in the low-x sea quarks, which has led to the increased precision on the predictions for W/Z production at the LHC.

Further evidence for the conclusion that the uncertainties on the gluon PDF at the input scale ($Q_0^2 = 7$ GeV^2, for ZEUS-S) are the major contributors to the uncertainty on the W/Z cross-sections at

[5]Contributing authors: Alessandro Tricoli, Amanda Cooper-Sarkar, Claire Gwenlan

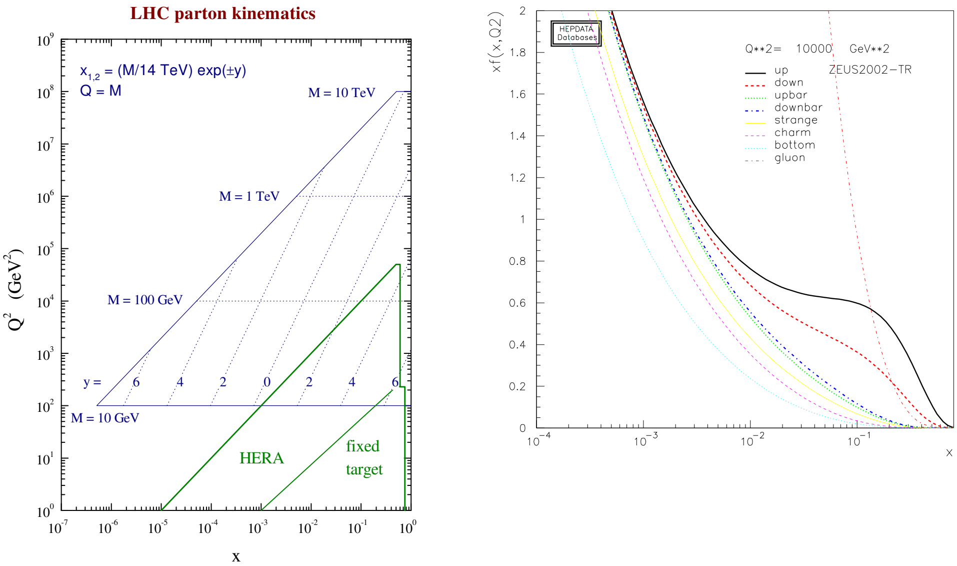


Fig. 2: Left plot: The LHC kinematic plane (thanks to James Stirling). Right plot: PDF distributions at $Q^2 = 10,000$ GeV^2.

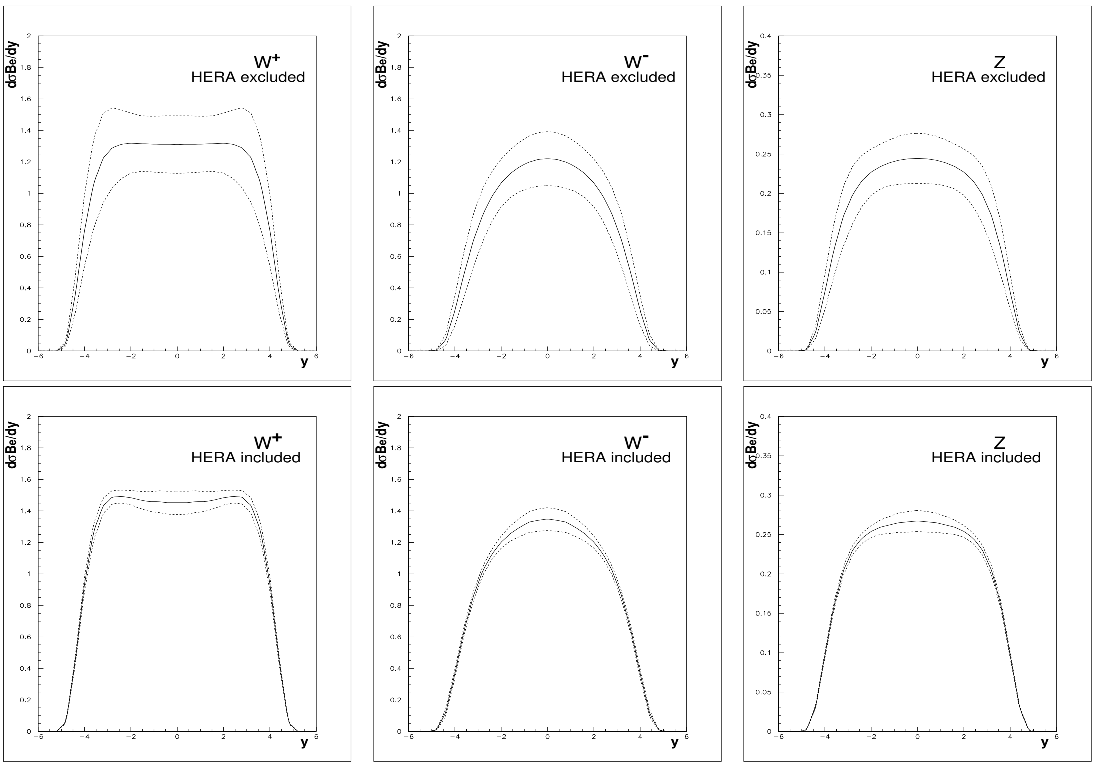


Fig. 3: LHC W^+, W^-, Z rapidity distributions and their PDF uncertainties (the full line shows the central value and the dashed lines show the spread of the uncertainty): Top Row: from the ZEUS-S global PDF analysis not including HERA data; left plot W^+; middle plot W^-; right plot Z: Bottom Row: from the ZEUS-S global PDF analysis including HERA data; left plot W^+; middle plot W^-; right plot Z

PDF Set	$\sigma(W^+).B(W^+ \to l^+\nu_l)$	$\sigma(W^-).B(W^- \to l^-\bar{\nu}_l)$	$\sigma(Z).B(Z \to l^+l^-)$
ZEUS-S no HERA	10.63 ± 1.73 nb	7.80 ± 1.18 nb	1.69 ± 0.23 nb
ZEUS-S	12.07 ± 0.41 nb	8.76 ± 0.30 nb	1.89 ± 0.06 nb
CTEQ6.1	11.66 ± 0.56 nb	8.58 ± 0.43 nb	1.92 ± 0.08 nb
MRST01	11.72 ± 0.23 nb	8.72 ± 0.16 nb	1.96 ± 0.03 nb

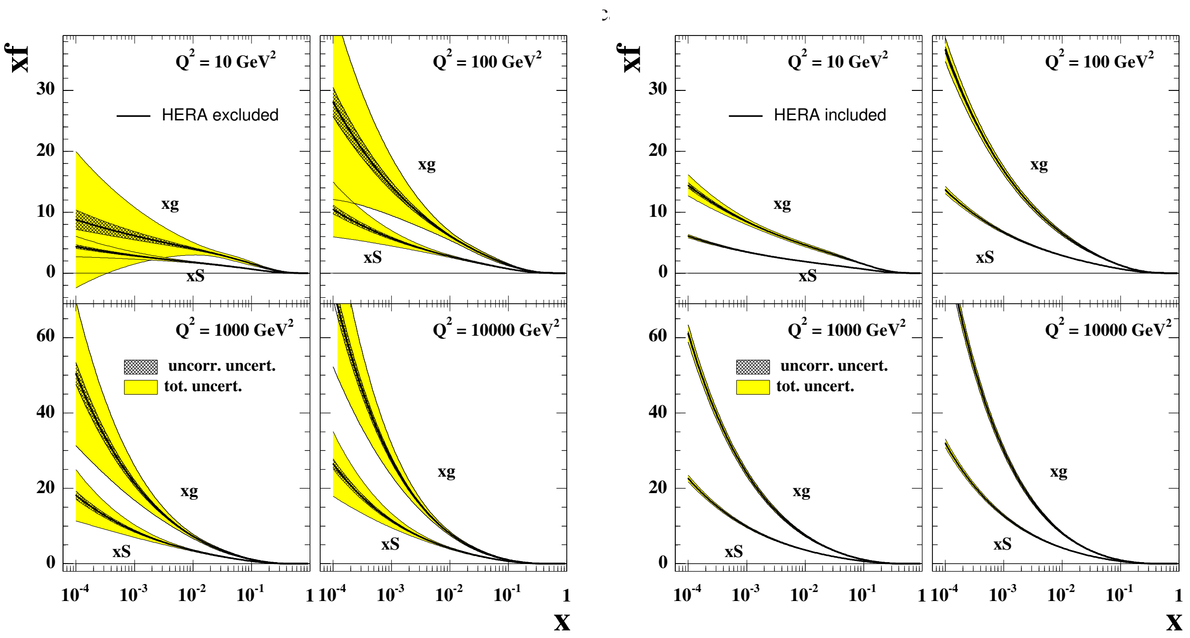


Fig. 4: Sea (xS) and gluon (xg) PDFs at various Q^2: left plot; from the ZEUS-S global PDF analysis not including HERA data; right plot: from the ZEUS-S global PDF analysis including HERA data. The inner cross-hatched error bands show the statistical and uncorrelated systematic uncertainty, the outer error bands show the total uncertainty including experimental correlated systematic uncertainties, normalisations and model uncertainty.

$Q^2 = M_W(M_Z)$, comes from decomposing the predictions down into their contributing eigenvectors. Fig 5 shows the dominant contributions to the total uncertainty from eigenvectors 3, 7, and 11 which are eigenvectors which are dominated by the parameters which control the low-x, mid-x and high-x, gluon respectively.

The post-HERA level of precision illustrated in Fig. 3 is taken for granted in modern analyses, such that W/Z production have been suggested as 'standard-candle' processes for luminosity measurement. However, when considering the PDF uncertainties on the Standard Model (SM) predictions it is necessary not only to consider the uncertainties of a particular PDF analysis, but also to compare PDF analyses. Fig. 6 compares the predictions for W^+ production for the ZEUS-S PDFs with those of the CTEQ6.1 [13] PDFs and the MRST01 [14] PDFs[6]. The corresponding W^+ cross-sections, for decay to leptonic mode are given in Table 1. Comparing the uncertainty at central rapidity, rather than the total cross-section, we see that the uncertainty estimates are rather larger: 5.2% for ZEUS-S; 8.7% for CTEQ6.1M and about 3.6% for MRST01. The difference in the central value between ZEUS-S and CTEQ6.1 is 3.5%. Thus the spread in the predictions of the different PDF sets is comparable to the uncertainty estimated by the individual analyses. Taking all of these analyses together the uncertainty at central rapidity is about 8%.

Since the PDF uncertainty feeding into the W^+, W^- and Z production is mostly coming from

[6]MRST01 PDFs are used because the full error analysis is available only for this PDF set.

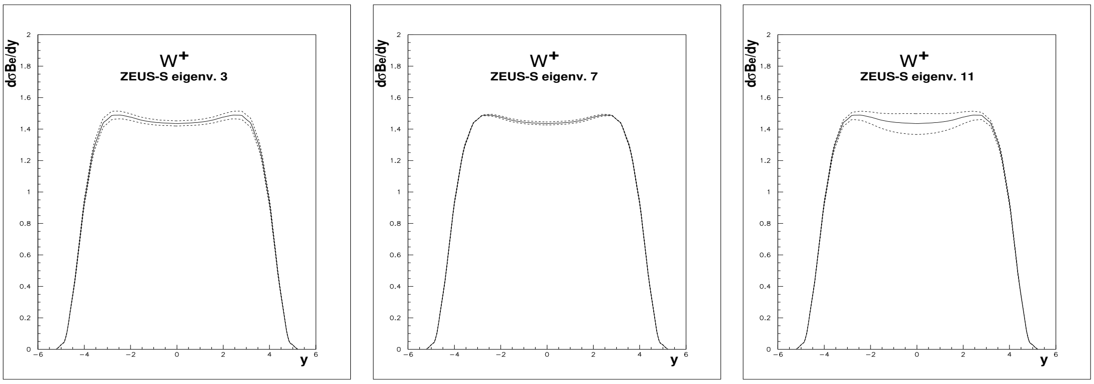


Fig. 5: LHC W^+ rapidity distributions and their PDF uncertainties due to the eigenvectors 3,7 and 11 of the ZEUS-S analysis.

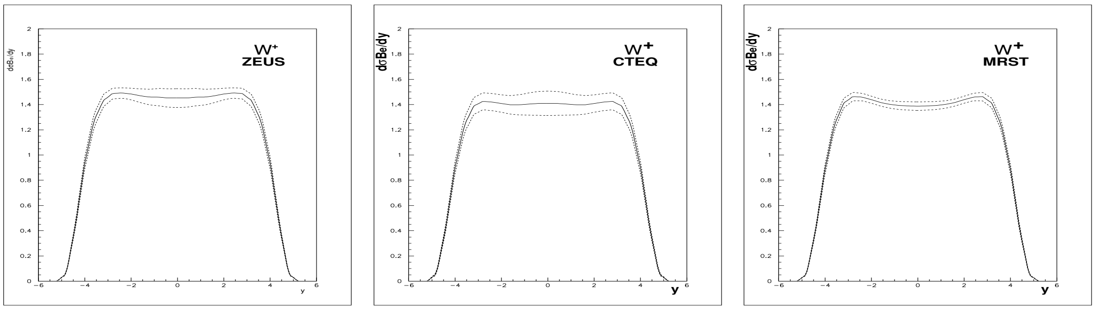


Fig. 6: LHC W^+ rapidity distributions and their PDF uncertainties: left plot, ZEUS-S PDFs; middle plot, CTEQ6.1 PDFs; right plot: MRST01 PDFs.

the gluon PDF, for all three processes, there is a strong correlation in their uncertainties, which can be removed by taking ratios. Fig. 7 shows the W asymmetry

$$A_W = (W^+ - W^-)/(W^+ + W^-).$$

for CTEQ6.1 PDFs, which have the largest uncertainties of published PDF sets. The PDF uncertainties on the asymmetry are very small in the measurable rapidity range. An eigenvector decomposition indicates that sensitivity to high-x u and d quark flavour distributions is now evident at large y. Even this residual flavour sensitivity can be removed by taking the ratio

$$A_{ZW} = Z/(W^+ + W^-)$$

as also shown in Fig. 7. This quantity is almost independent of PDF uncertainties. These quantities have been suggested as benchmarks for our understanding of Standard Model Physics at the LHC. However, whereas the Z rapidity distribution can be fully reconstructed from its decay leptons, this is not possible for the W rapidity distribution, because the leptonic decay channels which we use to identify the W's have missing neutrinos. Thus we actually measure the W's decay lepton rapidity spectra rather than the W rapidity spectra. The lower half of Fig. 7 shows the rapidity spectra for positive and negative leptons from W^+ and W^- decay and the lepton asymmetry,

$$A_l = (l^+ - l^-)/(l^+ + l^-).$$

A cut of, $p_{tl} > 25$ GeV, has been applied on the decay lepton, since it will not be possible to trigger on leptons with small p_{tl}. A particular lepton rapidity can be fed from a range of W rapidities so that the contributions of partons at different x values is smeared out in the lepton spectra, but the broad features of the W spectra and the sensitivity to the gluon parameters remain. The lepton asymmetry

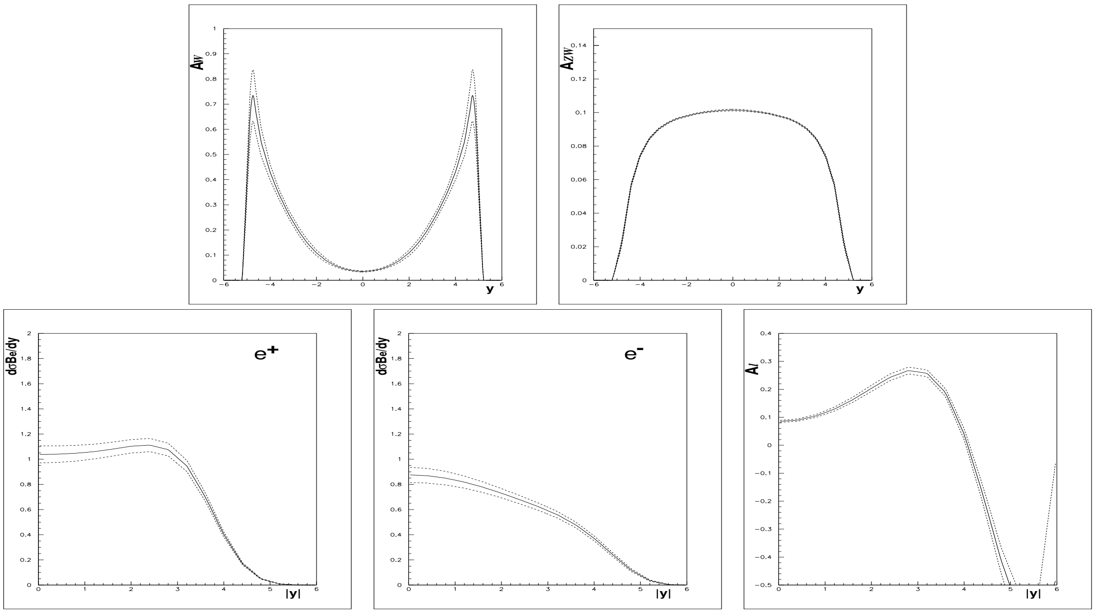


Fig. 7: Predictions for W, Z production at the LHC from the CTEQ6.1 PDFs. Top row: left plot, the W asymmetry, A_W; right plot, the ratio, A_{ZW}: Bottom row: left plot, decay e^+ rapidity spectrum; middle plot, decay e^- rapidity spectrum; right plot, lepton asymmetry, A_e

shows the change of sign at large y which is characteristic of the $V - A$ structure of the lepton decay. The cancellation of the uncertainties due to the gluon PDF is not so perfect in the lepton asymmetry as in the W asymmetry. Nevertheless in the measurable rapidity range sensitivity to PDF parameters is small. Correspondingly, the PDF uncertainties are also small (4%) and this quantity provides a suitable Standard Model benchmark.

In summary, these preliminary investigations indicate that PDF uncertainties on predictions for the W, Z rapidity spectra, using standard PDF sets which describe all modern data, have reached a precision of $\sim 8\%$. This may be good enough to consider using these processes as luminosity monitors. The predicted precision on ratios such as the lepton ratio, A_l, is better ($\sim 4\%$) and this measurement may be used as a SM benchmark. It is likely that this current level of uncertainty will have improved before the LHC turns on- see the contribution of C. Gwenlan ([15]) to these proceedings. The remainder of this contribution will be concerned with the question: how accurately can we measure these quantities and can we use the early LHC data to improve on the current level of uncertainty?

1.3.2 k-factor and PDF re-weighting

To investigate how well we can really measure W production we need to generate samples of Monte-Carlo (MC) data and pass them through a simulation of a detector. Various technical problems arise. Firstly, many physics studies are done with HERWIG (6.505) [16], which generates events at LO with parton showers to account for higher order effects. Distributions can be corrected from LO to NLO by k-factors which are applied as a function of the variable of interest. The use of HERWIG is gradually being superceded by MC@NLO (2.3) [17] but this is not yet implemented for all physics processes. Thus it is necessary to investigate how much bias is introduced by using HERWIG with k-factors. Secondly, to simulate the spread of current PDF uncertainties, it is necessary to run the MC with all of the eigenvector error sets of the PDF of interest. This would be unreasonably time-consuming. Thus the technique of PDF reweighting has been investigated.

One million $W \rightarrow e\nu_e$ events were generated using HERWIG (6.505). This corresponds to 43 hours of LHC running at low luminosity, $10fb^{-1}$. The events are split into W^+ and W^- events according to their Standard Model cross-section rates, 58%: 42% (the exact split depends on the input PDFs). These events are then weighted with k-factors, which are analytically calculated as the ratio of the NLO to LO cross-section as a function of rapidity for the same input PDF [18]. The resultant rapidity spectra for W^+, W^- are compared to rapidity spectra for $\sim 107,700$ events generated using MC@NLO(2.3) in

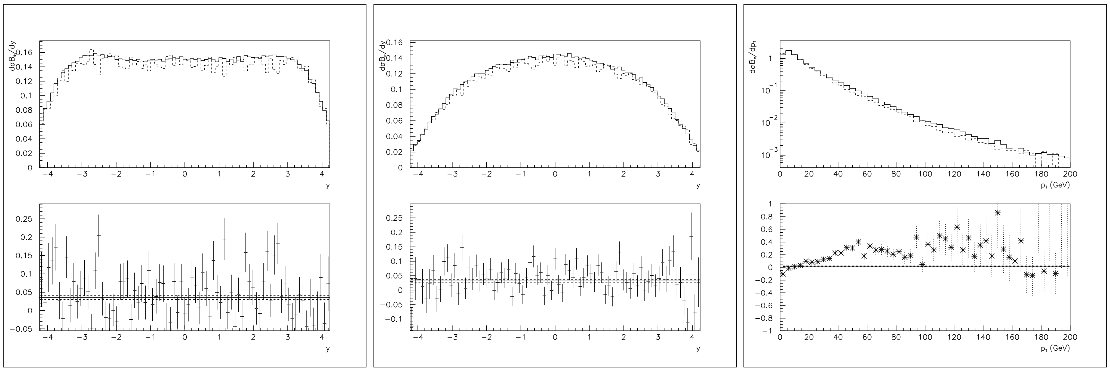


Fig. 8: Top Row: W rapidity and p_t spectra for events generated with HERWIG + k-Factors (full line), compared to those generated by MC@NLO (dashed line); left plot W^+ rapidity; middle plot W^- rapidity; right plot W^- p_t. Bottom row: the fractional differences of the spectra generated by HERWIG + k-factors and those generated by MC@NLO. The full line represents the weighted mean of these difference spectra and the dashed lines show its uncertainty

Fig 8[7]. The MRST02 PDFs were used for this investigation. The accuracy of this study is limited by the statistics of the MC@NLO generation. Nevertheless it is clear that HERWIG with k-factors does a good job of mimicking the NLO rapidity spectra. However, the normalisation is too high by 3.5%. This is not suprising since, unlike the analytic code, HERWIG is not a purely LO calculation, parton showering is also included. This normalisation difference is not too crucial since in an analysis on real data the MC will only be used to correct data from the detector level to the generator level. For this purpose, it is essential to model the shape of spectra to understand the effect of experimental cuts and smearing but not essential to model the overall normalisation perfectly. However, one should note that HERWIG with k-factors is not so successful in modelling the shape of the p_t spectra, as shown in the right hand plot of Fig. 8. This is hardly surprising, since at LO the W have no p_t and non-zero p_t for HERWIG is generated by parton showering, whereas for MC@NLO non-zero p_t originates from additional higher order processes which cannot be scaled from LO, where they are not present.

Suppose we generate W events with a particular PDF set: PDF set 1. Any one event has the hard scale, $Q^2 = M_W^2$, and two primary partons of flavours $flav_1$ and $flav_2$, with momentum fractions x_1, x_2 according to the distributions of PDF set 1. These momentum fractions are applicable to the hard process before the parton showers are implemented in backward evolution in the MC. One can then evaluate the probability of picking up the same flavoured partons with the same momentum fractions from an alternative PDF set, PDF set 2, at the same hard scale. Then the event weight is given by

$$\mathrm{PDF(re-weight)} = \frac{\mathrm{f_{PDF_2}(x_1, flav_1, Q^2).f_{PDF_2}(x_2, flav_2, Q^2)}}{\mathrm{f_{PDF_1}(x_1, flav_1, Q^2).f_{PDF_1}(x_2, flav_2, Q^2)}} \tag{1}$$

where $xf_{PDF}(x, flav, Q^2)$ is the parton momentum distribution for flavour, $flav$, at scale, Q^2, and momentum fraction, x. Fig. 9 compares the W^+ and W^- spectra for a million events generated using MRST02 as PDF set 1 and re-weighting to CTEQ6.1 as PDF set 2, with a million events which are directly generated with CTEQ6.1. Beneath the spectra the fractional difference between these distributions is shown. These difference spectra show that the reweighting is good to better than 1%, and there is no evidence of a y dependent bias. This has been checked for reweighting between MRST02, CTEQ6.1 and ZEUS-S PDFs. Since the uncertainties of any one analysis are similar in size to the differences between the analyses it is clear that the technique can be used to produce spectra for the eigenvector error PDF sets of each analysis and thus to simulate the full PDF uncertainties from a single set of MC generated events. Fig. 9 also shows a similar comparison for p_t spectra.

[7]In MC@NLO the hard emissions are treated by NLO computations, whereas soft/collinear emissions are handled by the MC simulation. In the matching procedure a fraction of events with negative weights is generated to avoid double counting. The event weights must be applied to the generated number of events before the effective number of events can be converted to an equivalent luminosity. The figure given is the effective number of events.

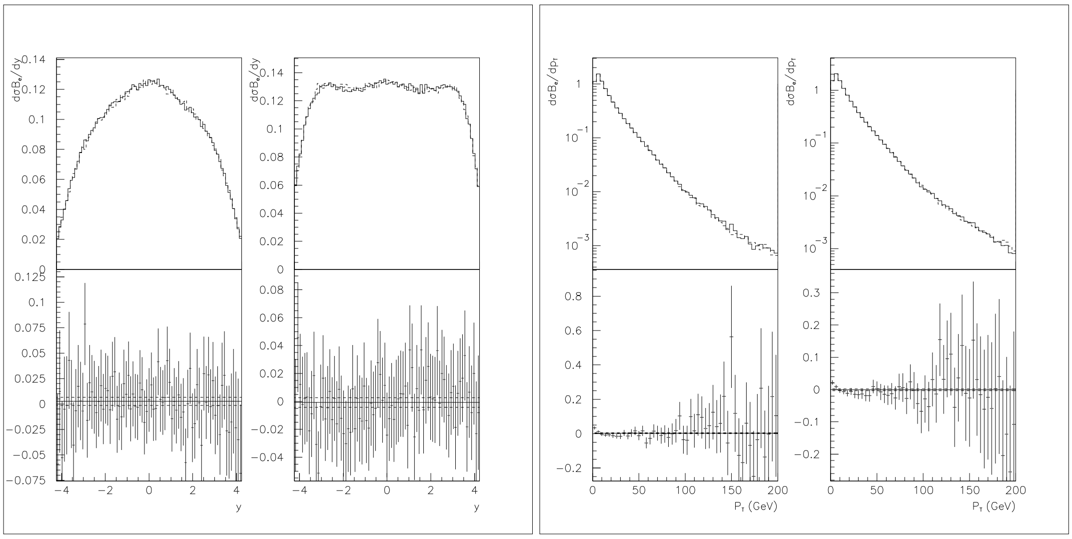


Fig. 9: Left side: W^- (left) and W^+ (right) rapidity spectra, for events generated with MRST02 PDFs reweighted to CTEQ6.1 PDFs (full line), compared to events generated directly with CTEQ6.1 PDFs (dashed line). The fractional difference between these spectra are also shown beneath the plots. The full line represents the weighted mean of these difference spectra and the dashed lines show its uncertainty. Right side: the same for p_t spectra.

1.3.3 Background Studies

To investigate the accuracy with which W events can be measured at the LHC it is necessary to make an estimate of the importance of background processes. We focus on W events which are identified through their decay to the $W \to e\, \nu_e$ channel. There are several processes which can be misidentified as $W \to e\nu_e$. These are: $W \to \tau\nu_\tau$, with τ decaying to the electron channel; $Z \to \tau^+\tau^-$ with at least one τ decaying to the electron channel (including the case when both τ's decay to the electron channel, but one electron is not identified); $Z \to e^+e^-$ with one electron not identified. We have generated one million events for each of these background processes, using HERWIG and CTEQ5L, and compared them to one million signal events generated with CTEQ6.1. We apply event selection criteria designed to eliminate the background preferentially. These criteria are:

- ATLFAST cuts (see Sec. 1.3.5)
- pseudorapidity, $|\eta| < 2.4$, to avoid bias at the edge of the measurable rapidity range
- $p_{te} > 25$ GeV, high p_t is necessary for electron triggering
- missing $E_t > 25$ GeV, the ν_e in a signal event will have a correspondingly large missing E_t
- no reconstructed jets in the event with $p_t > 30$ GeV, to discriminate against QCD background
- recoil on the transverse plane $p_t^{recoil} < 20$ GeV, to discriminate against QCD background

Table 2 gives the percentage of background with respect to signal, calculated using the known relative cross-sections of these processes, as each of these cuts is applied. After, the cuts have been applied the background from these processes is negligible. However, there are limitations on this study from the fact that in real data there will be further QCD backgrounds from $2 \to 2$ processes involving $q, \bar{q}, g$ in which a final state $\pi^0 \to \gamma\gamma$ decay mimics a single electron. A preliminary study applying the selection criteria to MC generated QCD events suggests that this background is negligible, but the exact level of QCD background cannot be accurately estimated without passing a very large number of events though a full detector simulation, which is beyond the scope of the current contribution.

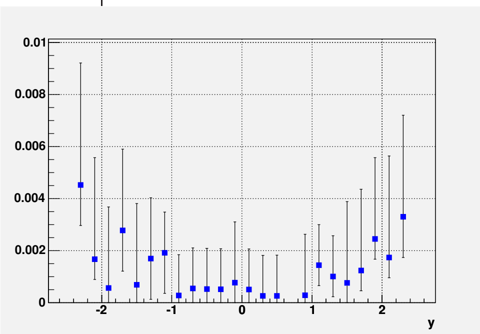
1.3.4 Charge misidentification

Clearly charge misidentification could distort the lepton rapidity spectra and dilute the asymmetry A_l.

$$A_{true} = \frac{A_{raw} - F^+ + F^-}{1 - F^- - F^+}$$

Table 2: Reduction of signal and background due to cuts

Cut	$W \to e\nu_e$		$Z \to \tau^+\tau^-$		$Z \to e^+e^-$		$W \to \tau\nu_\tau$	
	e^+	e^-	e^+	e^-	e^+	e^-	e^+	e^-
ATLFAST cuts	382,902	264,415	5.5%	7.9%	34.7%	50.3%	14.8%	14.9%
$\vert\eta\vert < 2.4$	367,815	255,514	5.5%	7.8%	34.3%	49.4%	14.7%	14.8%
$p_{te} > 25$ GeV	252,410	194,562	0.6%	0.7%	12.7%	16.2%	2.2%	2.3%
$p_{tmiss} > 25$ GeV	212,967	166,793	0.2%	0.2%	0.1%	0.2%	1.6%	1.6%
No jets with $P_t > 30$ GeV	187,634	147,415	0.1%	0.1%	0.1%	0.1%	1.2%	1.2%
$p_t^{recoil} < 20$ GeV	159,873	125,003	0.1%	0.1%	0.0%	0.0%	1.2%	1.2%



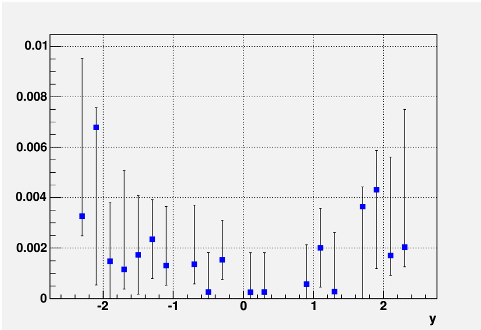


Fig. 10: The rates of charge misidentification as a function of rapidity for e^- misidentified as e^+ (left), e^+ misidentifed as e^- (right).

where A_{raw} is the measured asymmetry, A_{true} is the true asymmetry, F^- is the rate of true e^- misidentified as e^+ and F^+ is the rate of true e^+ misidentified as e^-. To make an estimate of the importance of charge misidentification we use a sample of $Z \to e^+e^-$ events generated by HERWIG with CTEQ5L and passed through a full simulation of the ATLAS detector. Events with two or more charged electromagnetic objects in the EM calorimeter are then selected and subject to the cuts; $|\eta| < 2.5$, $p_{te} > 25$ GeV, as usual and, $E/p < 2$, for bremsstrahlung rejection. We then look for the charged electromagnetic pair with invariant mass closest to M_Z and impose the cut, $60 < M_Z < 120$ GeV. Then we tag the charge of the better reconstructed lepton of the pair and check to see if the charge of the second lepton is the same as the first. Assuming that the pair really came from the decay of the Z this gives us a measure of charge misidentification. Fig 10 show the misidentification rates F^+, F^- as functions of pseudorapidity[8]. These rates are very small. The quantity A_l, can be corrected for charge misidentification applying Barlow's method for combining asymmetric errors [19]. The level of correction is 0.3% in the central region and 0.5% in the more forward regions.

1.3.5 Compare events at the generator level to events at the detector level

We have simulated one million signal, $W \to e\nu_e$, events for each of the PDF sets CTEQ6.1, MRST2001 and ZEUS-S using HERWIG (6.505). For each of these PDF sets the eigenvector error PDF sets have been simulated by PDF reweighting and k-factors have been applied to approximate an NLO generation. The top part of Fig. 11 shows the $e^\pm$ and A_l spectra at this generator level, for all of the PDF sets superimposed. The events are then passed through the ATLFAST fast simulation of the ATLAS detector. This applies loose kinematic cuts: $|\eta| < 2.5$, $p_{te} > 5$ GeV, and electron isolation criteria. It also smears the 4-momenta of the leptons to mimic momentum dependent detector resolution. We then apply the selection cuts described in Sec. 1.3.3. The lower half of Fig. 11 shows the $e^\pm$ and A_l spectra at the detector level after application of these cuts, for all of the PDF sets superimposed. The level of precision of each PDF set, seen in the analytic calculations of Fig. 6, is only slightly degraded at detector level, so

[8]These have been corrected for the small possibility that the better reconstructed lepton has had its charge misidentified as follows. In the central region, $|\eta| < 1$, assume the same probability of misidentification of the first and second leptons, in the more forward regions assume the same rate of first lepton misidentification as in the central region.

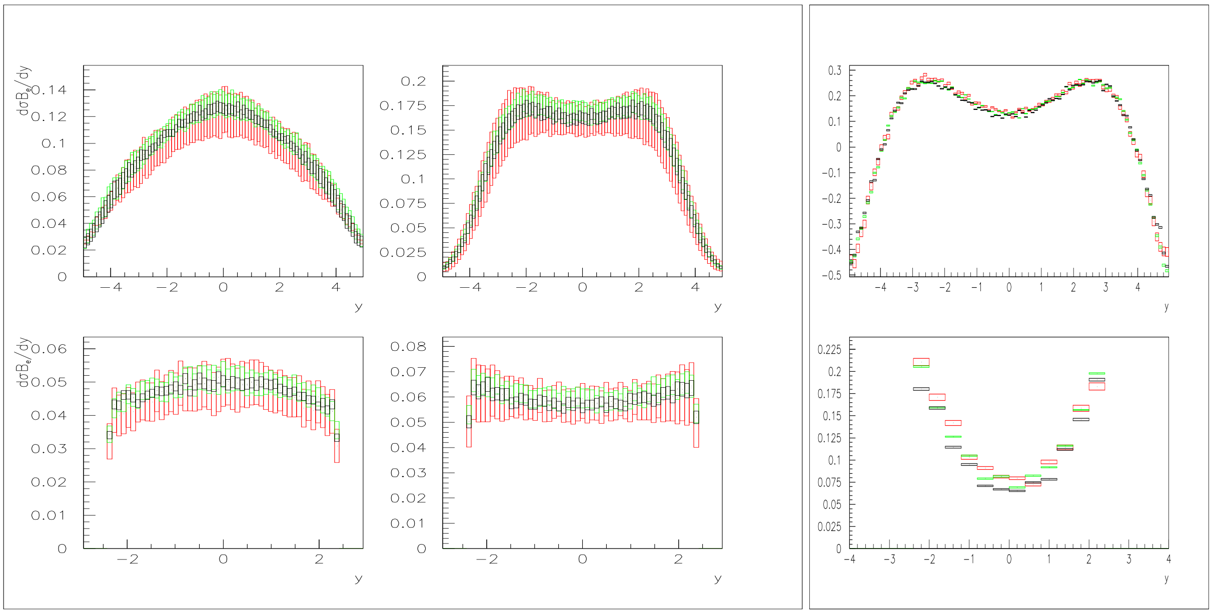


Fig. 11: Top row: e^-, e^+ and A_e rapidity spectra for the lepton from the W decay, generated using HERWIG + k factors and CTEQ6.1 (red), ZEUS-S (green) and MRST2001 (black) PDF sets with full uncertainties. Bottom row: the same spectra after passing through the ATLFAST detector simulation and selection cuts.

that a net level of PDF uncertainty at central rapidity of $\sim 8\%$ is maintained. The anticipated cancellation of PDF uncertainties in the asymmetry spectrum is also observed, within each PDF set, and the spread between PDF sets suggests that measurements which are accurate to better than $\sim 5\%$ could discriminate between PDF sets.

1.3.6 Using LHC data to improve precision on PDFs

The high cross-sections for W production at the LHC ensure that it will be the experimental systematic errors, rather than the statistical errors, which are determining. We have imposed a random 4% scatter on our samples of one million W events, generated using different PDFs, in order to investigate if measurements at this level of precision will improve PDF uncertainties at central rapidity significantly if they are input to a global PDF fit. Fig. 12 shows the e^+ and e^- rapidity spectra for events generated from the ZEUS-S PDFs ($|\eta| < 2.4$) compared to the analytic predictions for these same ZEUS-S PDFs. The lower half of this figure illustrates the result if these events are then included in the ZEUS-S PDF fit. The size of the PDF uncertainties, at $y = 0$, decreases from 5.8% to 4.5%. The largest improvement is in the PDF parameter λ_g controlling the low-x gluon at the input scale, Q_0^2: $xg(x) \sim x^{\lambda_g}$ at low-x, $\lambda_g = -0.199 \pm 0.046$, before the input of the LHC pseudo-data, compared to, $\lambda_g = -0.196 \pm 0.029$, after input. Note that whereas the relative normalisations of the e^+ and e^- spectra are set by the PDFs, the absolute normalisation of the data is free in the fit so that no assumptions are made on our ability to measure luminosity. Secondly, we repeat this procedure for events generated using the CTEQ6.1 PDFs. As shown in Fig. 13, the cross-section for these events is on the lower edge of the uncertainty band of the ZEUS-S predictions. If these events are input to the fit the central value shifts and the uncertainty decreases. The value of the parameter λ_g becomes, $\lambda_g = -0.189 \pm 0.029$, after input of these pseudo-data. Finally to simulate the situation which really faces experimentalists we generate events with CTEQ6.1, and pass them through the ATLFAST detector simulation and cuts. We then correct back from detector level to generator level using a different PDF set- in this case the ZEUS-S PDFs- since in practice we will not know the true PDFs. Fig. 14 shows that the resulting corrected data look pleasingly like CTEQ6.1, but they are more smeared. When these data are input to the PDF fit the central values shift and errors decrease just as for the perfect CTEQ6.1 pseudo-data. The value of λ_g becomes, $\lambda = -0.181 \pm 0.030$, after input of these pseudo-data. Thus we see that the bias introduced by the correction procedure from

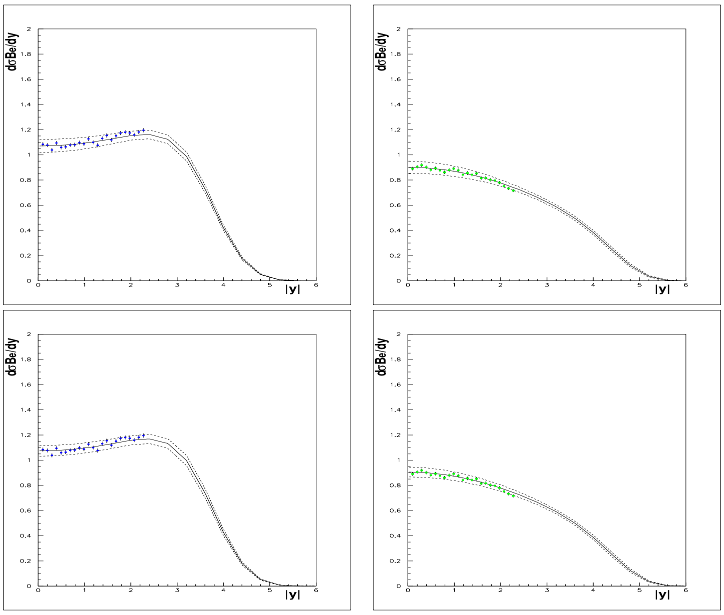


Fig. 12: Top row: e^+ and e^- rapidity spectra generated from ZEUS-S PDFs compared to the analytic prediction using ZEUS-S PDFs. Bottom row: the same lepton rapidity spectra compared to the analytic prediction AFTER including these lepton pseudo-data in the ZEUS-S PDF fit.

detector to generator level is small compared to the PDF uncertainty.

1.3.7 Conclusions and a warning: problems with the theoretical predictions at small-x?

We have investigated the PDF uncertainty on the predictions for W and Z production at the LHC, taking into account realistic expectations for measurement accuracy and the cuts on data which will be needed to identify signal events from background processes. We conclude that at the present level of PDF uncertainty the decay lepton asymmetry, A_l, will be a useful standard model benchmark measurement, and that the decay lepton spectra can be used as a luminosity monitor which will be good to $\sim 8\%$. However, we have also investigated the measurement accuracy necessary for early measurements of these decay lepton spectra to be useful in further constraining the PDFs. A systematic measurement error of less than $\sim 4\%$ would provide useful extra constraints.

However, a caveat is that the current study has been performed using standard PDF sets which are extracted using NLO QCD in the DGLAP [20, 21, 22, 23] formalism. The extension to NNLO is straightforward, giving small corrections $\sim 1\%$. PDF analyses at NNLO including full accounting of the PDF uncertainties are not extensively available yet, so this small correction is not pursued here. However, there may be much larger uncertainties in the theoretical calculations because the kinematic region involves low-x. There may be a need to account for $ln(1/x)$ resummation (first considered in the BFKL [24, 25, 26] formalism) or high gluon density effects. See reference [27] for a review.

The MRST group recently produced a PDF set, MRST03, which does not include any data for $x < 5 \times 10^{-3}$. The motivation behind this was as follows. In a global DGLAP fit to many data sets there is always a certain amount of tension between data sets. This may derive from the use of an inappropriate theoretical formalism for the kinematic range of some of the data. Investigating the effect of kinematic cuts on the data, MRST found that a cut, $x > 5 \times 10^{-3}$, considerably reduced tension between the remaining data sets. An explanation may be the inappropriate use of the DGLAP formalism at small-x. The MRST03 PDF set is thus free of this bias BUT it is also only valid to use it for $x > 5 \times 10^{-3}$. What is needed is an alternative theoretical formalism for smaller x. However, the MRST03 PDF set may be used as a toy PDF set, to illustrate the effect of using very different PDF sets on our predictions. A comparison of Fig. 15 with Fig. 3 or Fig. 6 shows how different the analytic predictions are from the conventional ones, and thus illustrates where we might expect to see differences due to the need for an alternative formalism at small-x.

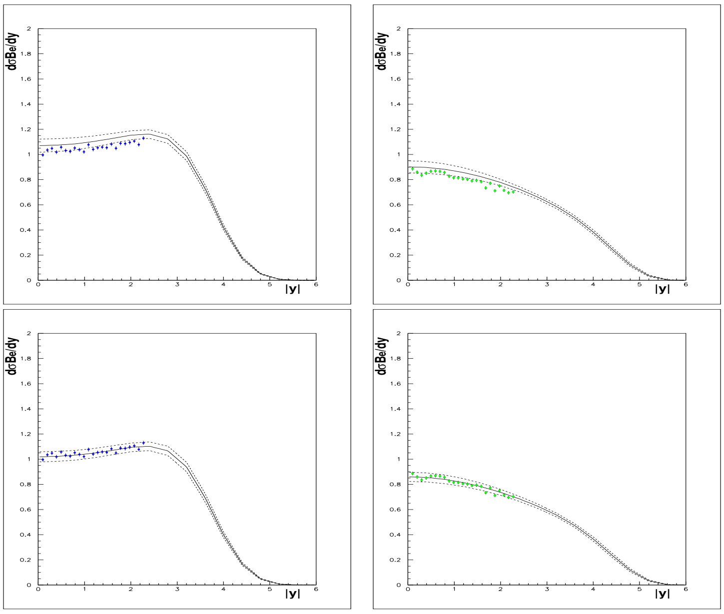


Fig. 13: Top row: e^+ and e^- rapidity spectra generated from CTEQ6.1 PDFs compared to the analytic prediction using ZEUS-S PDFs. Bottom row: the same lepton rapidity spectra compared to the analytic prediction AFTER including these lepton pseudo-data in the ZEUS-S PDF fit.

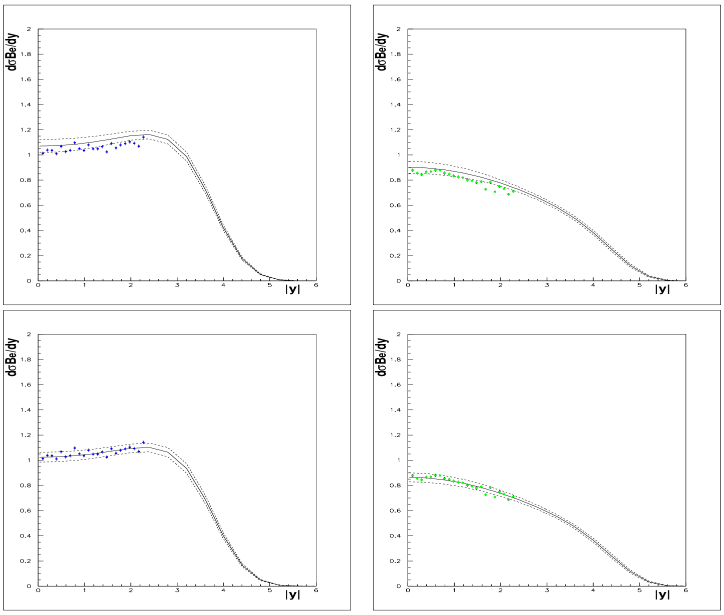


Fig. 14: Top row: e^+ and e^- rapidity spectra generated from CTEQ6.1 PDFs, which have been passed through the ATLFAST detector simulation and corrected back to generator level using ZEUS-S PDFs, compared to the analytic prediction using ZEUS-S PDFs. Bottom row: the same lepton rapidity spectra compared to the analytic prediction AFTER including these lepton pseudo-data in the ZEUS-S PDF fit.

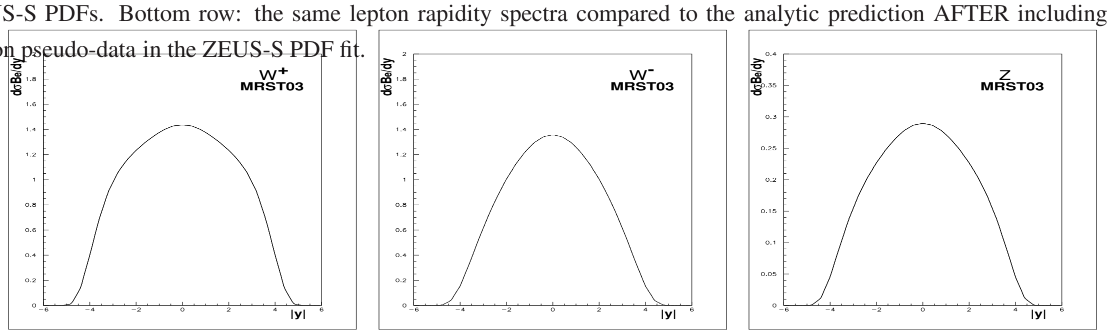


Fig. 15: LHC W^+, W^-, Z rapidity distributions for the MRST03 PDFs: left plot W^+; middle plot W^-; right plot Z

1.4 W and Z production at the LHC[9]

The study of the production at the LHC of the electroweak bosons W and Z with subsequent decays in leptonic final states will provide several precision measurements of Standard Model parameters such as the mass of the W boson or the weak mixing angle from the Z boson forward-backward asymmetry. Given their large cross section and clean experimental signatures, the bosons will furthermore serve as calibration tool and luminosity monitor. More challenging, differential cross sections in rapidity or transverse momentum may be used to further constrain parton distribution functions. Eventually these measurements for single inclusive boson production may be applied to boson pair production in order to derive precision predictions for background estimates to discovery channels like $H \to W^+W^-$.

This contribution is devoted to the estimation of current uncertainties in the calculations for Standard Model cross sections involving W and Z bosons with particular emphasis on the PDF and perturbative uncertainties. All results are obtained at NLO with MCFM [28] version 4.0 interfaced to LHAPDF [12] for a convenient selection of various PDF families and evaluation of their intrinsic uncertainties. The cross sections are evaluated within a typical experimental acceptance and for momentum cuts summarised in Table 3. The electromagnetic decays of W and Z are considered (massless leptons) and the missing transverse energy is assigned to the neutrino momentum sum (in case of W decays). Jets in the processes $W/Z + jets$ are produced in an inclusive mode with at least one jet in the event

Table 3: Experimental acceptance cuts used for the calculation of cross-sections.

Observable	cut
p_T^{lept}	> 25 GeV
p_T^{jet}	> 25 GeV
$\lvert\eta_{lept}\rvert$	< 3.0
$\lvert\eta_{jet}\rvert$	< 4.0
$R(lepton - jet)$	> 0.8
$R(lepton - lepton)$	> 0.2
E_T^{miss}	>25 GeV

reconstructed with the k_T-algorithm. MCFM includes one- and two-jet processes at NLO and three-jet processes at LO. In the case of boson pair production the cuts of Table 3 can only be applied to the two leading leptons, hence a complete acceptance is assumed for additional leptons e.g. from ZZ or WZ decays.

The calculations with MCFM are carried out for a given fixed set of electroweak input parameters using the effective field theory approach [28]. The PDF family CTEQ61 provided by the CTEQ collaboration [29] is taken as nominal PDF input while MRST2001E given by the MRST group [30] is considered for systematic purposes. The difference between CTEQ61 and MRST2001E alone can't be considered as systematic uncertainty but merely as cross-check. The systematic uncertainty is therefore estimated for each family separately with the family members, 40 for CTEQ61 and 30 for MRST2001E, which are variants of the nominal PDF obtained with different assumptions while maintaining a reasonable fit of the input data. The value of α_s is not a free input parameter for the cross section calculation but taken from the corresponding value in the PDF.

Important input parameters are renormalisation and factorisation scales. The central results are obtained with $\mu_R = \mu_F = M_V$, $V = W, Z$ for single boson production and $\mu_R = \mu_F = M_V + M_V'$ for pair production (V' being the second boson in the event). Missing higher orders are estimated by

[9]Contributing author: Hasko Stenzel

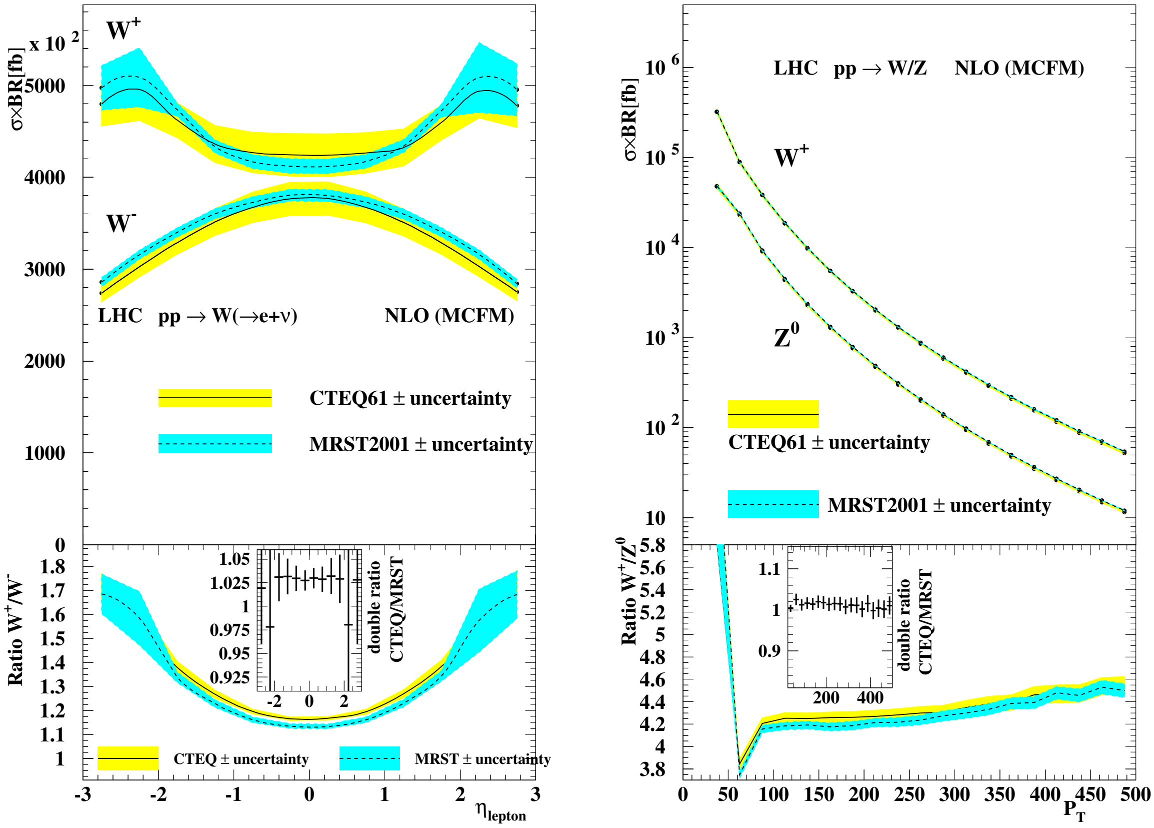


Fig. 16: Left: pseudo-rapidity distribution of the decay lepton from inclusive W production and right: p_T spectra of W and Z. The bands represent the PDF-uncertainty. The lower inserts show on the left side the ratio W^+/W^- resp. the double-ratio CTEQ/MRST and on the right side the ratios for W^+/Z^0.

a variation of the scales in the range $1/2 \leq x_{\mu R} \leq 2$ and independently $1/2 \leq x_{\mu F} \leq 2$ where $\mu = x_\mu \cdot M_V$, following prescriptions applied to other processes [31], keeping in mind that the range of variation of the scales is purely conventional.

1.4.1 Single W and Z cross sections

Detailed studies of single W and Z production including detector simulation are presented elsewhere in these proceedings, here these channels are mainly studied for comparison with the associated production with explicitly reconstructed jets and with pair production. The selected process is inclusive in the sense that additional jets, present in the NLO calculation, are not explicitly reconstructed. The experimentally required lepton isolation entailing a jet veto in a restricted region of phase space is disregarded at this stage.

As an example the pseudo-rapidity distribution of the lepton from W decays and the p_T spectra for Z and W^+ are shown in fig. 16. The cross section for W^+ is larger than for W^- as a direct consequence of the difference between up- and down-quark PDFs, and this difference survives in the pseudo-rapidity distribution of the decay lepton with a maximum around $|\eta|$=2.5. In the central part the PDF uncertainty, represented by the bands in fig. 16, amounts to about 5% for CTEQ and 2% for MRST, and within the uncertainty CTEQ and MRST are fully consistent. Larger differences are visible in the peaks for the W^+, where at the same time the PDF uncertainty increases. In the ratio W^+/W^- the PDF uncertainty is reduced to about 1-2% in the central region and a difference of about 3% is observed between CTEQ and MRST, as can be seen from the double-ratio CTEQ/MRST. The uncertainty of the double ratio is calculated from the CTEQ uncertainty band alone.

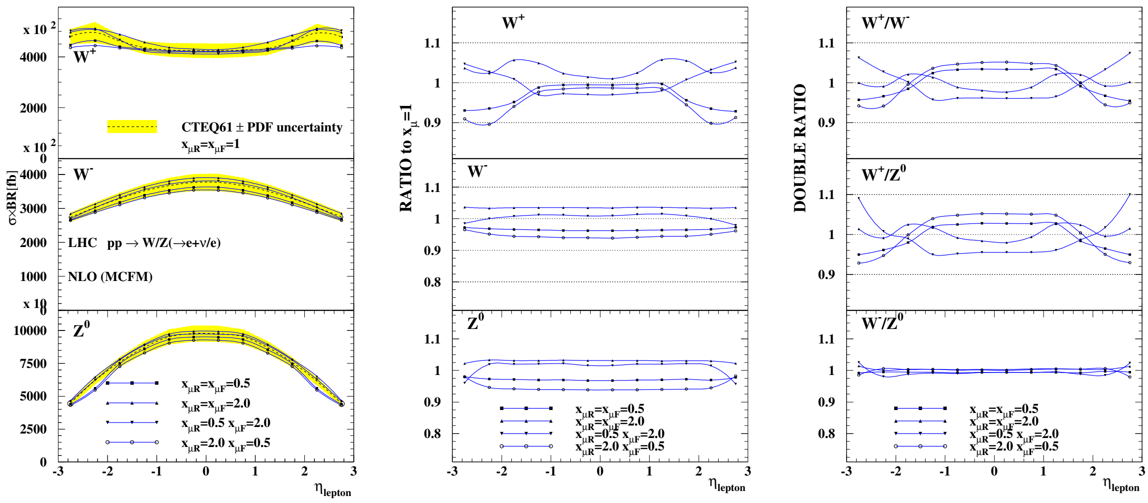


Fig. 17: Left: pseudo-rapidity distribution of the decay lepton from inclusive W/Z production for different values of $x_{\mu R}$ and $x_{\mu F} = 1$, centre: the ratio of predictions with respect to $x_{\mu} = 1$ and right: double ratio V/V' of cross sections for actual scale settings normalised to the nominal scale.

In the case of Z production the rapidity and p_T spectra can be fully reconstructed from the e^+e^- pair. A measurement of the Z p_T spectrum may be used to tune the Monte Carlo description of W p_T, which is relevant for measurements of the W mass. The p_T spectra are shown in the right part of fig. 16. The total yield for W^+ is about six times larger than for Z^0 but for $p_T > 150$ GeV the ratio stabilises around 4.5. At small values of p_T the fixed-order calculation becomes trustless and should be supplemented by resummed calculations. The PDF uncertainties for the p_T spectra themselves are again about 5% and about 2% in the ratio, CTEQ and MRST being consistent over the entire p_T range.

The perturbative uncertainties are estimated by variations of the renormalisation and factorisation scales in by a factor of two. The scale variation entails a global change in the total cross section of the order of 5%. The η distribution of leptons from W/Z decays are shown in fig. 17, comparing the nominal cross section with $x_{\mu R} = x_{\mu F} = 1$, to alternative scale settings. The nominal cross section is drawn with its PDF uncertainty band, illustrating that the perturbative uncertainties are of the same size. For W^- and Z^0 the shape of the distribution is essentially unaltered, but for W^+ the region around the maxima is changed more than the central part, leading to a shape deformation. The scale variation uncertainty is strongly correlated for W^- and Z^0 and cancels in the ratio W^-/Z^0, but for W^+ it is almost anti-correlated with W^- and Z^0 and partly enhanced in the ratio.

Globally the perturbative uncertainty is dominated by the asymmetric scale setting $x_{\mu R} = 2, x_{\mu R} = 1/2$ for which a change of -5% is observed, the largest upward shift of 3.5% is obtained for $x_{\mu R} = 2, x_{\mu R} = 2$, locally the uncertainty for W^+ can be much different. It can be expected that the perturbative uncertainties are reduced for NNLO calculations to the level of 1%.

The integrated cross sections and systematic uncertainties within the experimental acceptance are summarised in Table 4.

1.4.2 $W/Z + jet$ production

In the inclusive production of $W/Z + jet$ at least one jet is requested to be reconstructed, isolated from any lepton by $R > 0.8$. Additional jets are in case of overlap eventually merged at reconstruction level by the k_T-prescription. Given the presence of a relatively hard ($p_T > 25$ GeV) jet, it can be expected that PDF- and perturbative uncertainties are different than for single boson production. The study of this

Table 4: Total cross-sections and systematic uncertainties within the experimental acceptance.

	W^+	W^-	Z^0
CTEQ61 [pb]	5438	4002	923.9
Δ_{PDF}^{CTEQ} [pb]	± 282	± 221	± 49.1
Δ_{PDF}^{CTEQ} [%]	± 5.2	± 5.5	± 5.3
MRST [pb]	5480	4110	951.1
Δ_{PDF}^{MRST} [pb]	± 103	± 83.4	± 17.4
Δ_{PDF}^{MRST} [%]	± 1.9	± 2.1	± 1.9
Δ_{pert} [%]	$+3.5$ -5.2	$+3.5$ -5.4	$+3.1$ -5.5

process at the LHC, other than being a stringent test of perturbative QCD, may in addition contribute to a better understanding of the gluon PDF.

The first difference with respect to single boson production appears in the lepton pseudo-rapidities, shown in fig. 18. The peaks in the lepton spectrum from W^+ disappeared, the corresponding spectrum from W^- is stronger peaked at central rapidity while the ratio W^+/W^- with jets is essentially the same as without jets. The PDF uncertainties are slightly smaller (4.2-4.4%) compared to single bosons. The jet pseudo-rapidities are shown in the right part of fig. 18, they are much stronger peaked in the central region but the ratio W^+/W^- for jets is similar to the lepton ratio.

The transverse momenta of associated jets from $W/Z + jet$ production is shown in fig. 19, the spectra are steeply falling and the ratio W^+/W^- is increasing from 1.3 at low p_T to almost 2 at 500 GeV p_T.

The perturbative uncertainties are investigated in the same way as for the single boson production and are shown in fig. 20. The scale variation entails here a much larger uncertainty between 8 and 10%, almost twice as large as for single bosons. In contrast to the latter case, the scale variation is correlated for W and Z and cancels in the ratio W^+/W^-, with an exception for W^- where a bump appears at $|\eta| = 1.8$ for $x_{\mu R} = 2$.

The total cross sections and their systematic uncertainties are summarised in Table 5.

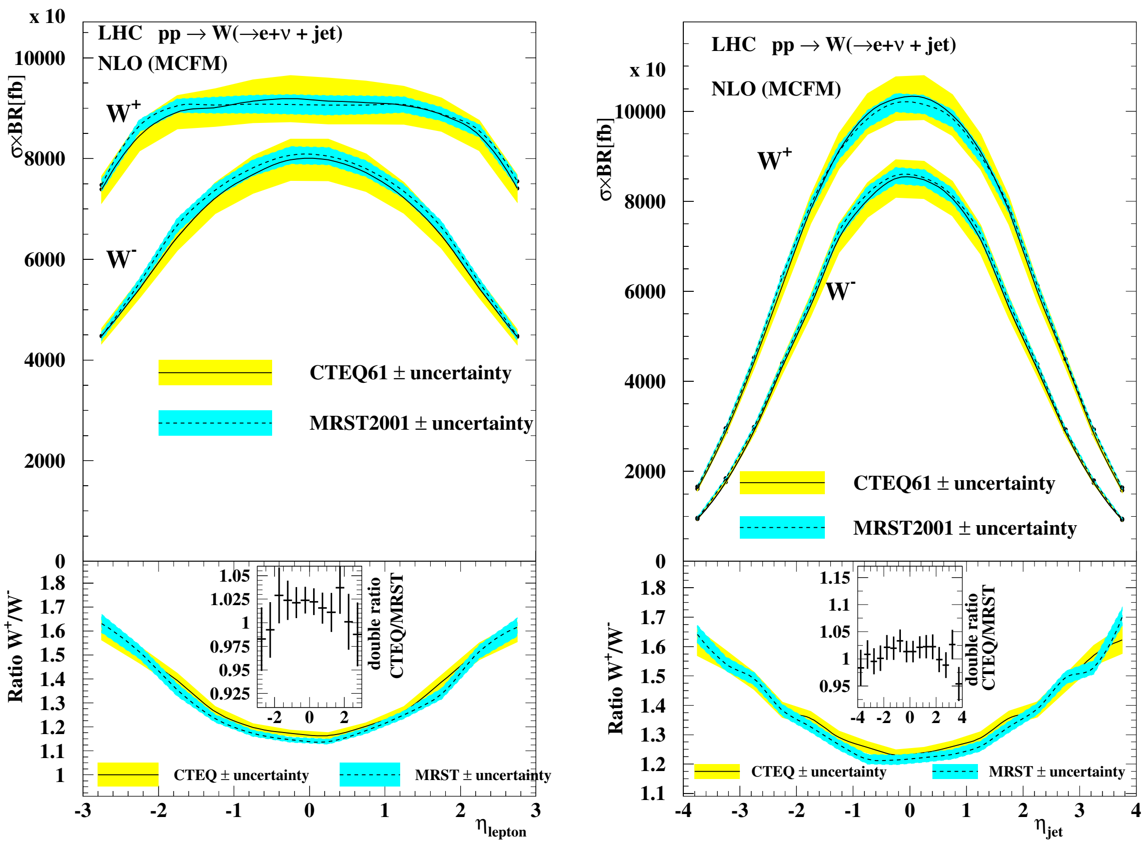


Fig. 18: Left: pseudo-rapidity distribution of the decay lepton from inclusive W+jet production and right: pseudo-rapidity of the associated leading jet. The bands represent the PDF-uncertainty.

Table 5: Total cross-sections and systematic uncertainties within the experimental acceptance for $W/Z + jet$ processes.

	$W^+ + jet$	$W^- + jet$	$Z^0 + jet$
CTEQ61 [pb]	1041	784.5	208.1
Δ_{PDF}^{CTEQ} [pb]	± 44.1	± 34.3	± 9.01
Δ_{PDF}^{CTEQ} [%]	± 4.2	± 4.4	± 4.3
MRST [pb]	1046	797.7	211.3
Δ_{PDF}^{MRST} [pb]	± 17.6	± 14.8	± 3.67
Δ_{PDF}^{MRST} [%]	± 1.7	± 1.9	± 1.8
Δ_{pert} [%]	$+8.7$ -9.8	$+8.9$ -10.0	$+7.6$ -9.1

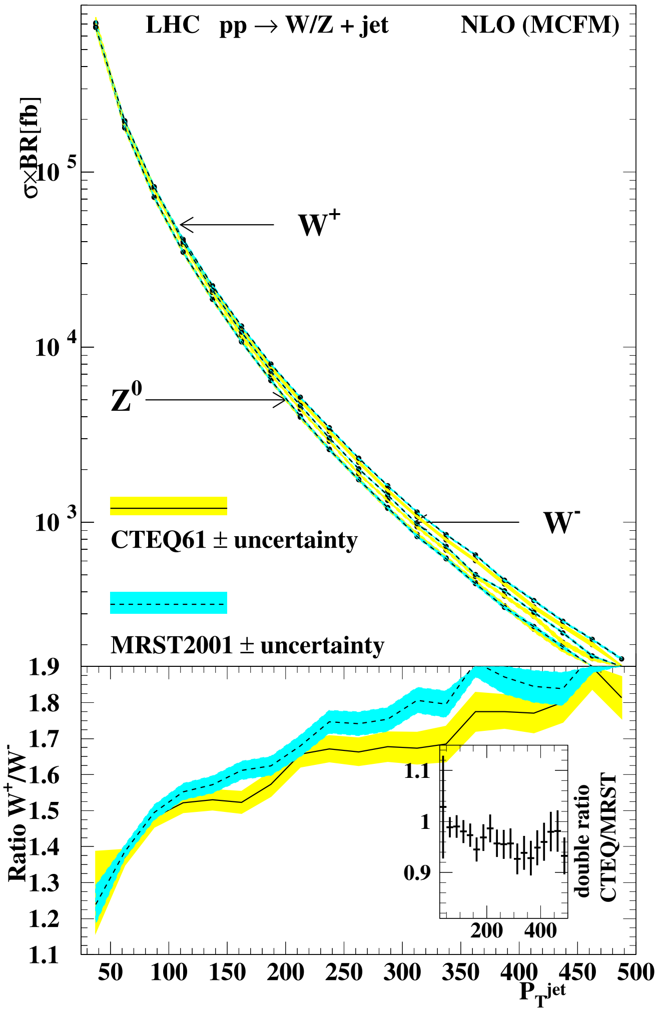


Fig. 19: Transverse momentum distribution of the jet from inclusive $W/Z + jet$ production

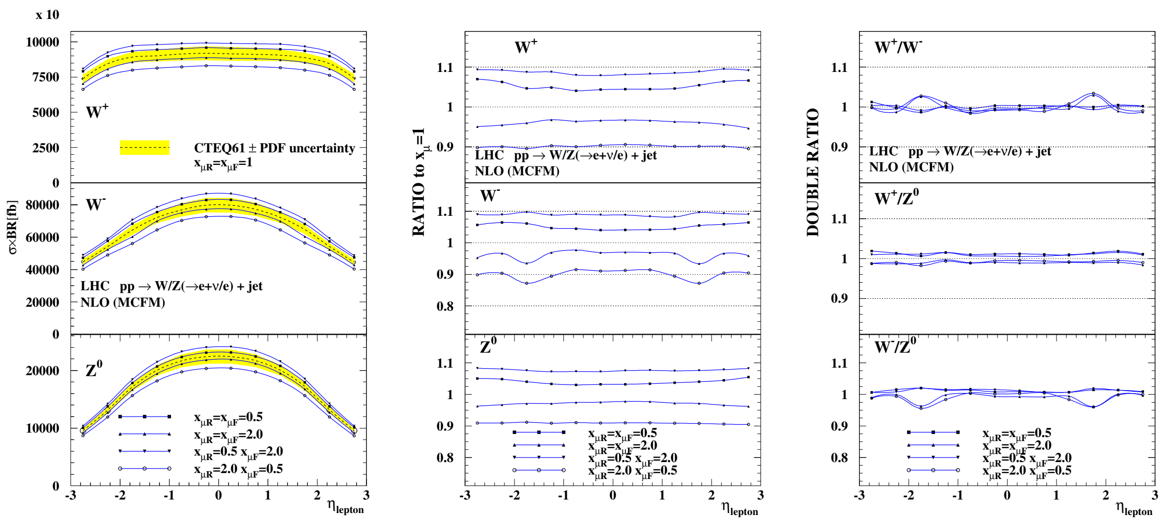


Fig. 20: Left: pseudo-rapidity distribution of the decay lepton from inclusive $W/Z + jet$ production for different values of $x_{\mu R}$ and $x_{\mu F} = 1$, centre: the ratio of predictions with respect to $x_{\mu} = 1$ and right: double ratio V/V' of cross sections for actual scale settings normalised to the nominal scale.

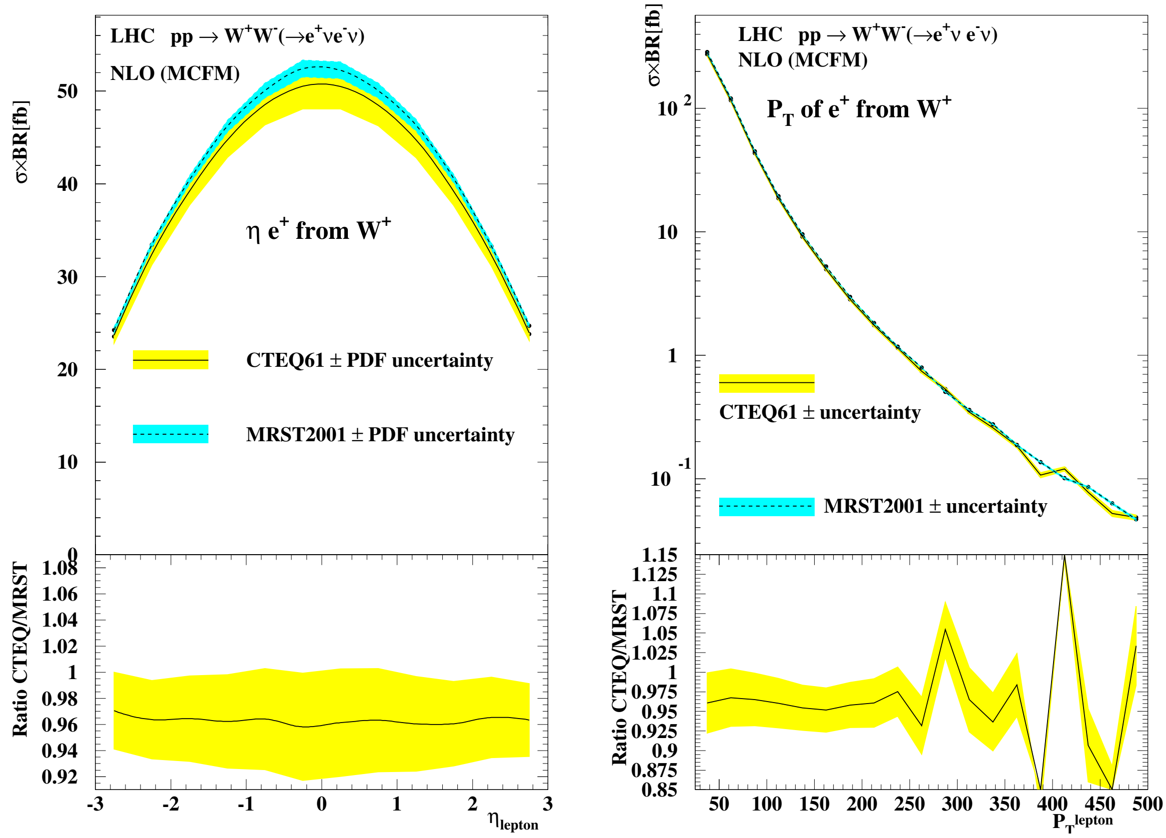


Fig. 21: Left: pseudo-rapidity distribution of the decay lepton from inclusive WW production and right: transverse momentum of the decay lepton.

1.4.3 Vector Boson pair production

In the Standard Model the non-resonant production of vector bosons pairs in the continuum is suppressed by factors of 10^4-10^5 with respect to single Boson production. The cross sections for WW, WZ and ZZ within the experimental acceptance range from 500 fb (WW) to 10 fb (ZZ). Given the expected limited statistics for these processes, the main goal of their experimental study is to obtain the best estimate of the background they represent for searches of the Higgs boson or new physics yielding boson pairs.

The selection of boson pairs follows in extension the single boson selection cuts applied to 2, 3 or 4 isolated leptons. Again real gluon radiation and virtual loops have been taken into account at NLO but without applying lepton-jet isolation cuts. Lepton-lepton separation is considered only for the two leading leptons.

The pseudo-rapidity and transverse momentum distributions taking the e^+ from W^+W^- production as example are shown in fig.21. The pseudo-rapidity is strongly peaked and the cross section at $\eta = 0$ twice as large as at $|\eta| = 3$. The PDF uncertainties are smaller than for single bosons, between 3.5 and 4 $\%$.

The same shape of lepton distributions is also found for the other lepton and for the other pair production processes, as shown for the W^-Z^0 case in fig.22.

The rapidity distribution of the leading Z^0 from ZZ production is shown in the left part of fig.23. With both Z's being fully reconstructed, the invariant mass of the ZZ system can be compared in the right part of fig.23 to the invariant mass spectrum of the Higgs decaying into the same final state for an intermediate mass of $m_H = 200$ GeV. In this case a clear peak appears at low invariant masses above the continuum, and the mass spectrum is also harder at high masses in presence of the Higgs.

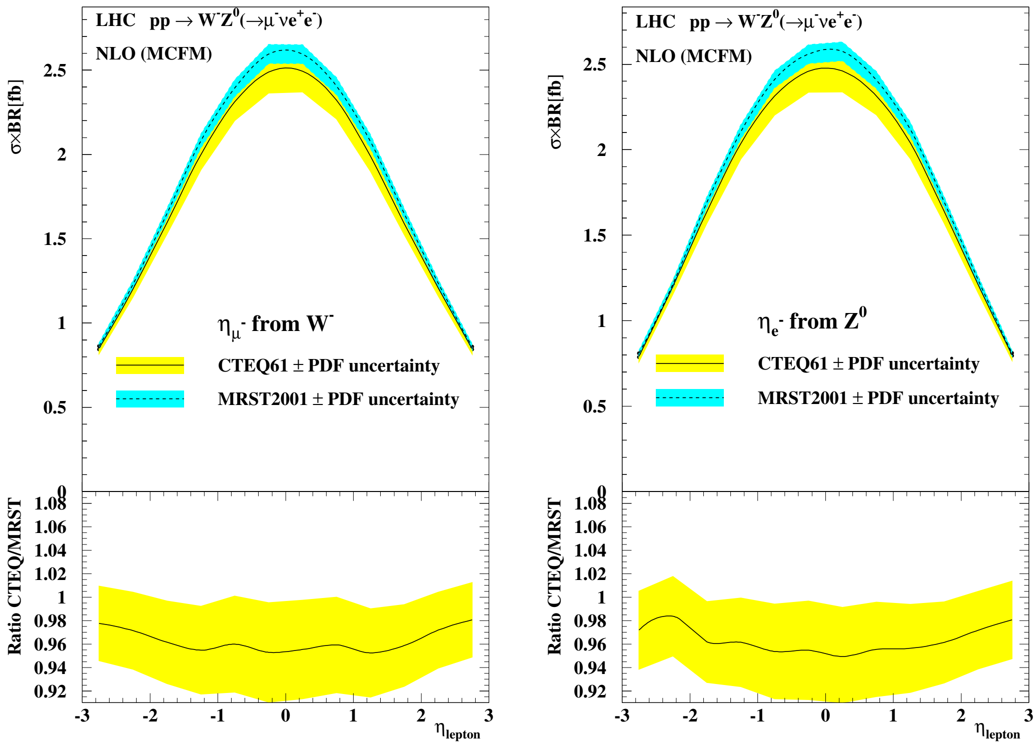


Fig. 22: Left: pseudo-rapidity distribution of the decay lepton of the W^- from inclusive W^-Z^0 production and right: pseudo-rapidity distribution of a decay lepton of the Z^0.

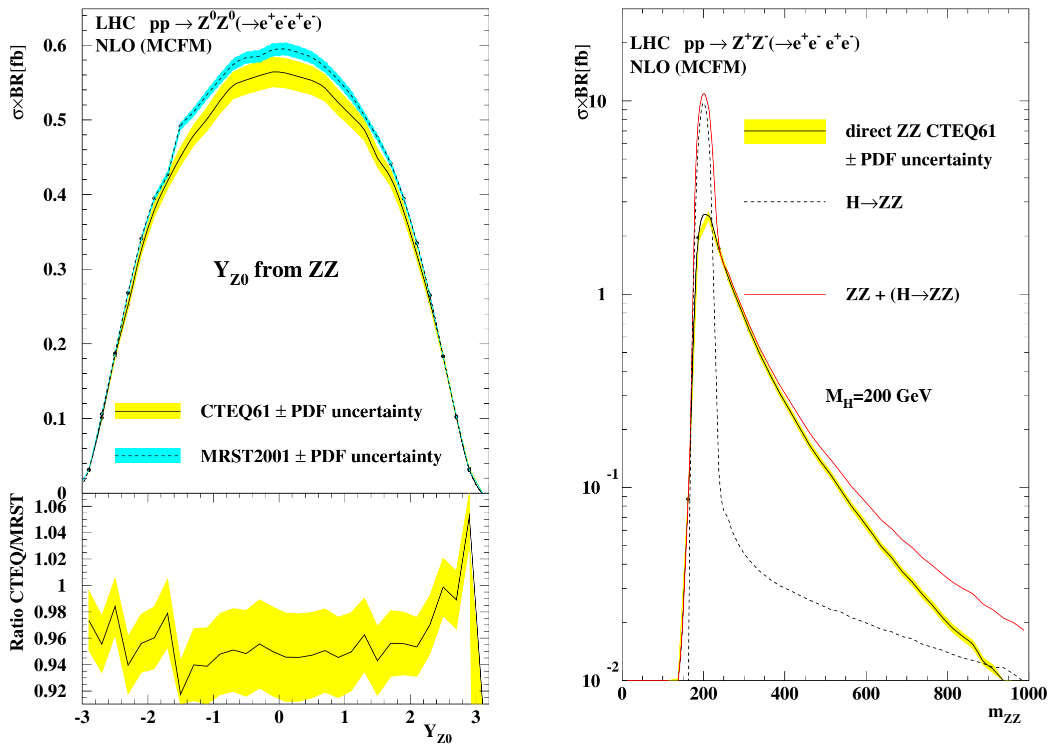


Fig. 23: Left: rapidity distribution of the leading Z from inclusive ZZ production and right: invariant mass of the ZZ pair for non-resonant continuum production compared to resonant pair production via the SM Higgs decay.

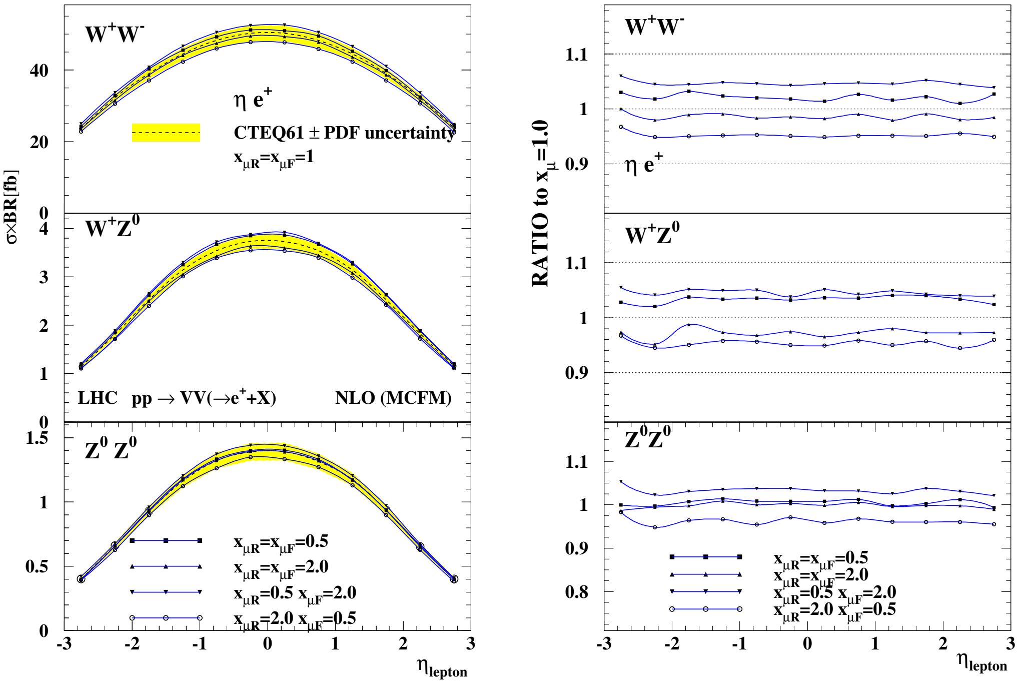


Fig. 24: Left: pseudo-rapidity distributions of leptons from various boson pair production processes and different scale settings and right: ratio of predictions relative to $x_\mu = 1$.

The perturbative uncertainties, obtained as for the other processes, are shown in fig.24 for the lepton distributions. The systematic uncertainties range from 3.3 to 4.9 % and are slightly smaller than for single bosons, given the larger scale $\mu = 2M_V$ and better applicability of perturbative QCD. The perturbative uncertainty is essentially constant across the pseudo-rapidity and largely correlated between different pair production processes.

The ratio of boson pair production to single Z production is of particular interest, as similar quark configurations contribute to both process types, though evidently in a somewhat different x, Q^2 regime. This ratio is shown in fig.25 for the lepton distribution, given the different shapes of pseudo-rapidity is not flat but its PDF uncertainty is reduced to the level of 2 %. The perturbative uncertainties of the VV/Z ratio, however, are only reduced for the ZZ/Z case and even slightly larger for other ratios because the scale variations have partly an opposite effect on the cross sections for Z and e.g. WW production.

The total cross sections and their systematic uncertainties are summarised in Table 6.

1.5 Study of next-to-next-to-leading order QCD predictions for W and Z production at LHC[10]

It has been in 2004 that the first differential next-to-next-to-leading order (NNLO) QCD calculation for vector boson production in hadron collisions was completed by Anastasiou *et al.* [32]. This group has calculated the rapidity dependence for W and Z production at NNLO. They have shown that the perturbative expansion stabilizes at this order in perturbation theory and that the renormalization and factorization scale uncertainties are drastically reduced, down to the level of one per-cent. It is therefore interesting to perform a more detailed study of these NNLO predictions for various observables which can be measured at LHC, as well as to investigate their systematic uncertainties.

In the study presented here we have calculated both the differential (in rapidity) and inclusive

[10]Contributing author:Günther Dissertori

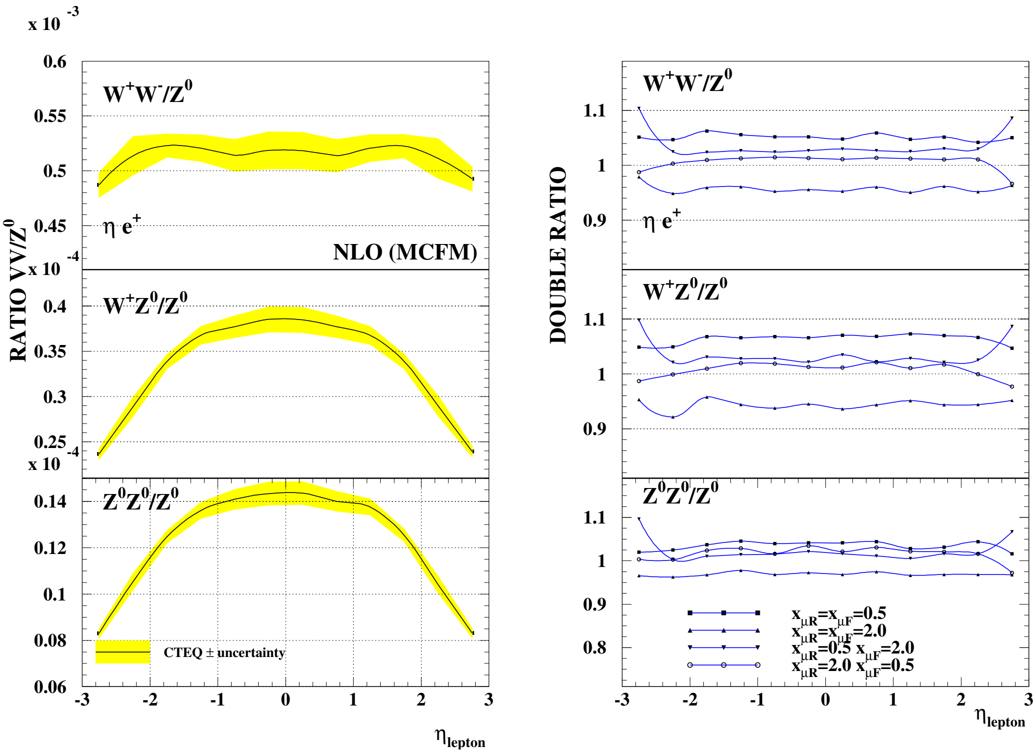


Fig. 25: Left: the ratio of pseudo-rapidity distributions of leptons from boson pair production processes normalised to single Z production and right: the double ratio VV/Z of predictions for different scales relative to $x_\mu = 1$.

Table 6: Total cross-sections and systematic uncertainties within the experimental acceptance for pair production processes.

	WW	ZZ	W^+Z^0	W^-Z^0
CTEQ61 [fb]	475.7	11.75	31.81	20.77
Δ_{PDF}^{CTEQ} [fb]	± 17.0	± 0.48	± 1.12	± 0.80
Δ_{PDF}^{CTEQ} [%]	± 3.6	± 4.1	± 3.5	± 3.8
MRST [fb]	494.2	12.34	32.55	21.62
Δ_{PDF}^{MRST} [fb]	± 6.3	± 0.19	± 0.49	± 0.41
Δ_{PDF}^{MRST} [%]	± 1.3	± 1.6	± 1.5	± 1.9
Δ_{pert} [%]	$+4.6$ -4.9	$+3.3$ -3.8	$+4.6$ -4.7	$+4.8$ -4.7

cross sections for W, Z and high-mass Drell-Yan (Z/γ^*) production. Here "inclusive" refers to the results obtained by integrating the differential cross sections over a rapidity range similar to the experimentally accessible region, which might be more relevant than the complete cross section which also includes the large-rapidity tails.

Such a prediction would then be compared to the experimental measurements at LHC, which will allow for precise tests of the Standard Model as well as to put strong constraints on the parton distribution functions (PDFs) of the proton. It is clear that in the experiment only the rapidity and transverse momenta of the leptons from the vector boson decays will be accessible, over a finite range in phase space. In order to compute the rapidity of the vector boson by taking into account the finite experimental lepton acceptance, Monte Carlo simulations have to be employed which model vector boson production at the best possible precision in QCD, as for example the program MC@NLO [17]. The so computed acceptance corrections will include further systematic uncertainties, which are not discussed here.

1.5.1 Parameters and analysis method

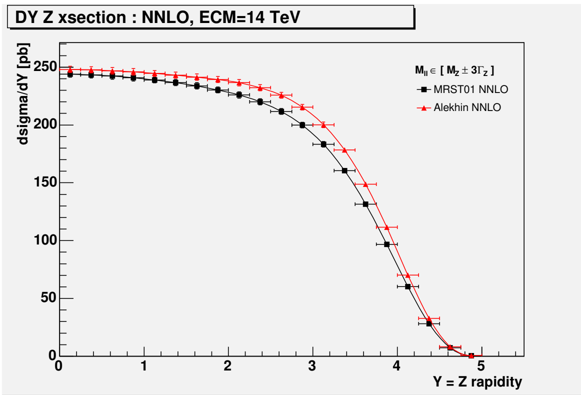
The NNLO predictions have been implemented in the computer code VRAP [33], which has been modified in order to include ROOT [34] support for producing ntuples, histograms and plots. The code allows to specify the collision energy (14 TeV in our case), the exchanged vector boson (γ^*, Z, Z/γ^*, W^+, W^-), the scale Q of the exchanged boson ($M_\mathrm{Z}, M_\mathrm{W}$ or off-shell, e.g. $Q = 400\,\mathrm{GeV}$), the renormalization and factorization scales, the invariant mass of the di-lepton system (fixed or integrated over a specified range), the value of the electro-magnetic coupling ($\alpha_\mathrm{QED} = 1/128$ or $\alpha_\mathrm{QED}(Q)$) and the number of light fermions considered. Regarding the choice of pdfs, the user can select a pdf set from the MRST2001 fits [35] or from the ALEKHIN fits [36], consistent at NNLO with variable flavour scheme. We have chosen the MRST2001 NNLO fit, mode 1 with $\alpha_s(M_\mathrm{Z}) = 0.1155$ [35], as reference set.

The program is run to compute the differential cross section $d\sigma/dY$, Y being the boson rapidity, at a fixed number of points in Y. This result is then parametrized using a spline interpolation, and the thus found function can be integrated over any desired rapidity range, such as $|Y| < 2, |Y| < 2.5$ or $|Y| < 3$, as well as over finite bins in rapidity. For the study of on-shell production the integration range over the di-lepton invariant mass M_{ll} was set to $M_V - 3\Gamma_V < M_{ll} < M_V + 3\Gamma_V$, with M_V and Γ_V the vector boson mass and width. This simulates an experimental selection over a finite signal range.

The systematic uncertainties have been divided into several categories: The PDF uncertainty is estimated by taking the maximum deviation from the reference set when using different PDFs from within the MRST2001 set or the ALEKHIN set. The latter difference is found to give the maximal variation in all of the investigated cases. The renormalization and factorization scales $\mu = \mu_\mathrm{R}, \mu_\mathrm{F}$ have been varied between $0.5 < \mu/Q < 2$, both simultaneously as well as fixing one to $\mu = Q$ and varying the other. The maximum deviation from the reference setting $\mu = Q$ is taken as uncertainty. The observed difference when using either a fixed or a running electro-magnetic coupling constant is also studied as possible systematic uncertainty due to higher-order QED effects. Since it is below the one per-cent level, it is not discussed further. Finally, in the case of Z production it has been checked that neglecting photon exchange and interference contributions is justified in view of the much larger PDF and scale uncertainties.

1.5.2 Results for W and Z production

In Figure 26 the results for Z production at LHC are shown for two different choices of PDF set, as a function of the boson rapidity. It can be seen that the predictions differ by about 2% at central rapidity, and the difference increases to about 5% at large rapidity. A similar picture is obtained when integrating the differential cross section up to rapidities of 2, 2.5 and 3 (Table 7). The more of the high-rapidity tail is included, the larger the uncertainty due to the PDF choice. From Table 1 it can also be seen that the scale uncertainties are slightly below the one per-cent level. It is worth noting that the choice of the



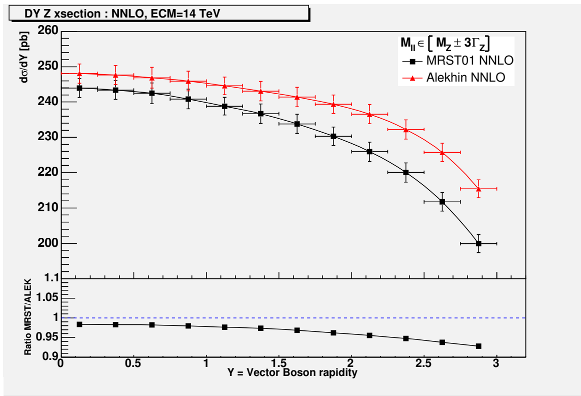


Fig. 26: Left : Drell-Yan Z production cross section ($\times$ BR) at LHC energies, as a function of the Z rapidity, for two different PDF choices. Right : Zoom into a restricted rapidity region, with the ratio of the predictions for the two different PDF sets as lower inset. The error bars indicate the scale uncertainties.

integration range over the di-lepton invariant mass can have a sizeable impact on the cross section. For example, increasing the range from the standard value to $66\,\mathrm{GeV} < M_{\mathrm{Z}} < 116\,\mathrm{GeV}$ increases the cross section by 8%.

Channel	Z prod.			W prod.		
range	$\|Y\| < 2$	$\|Y\| < 2.5$	$\|Y\| < 3$	$\|Y\| < 2$	$\|Y\| < 2.5$	$\|Y\| < 3$
cross section [nb]	0.955	1.178	1.384	9.388	11.648	13.800
Δ PDF [%]	2.44	2.95	3.57	5.13	5.47	5.90
Δ scale [%]	0.85	0.87	0.90	0.99	1.02	1.05

Table 7: NNLO QCD results for W and Z production at the LHC for the integration over different rapidity ranges. Also given are the relative uncertainties due to the choice of the PDFs and of the renormalization and factorization scale. The numbers include the branching ratio $Z(W) \rightarrow ee(e\nu)$.

The results for W production (Table 7) have been obtained by first calculating separately the cross sections for W^+ and W^- production, and then adding these up. Again we observe an increase of the PDF uncertainty when going to larger rapidity ranges. Compared to the Z production, here the PDF uncertainties are larger, between 5 and 6%, whereas the scale uncertainties are of the same level, $\approx 1\%$. It is interesting to note that the PDF uncertainty for W^- production is about 10 - 20% (relative) lower than that for W^+.

A considerable reduction in systematic uncertainty can be obtained by calculating cross section ratios. Two options have been investigated, namely the ratios $\sigma(\mathrm{W}^+)/\sigma(\mathrm{W}^-)$ and $\sigma(\mathrm{W})/\sigma(\mathrm{Z})$. As can be seen from Figure 27, the PDF uncertainties are reduced to the 0.7% level in the former ratio, and to about 2% in the latter. The scale uncertainties are reduced to the 0.15% level in both cases. Taking such ratios has also the potential advantage of reduced experimental systematic uncertainties, such as those related to the acceptance corrections.

1.5.3 Results for high-mass Drell-Yan processes

Similarly to on-shell W and Z production we have also analyzed the high-mass Drell-Yan process, namely Z/γ^* production at a scale of $Q = 400$ GeV. In this case the di-lepton invariant mass has been integrated over the range $M_{ll} = 400 \pm 50$ GeV. Here the PDF uncertainties are found between 3.7% and 5.1% for the various integration ranges over rapidity, somewhat larger than for on-shell pro-

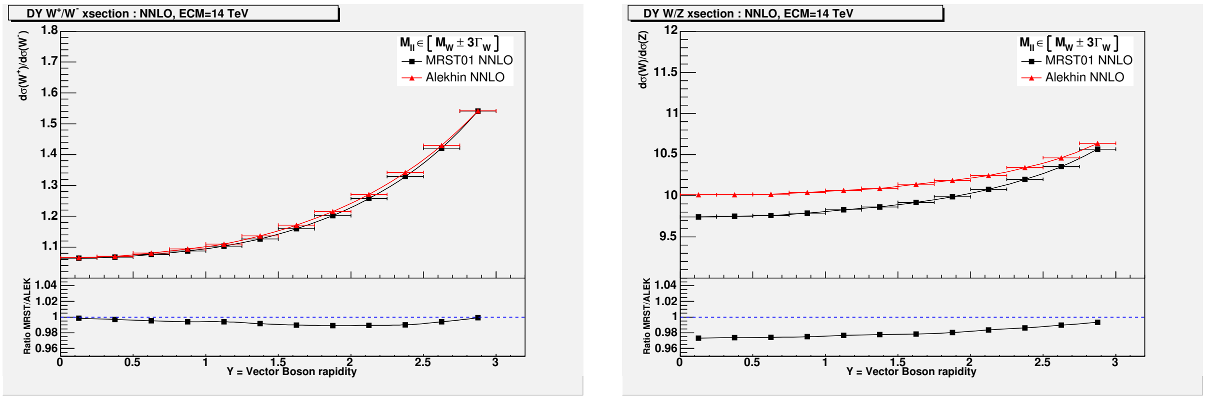


Fig. 27: Ratio of the production cross sections for W^+, W^- (left), and W, Z (right), as a function of rapidity, for two different PDF sets. The inserts show the ratios of the results for the two PDF choices.

duction. However, by normalizing the high-mass production cross section to the on-shell case, the PDF uncertainties are considerably reduced, being 1.2 - 1.5%.

The systematic uncertainties related to the renormalization and factorization scale are reduced (Δ scale $\approx 0.2\%$) when going to the high-mass exchange, as expected from perturbative QCD with a decreasing strong coupling constant. In this case a normalization of the cross section to the on-shell case does not give an improvement. However, since the scale uncertainties are well below the PDF uncertainties, this is less of an issue for the moment.

1.5.4 Summary

We have studied NNLO QCD predictions for W and Z production at LHC energies. We have identified the choice of PDF set as the dominant systematic uncertainty, being between 3 and 6%. The choice of the renormalization and factorization scale leads to much smaller uncertainties, at or below the 1% level. In particular we have shown that the systematic uncertainties can be sizeably reduced by taking ratios of cross sections, such as $\sigma(\mathrm{W}^+)/\sigma(\mathrm{W}^-)$, $\sigma(\mathrm{W})/\sigma(\mathrm{Z})$ or $\sigma(\mathrm{Z}/\gamma^*, Q = 400\,\mathrm{GeV})/\sigma(\mathrm{Z}/\gamma^*, Q = M_{\mathrm{Z}})$. For such ratios it can be expected that also part of the experimental uncertainties cancel. With theoretical uncertainties from QCD at the few per-cent level the production of W and Z bosons will most likely be the best-known cross section at LHC.

Concerning the next steps, it should be considered that at this level of precision it might become relevant to include also higher-order electro-weak corrections. In addition, since experimentally the boson rapidity will be reconstructed from the measured lepton momenta, a detailed study is needed to evaluate the precision at which the acceptance correction factors for the leptons from the boson decays can be obtained. For this Monte Carlo programs such as MC@NLO should be employed, which combine next-to-leading-order matrix elements with parton showers and correctly take account of spin correlations.

References

[1] Dittmar, M. and Pauss, F. and Zurcher, D., Phys. Rev. **D56**, 7284 (1997).

[2] Acosta, D. et al., Phys. Rev. Lett. **94**, 091803 (2005).

[3] Bellavance, A. M., *W / z production cross sections and asymmetries at e(cm) = 2-tev*. Preprint hep-ex/0506025, 2005.

[4] Abazov, V. M. et al., Phys. Rev. Lett. **95**, 051802 (2005).

[5] Taffard, A., *Top production cross section from cdf.* To appear in the proceedings of 32nd International Conference on High-Energy Physics (ICHEP 04), Beijing, China, 16-22 Aug 2004.

[6] Bortoletto, D., *Searching for susy at the tevatron.* Preprint hep-ex/0412013, 2004.

[7] Heister, A. et al., Eur. Phys. J. **C38**, 147 (2004).

[8] Abbiendi, G. et al., Phys. Lett. **B609**, 212 (2005).

[9] Heister, A. et al., Eur. Phys. J. **C28**, 1 (2003).

[10] Abdallah, J. et al., Eur. Phys. J. **C37**, 405 (2004).

[11] Chekanov, S. et al., Phys. Rev. **D67**, 012007 (2003).

[12] Giele, W. and Whalley, M. R., *Lhapdf v3: Les houches accord pdf.* Available on `http://durpdg.dur.ac.uk/lhapdf/`.

[13] Pumplin, J. et al., JHEP **07**, 012 (2002).

[14] Martin, A. D. et al., Eur. Phys. J **C23**, 73 (2002).

[15] G. Laštovička-Medin et al., *Precision limits for hera dis cross section measurement*. These proceedings.

[16] Corcella, G. et al., JHEP **01**, 010 (2001).

[17] Frixione, S. and Webber, B. R., JHEP **06**, 029 (2002).

[18] Stirling, W. J., private communication.

[19] Barlow, R., ECONF **C030908**, WEMT002 (2003).

[20] Altarelli, G. and Parisi, G., Nucl.Phys. **B126**, 298 (1977).

[21] Gribov, V.N. and Lipatov, L.N., Sov.J.Nucl.Phys **15**, 438 (1972).

[22] Lipatov, L. N., Sov.J.Nucl.Phys **20**, 94 (1975).

[23] Dokshitzer, Yu. L., JETP **46**, 641 (1977).

[24] Lipatov, L. N., Sov. J. Nucl. Phys. **23**, 338 (1976).

[25] Kuraev, E. A. and Lipatov, L. N. and Fadin, Victor S., Sov. Phys. JETP **45**, 199 (1977).

[26] Balitsky, I. I. and Lipatov, L. N., Sov. J. Nucl. Phys. **28**, 822 (1978).

[27] Devenish, R. C. E. and Cooper-Sarkar, A. M., *Deep Inelastic Scattering*. Oxford Unviersity Press, Oxford, 2004.

[28] Campbell, J. M. and Ellis, R. K., *Mcfm v4.0: A monte carlo for femtobarn processes at hadron colliders*. Available on `http://mcfm.fnal.gov//`.

[29] Stump, D. et al., JHEP **10**, 046 (2003).

[30] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C28**, 455 (2003).

[31] Jones, R. W. L. and Ford, M. and Salam, G. P. and Stenzel, Hasko and Wicke, Daniel, JHEP **12**, 007 (2003).

[32] Anastasiou, C. and Dixon, L. J. and Melnikov, K. and Petriello, F., Phys. Rev. **D69**, 094008 (2004).

[33] Dixon, L., private communication.

[34] Brun, R. and Rademakers, F., Nucl. Instrum. Meth. **A389**, 81 (1997).

[35] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Phys. Lett. **B531**, 216 (2002).

[36] Alekhin, S., Phys. Rev. **D68**, 014002 (2003).

Experimental determination of Parton Distributions

T. Carli, A. Cooper-Sarkar, J. Feltesse, A. Glazov, C. Gwenlan, M. Klein, T. Laštovička G. Laštovička-Medin, S. Moch, B. Reisert G. Salam, F. Siegert

1 Introduction [1]

With HERA currently in its second stage of operation, it is possible to assess the potential precision limits of HERA data and to estimate the potential impact of the measurements which are expected at HERA-II, in particular with respect to the PDF uncertainties.

Precision limits of the structure function analyses at HERA are examined in [1]. Since large amounts of luminosity are already collected, the systematic uncertainty becomes most important. A detailed study of error sources with particular emphasis on correlated errors for the upcoming precision analysis of the inclusive DIS cross section at low Q^2 using 2000 data taken by the H1 experiment is presented. A new tool, based on the ratio of cross sections measured by different reconstruction methods, is developed and its ability to qualify and unfold various correlated error sources is demonstrated.

An important issue is the consistency of the HERA data. In section 3, the H1 and ZEUS published PDF analyses are compared, including a discussion of the different treatments of correlated systematic uncertainties. Differences in the data sets and the analyses are investigated by putting the H1 data set through both PDF analyses and by putting the ZEUS and H1 data sets through the same (ZEUS) analysis, separately. Also, the HERA averaged data set (section 4) is put through the ZEUS PDF analysis and the result is compared to that obtained when putting the ZEUS and H1 data sets through this analysis together, using both the Offset and Hessian methods of treating correlated systematic uncertainties.

The HERA experimental data can not only be cross checked with respect to each other but also combined into one common dataset, as discussed in section 4. In this respect, a method to combine measurements of the structure functions performed by several experiments in a common kinematic domain is presented. This method generalises the standard averaging procedure by taking into account point-to-point correlations which are introduced by the systematic uncertainties of the measurements. The method is applied to the neutral and charged current DIS cross section data published by the H1 and ZEUS collaborations. The averaging improves in particular the accuracy due to the cross calibration of the H1 and ZEUS measurements.

The flavour decomposition of the light quark sea is discussed in [2]. For low x and thus low Q^2 domain at HERA only measurement of the photon exchange induced structure functions F_2 and F_L is possible, which is insufficient to disentangle individual quark flavours. A general strategy in this case is to assume flavour symmetry of the sea. [2] considers PDF uncertainties if this assumption is released. These uncertainties can be significantly reduced if HERA would run in deuteron-electron collision mode.

The impact of projected HERA-II data on PDFs is estimated in section 7. In particular, next-to-leading order (NLO) QCD predictions for inclusive jet cross sections at the LHC centre-of-mass energy are presented using the estimated PDFs. A further important measurement which could improve understanding of the gluon density at low x and, at the same time, provide consistency checks of the low Q^2 QCD evolution is the measurement of the longitudinal structure function F_L. Perspectives of this measurement are examined in section 5, while the impact of this measurement is also estimated in section 7.

Further improvements for consistently including final-state observables in global QCD analyses are discussed in section 8. There, a method for "a posteriori" inclusion of PDFs, whereby the Monte Carlo run calculates a grid (in x and Q) of cross section weights that can subsequently be combined with

[1] Subsection coordinators: A. Glazov, S. Moch

an arbitrary PDF. The procedure is numerically equivalent to using an interpolated form of the PDF. The main novelty relative to prior work is the use of higher-order interpolation, which substantially improves the tradeoff between accuracy and memory use. An accuracy of about 0.01% has been reached for the single inclusive cross-section in the central rapidity region $|y| < 0.5$ for jet transverse momenta from 100 to $5000\mathrm{GeV}$. This method will make it possible to consistently include measurements done at HERA, Tevatron and LHC in global QCD analyses.

2 Precision Limits for HERA DIS Cross Section Measurement [2]

The published precision low Q^2 cross section data [3] of the H1 experiment became an important data set in various QCD fit analyses [3, 4, 5, 6]. Following success of these data the H1 experiment plans to analyse a large data sample, taken during 2000 running period[3], in order to reach precision limits of low Q^2 inclusive cross sections measurements at HERA. The precision is expected to approach 1% level.

The aim of this contribution is to calculate realistic error tables for 2000 H1 data and pursue paths how to reach such a high precision. Correlated error sources are studied in particular and a new tool, based on the ratio of cross sections measured by different reconstruction methods, is developed. All errors, including correlated errors, are treated in the same manner as in [3]. Error tables are provided and used in QCD fit analysis, see Sec 7, in order to study the impact of the new data on PDFs. The new data are expected to reach higher precision level than [3] due to the following reasons

- Larger data statistics - Statistical errors will decrease by factor of $1.5 - 2$, compared to [3], depending on the kinematic region.
- Very large Monte Carlo simulations (MC) - Due to a progress in computing a number of simulated events can be significantly increased in order to minimise statistical error of MC, to understand uncorrelated errors and to estimate correlated errors more precisely.
- During past years increasing knowledge, arriving from various H1 analyses, enabled better understanding of the detector and its components as well as improving quality of MC.
- Data taking in 2000 was particularly smooth. Both HERA and H1 were running at peak performance for HERA-I running period.

This contribution uses existing 2000 data and MC ntuples along with the full analysis chain. It applies all preliminary technical work done on these data, including calibration, alignment, trigger studies etc. Quoted errors are assumed to be achieved in the final version of analysis yet the analysis has not been finalised, all the numbers in the paper are preliminary and may change in the publication.

The uncertainties of the cross section measurement are divided into a number of different types. Namely, these are *statistical uncertainties* of the data, *uncorrelated systematics* and *correlated systematics*. The term 'correlated' refers to the fact that cross section measurements in kinematic bins are affected in a correlated way while different correlated systematic error sources are considered uncorrelated among each other. The classification of the systematic errors into types is sometimes straightforward (MC statistics is uncorrelated error source) but sometimes is rather arbitrary (radiative corrections are assumed to be uncorrelated error source). The main goal of this classification is to preserve correlation between data points while keeping the treatement as simple as possible.

The cross section uncertainties depend on the method used to reconstruct event kinematics. There are various methods existing, involving a measurement of the scattered electron as well as of the hadronic finale state. In the following two of them, so called *electron method* and *sigma method*, are employed [7]. The electron method uses only the measurement of the scattered electron, namely its energy and polar angle, while the sigma method uses both the scattered electron and the hadronic final state. An advantage of the sigma method is a proper treatment of QED radiation from the incoming beam electron (ISR).

[2]Contributing authors: G. Laštovička-Medin, A. Glazov, T. Laštovička

[3]Data statistics will be increased further by adding data taken in year 1999.

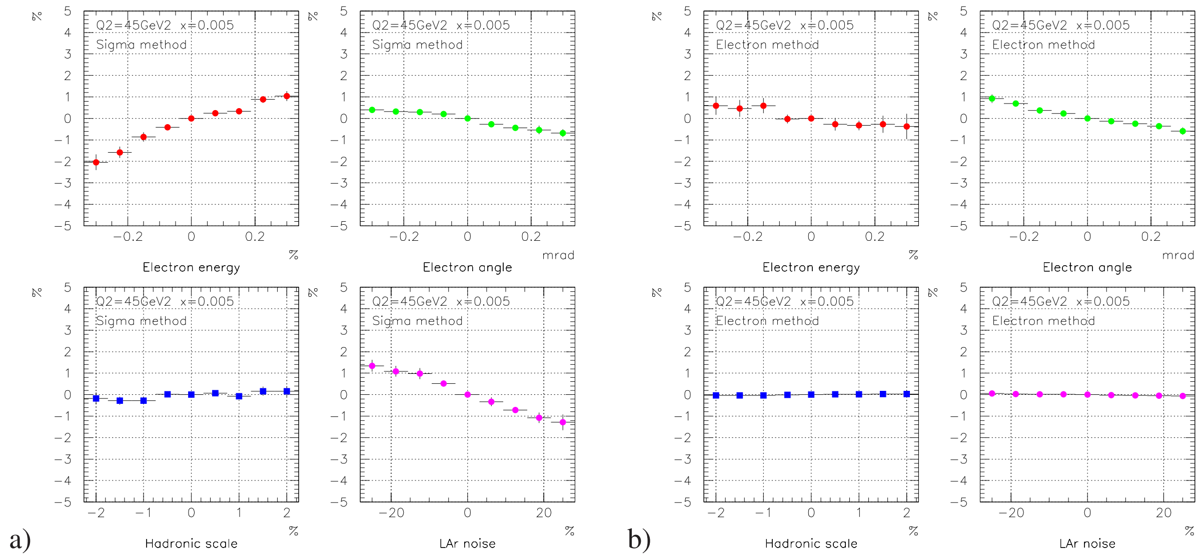


Fig. 1: A scan of the cross section measurement change in % depending on a variation of (from top-left) electron energy, electron polar angle, hadronic final state calibration scale and noise level in LAr calorimeter (bottom-right). The sigma method (a) and the electron method (b) were used to reconstruct kinematics of events.

The *statistical uncertainty* of the data is typically 0.5-1%, depending on the kinematic region analysed and the definition of the kinematic bins. In the following we adapt the bin definition used in [3], apart from merging bins at low y which was done in the published data in order to increase statistics.

The *uncorrelated systematics* consists from various contributions. A cross section uncertainty due to the Monte Carlo statistics is the one with very good potential to be minimised. In the following we assume 100 million simulated events to be used in analysis of 2000 data. Estimates were calculated with available 12 million simulated events and corresponding statistical errors scaled by a factor of $\sqrt{100/12}$. As a result the uncertainty is very small and typically on the level of few permile.

Additional contributions to the uncorrelated systematics are efficiencies. We assume for trigger efficiency 0.3% and backward tracker tracker efficiency 0.3% uncertainty. Radiative corrections are expected to affect the final cross section by 0.4%.

Effect of *correlated uncertainties* on the cross section measurement is studied in the following manner. Particular source of correlated uncertainty, for instance the scattered electron energy measurement, is varied by assumed error and the change of the measured cross section is quoted as the corresponding cross section measurement error. An example of cross section change on various correlated error source is shown in Fig. 1 for bin of $Q^2 = 45\,\mathrm{GeV}^2$ and $x = 0.005$. The kinematics of events was reconstructed with the sigma method (a) and the electron method (b). Errors are calculated as so called standard errors of the mean in calculation of which the available Monte Carlo sample was split into nine sub-samples. It is clearly seen that the cross section measurement with the sigma method in this kinematic bin is particularly sensitive to the electron energy measurement (top-left) and to noise description in LAr calorimeter (bottom-right). On the contrary, the electron polar angle measurement and the calibration of the hadronic final state play a little role. The electron method is mainly sensitive to the electron energy measurement. The importance of the systematic sources vary from bin to bin.

There are five individual sources contributing to the correlated cross section uncertainties:

- Uncertainties of 0.15% at $E_e = 27\,\mathrm{GeV}$ and 1% at 7 GeV are assigned to the electron energy scale for the backward calorimeter. The uncertainty is treated as a linear function of E_e interpolating between the results at 27 GeV and 7 GeV.
- The uncertainty on the scattered electron polar angle measurement is 0.3 mrad . The corresponding error on the cross section measurement is typically well below 1% but may be larger at lowest

Q^2	x	y	σ_r	R	F_2	Tot.(%)	Sta.	Uncorr.	Corr.	E_e	θ	Ehad	Noise	γp
25	0.0005	0.493	1.391	0.261	1.449	0.88	0.47	0.63	0.41	0.19	0.21	0.22	0.15	0.13
25	0.0008	0.308	1.251	0.261	1.268	0.91	0.43	0.62	0.51	0.34	0.37	0.02	0.04	0
25	0.0013	0.19	1.138	0.248	1.143	0.94	0.44	0.62	0.56	0.45	0.33	0.03	0.02	0
25	0.002	0.123	1.041	0.236	1.042	0.9	0.45	0.62	0.47	0.13	0.45	0.03	0.05	0
25	0.0032	0.077	0.842	0.254	0.843	1.42	0.5	0.63	1.17	0.74	0.36	0.17	0.8	0
25	0.005	0.049	0.745	0.243	0.745	1.17	0.52	0.63	0.83	0.59	0.42	0.25	0.33	0
25	0.008	0.031	0.667	0.225	0.667	1.22	0.56	0.64	0.87	0.43	0.35	0.66	0.09	0
25	0.013	0.019	0.586	0.214	0.586	2.02	0.65	0.66	1.8	0.67	0.57	1.43	0.65	0
25	0.02	0.012	0.569	0.159	0.569	5.77	0.86	0.71	5.66	0.83	0.52	3.51	4.33	0
25	0.032	0.008	0.553	0.065	0.553	10.64	1.34	0.88	10.52	0.93	0.64	3.86	9.72	0

Table 1: An example of the error table for $Q^2 = 25\,\mathrm{GeV}^2$ for 2000 data, large Monte Carlo sample and suppressed systematic errors compared to [1], see text for details. Absolute errors are shown. The table format is identical to the one published in [1].

Q^2	x	y	σ_r	R	F_2	Tot.(%)	Sta.	Uncorr.	Corr.	E_e	θ	Ehad	Noise	γp
25	0.0005	0.553	1.345	0.248	1.417	2.41	1.04	1.81	1.21	-1.04	-0.37	0.25	0.04	-0.41
25	0.0008	0.346	1.242	0.243	1.263	1.94	0.67	1.62	0.85	-0.6	-0.6	0.04	0.02	-0.07
25	0.0013	0.213	1.091	0.238	1.097	1.78	0.66	1.36	0.93	-0.64	-0.69	0	0	0
25	0.002	0.138	0.985	0.236	0.987	2.89	0.76	1.43	2.4	1.78	-0.7	0.17	1.34	0
25	0.0032	0.086	0.879	0.234	0.88	2.78	0.79	1.46	2.23	1.8	-0.77	-0.23	0.92	0
25	0.005	0.055	0.754	0.234	0.754	2.38	0.85	1.49	1.64	1.01	-0.58	0.16	1.03	0
25	0.008	0.034	0.663	0.234	0.663	2.52	0.92	1.54	1.78	1.11	-0.68	-0.72	0.84	0
25	0.0158	0.018	0.547	0.226	0.547	3.71	0.85	1.49	3.29	1.36	-0.88	-2.44	-1.42	0
25	0.05	0.005	0.447	0.148	0.447	7.54	1.28	3.35	6.64	0.99	-0.68	-3.28	-5.62	0

Table 2: An example of the full error table for $Q^2 = 25\,\mathrm{GeV}^2$, published H1 data. The definition of kinematic bins is not identical to that in Table 1, some bins were merged to enlarge statistics.

values of Q^2.

- The uncertainty on the hadronic energy scale comprises a number of systematic error sources corresponding to the $E-p_z$ decomposition: an uncertainty of the hadronic energy scale calibration of 2% for the central and forward calorimeter, an uncertainty of 3% for the fraction carried by tracks and a 5% uncertainty of the hadronic energy scale measured in backward calorimeter.
- The uncertainty on the hadronic energy scale is further affected by the subtracted noise in the calorimetery. The noise is described to the level of 10% and the corresponding error is propagated to the cross section uncertainty. The largest influence is in the low y region, which is measured with the sigma method.
- The uncertainty due to the photoproduction background at large y is estimated from the normalisation error of the PHOJET simulations to about 10%. At low and medium values of $y \lesssim 0.5$ it is negligible.

The total systematic error is calculated from the quadratic summation over all sources of the uncorrelated and correlated systematic uncertainties. The total error of the DIS cross section measurement is obtained from the statistical and systematical errors added in quadrature.

An example of the full error table for kinematic bin of $Q^2 = 25\,\mathrm{GeV}^2$ is shown in Table 1. For a comparison the corresponding part of the published data from [3] is presented in Table 2. One can see that precision about 1% can be reached especially in four lowest x bins, where the electron method was used to reconstruct the event kinematics. The key contributions to the seen improvement in the cross section measurement precision are the electron energy measurement, very large Monte Carlo statistics, well understood noise in LAr calorimeter and precisely controlled efficiencies entering the analysis.

Full error table, covering the kinematic region of $5 \leq Q^2 \leq 150\,\mathrm{GeV}^2$ and $0.01 \leq y \leq 0.6$ was produced. The electron method was applied for kinematic bins at $y > 0.1$ while the sigma method otherwise. The measurement of the proton structure function F_2 was simulated using fractal parametri-

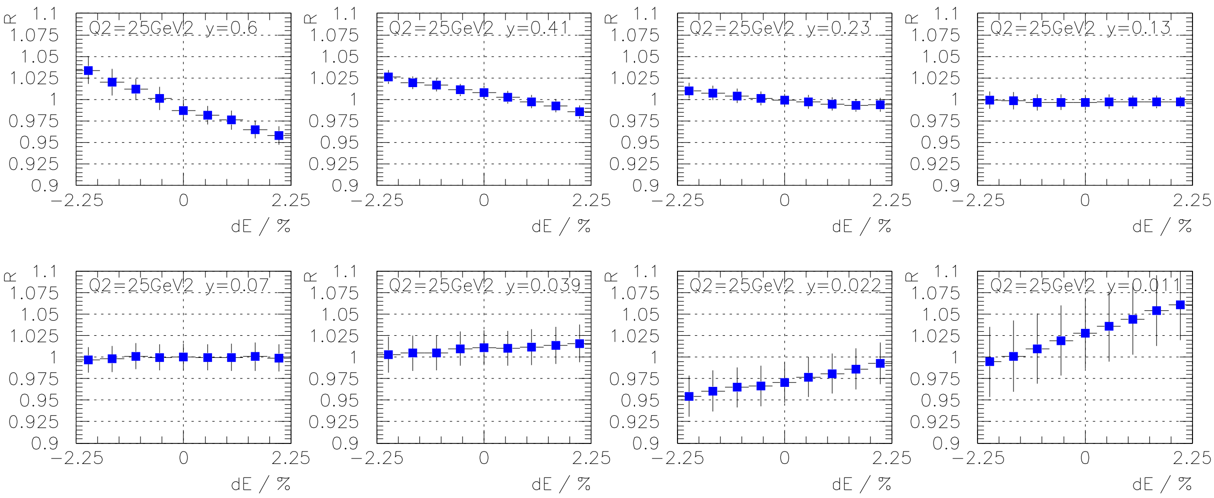


Fig. 2: A scan of the cross section ratio R in bins of Q^2 and y as a function of the hadronic final state calibration variation.

sation [8] for central values, accounting for all sources of correlated and uncorrelated errors. This table was used to estimate effect of precise low Q^2 data on the determination of proton PDFs from QCD fits.

The fact that different kinematics reconstruction methods are affected differently by the correlated systematic uncertainties may be employed as a tool to estimate these uncertainties. We define

$$R_i = \frac{\sigma_r^{el,i}}{\sigma_r^{\Sigma,i}} \tag{1}$$

to be the cross section measurement ratio, where the reduced cross section $\sigma_r^{el,i}$ and $\sigma_r^{\Sigma,i}$ is measured using the electron method and the sigma method, respectively. Kinematic bins, indexed by i, cover a region of the analysis phase space where both reconstruction methods are applicable for the measurement. The statistical error of R_i measurement is again evaluated by splitting the sample to a number of sub-samples and calculating the standard error of the mean. An example of a scan of the cross section ratio R_i dependence on the hadronic final state calibration variation in a bin of $Q^2 = 25\,\mathrm{GeV}^2$ and various inelasticity y is shown in Fig. 2.

An error of a particular correlated uncertainty source j can be estimated by searching for lowest $\chi^2 = \sum_i (R_i(\alpha_j) - 1)^2/\sigma_i^2$, where summation runs over kinematic bins, σ_i is the error of R_i measurement and α_j is the variation of the source j. However, since there is a number of correlated error sources the correct way to find correlated uncertainties is account for all of them.

Unfolding of the correlated error sources can be linearised and directly solved by minimising the following function:

$$\mathcal{L} = \sum_i \frac{1}{\sigma_i^2}(R_i + \sum_j \alpha_j \frac{\partial R_i}{\partial \alpha_j} - 1)^2. \tag{2}$$

The partial derivatives $\frac{\partial R_i}{\partial \alpha_j}$ for systematic source α_j are obtained from linear fits to distributions as shown in Fig. 2. Parameters α_j and their respective errors are obtained by matrix inversion technique.

The procedure was tested on available Monte Carlo sample for 2000 H1 data. Half of the sample, six million events, was used to simulate data. Full analysis chain was applied to measure the cross section and thus R_i. Kinematic bins were selected according to $15 \leq Q^2 \leq 60\,\mathrm{GeV}^2$ and $0.011 \leq y \leq 0.6$, i.e. in the main region of the data. The results are shown in Fig. 3. Closed points correspond to unfolded errors of the electron energy measurement (top-left), hadronic final state calibration and noise in the LAr calorimeter (bottom-left). There is no sensitivity observed to the electron polar angle measurement. All values are within statistical errors compatible with zero, as expected. For the final analysis the statistical

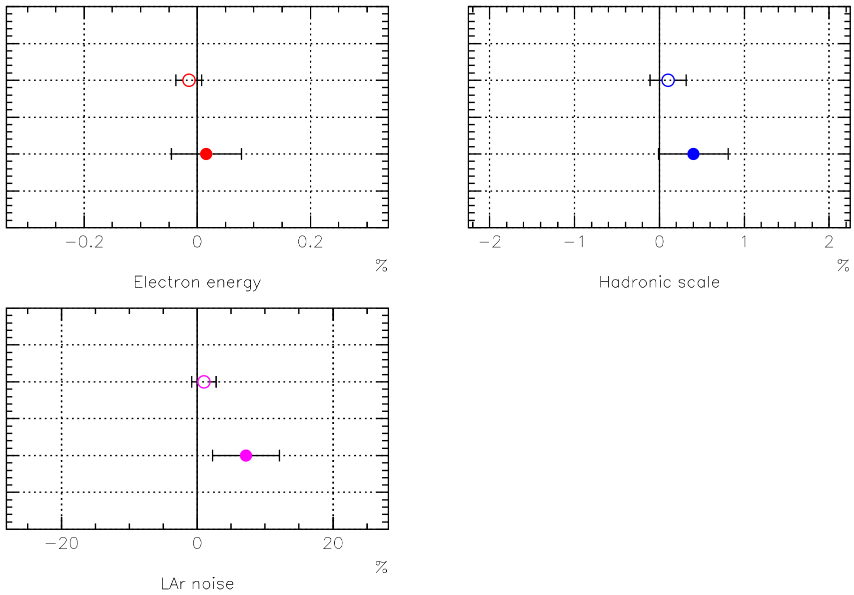


Fig. 3: Errors on the electron energy measurement (top-left), hadronic scale calibration (top-right) and noise in LAr calorimeter (bottom-left). Open points correspond to χ^2 scan in one correlated error source. Closed points show the result of complete unfolding, taking into account correlations.

errors are expected to be approximately three times smaller due to the significantly larger statistics than used for the presented study. This will enable the method to gain sufficient control over systematic correlated errors. Apart from being able to evaluate calibration of the scattered electron and of the hadronic final state, it gives a very good handle on the LAr calorimeter noise.

For a comparison, open points in Fig. 3 correspond to a χ^2 scan in one correlated error source. The statistical errors are smaller, as expected, and compatible with zero. However, the unfolding method is preferred since it takes into account all correlated error sources correctly.

In summary, a study of the DIS cross section uncertainties realistically achievable at HERA has been performed. For $x \in 0.001 - 0.01$ a precision of 1% can be reached across for a wide range of $Q^2 \in 5 - 150$ GeV^2, allowing improved estimate of W, Z production cross section in the central rapidity region of LHC. The accuracy of the DIS cross section measurement can be verified using different kinematic reconstruction methods available at the HERA collider.

3 Comparison and combination of ZEUS and H1 PDF analyses [4]

Parton Density Function (PDF) determinations are usually global fits [5, 4, 9], which use fixed target DIS data as well as HERA data. In such analyses the high statistics HERA NC e^+p data, which span the range $6.3 \times 10^{-5} < x < 0.65, 2.7 < Q^2 < 30,000\text{GeV}^2$, have determined the low-x sea and gluon distributions, whereas the fixed target data have determined the valence distributions and the higher-x sea distributions. The ν-Fe fixed target data have been the most important input for determining the valence distributions, but these data suffer from uncertainties due to heavy target corrections. Such uncertainties are also present for deuterium fixed target data, which have been used to determine the shape of the high-x d-valence quark.

HERA data on neutral and charged current (NC and CC) e^+p and e^-p inclusive double differential cross-sections are now available, and have been used by both the H1 and ZEUS collaborations [10, 11] in order to determine the parton distributions functions (PDFs) using data from within a single experiment. The HERA high Q^2 cross-section data can be used to determine the valence distributions, thus

[4]Contributing authors: A. Cooper-Sarkar, C. Gwenlan

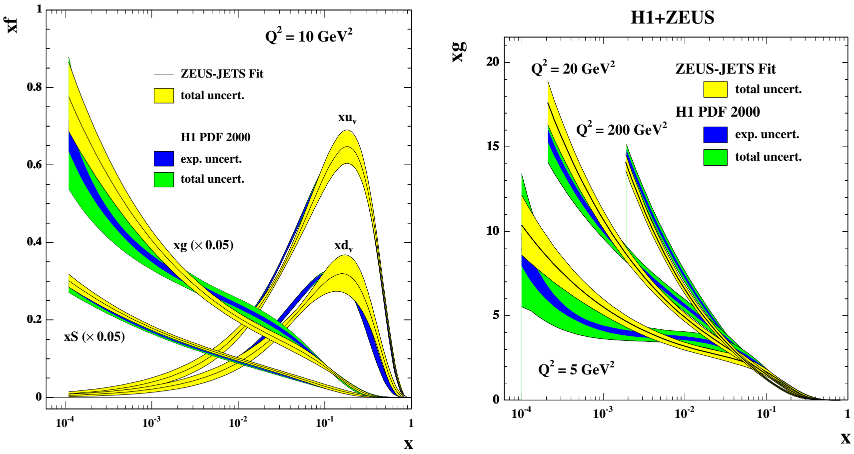


Fig. 4: Left plot: Comparison of PDFs from ZEUS and H1 analyses at $Q^2 = 10\text{GeV}^2$. Right plot: Comparison of gluon from ZEUS and H1 analyses, at various Q^2. Note that the ZEUS analysis total uncertainty includes both experimental and model uncertainties.

eliminating uncertainties from heavy target corrections. The PDFs are presented with full accounting for uncertainties from correlated systematic errors (as well as from statistical and uncorrelated sources). Peforming an analysis within a single experiment has considerable advantages in this respect, since the global fits have found significant tensions between different data sets, which make a rigorous statistical treatment of uncertainties difficult.

Fig. 4 compares the results of the H1 and ZEUS analyses. Whereas the extracted PDFs are broadly compatible within errors, there is a noticeable difference in the shape of the gluon PDFs. Full details of the analyses are given in the relevant publications, in this contribution we examine the differences in the two analyses, recapping only salient details.

3.1 Comparing ZEUS and H1 published PDF analyses

The kinematics of lepton hadron scattering is described in terms of the variables Q^2, the invariant mass of the exchanged vector boson, Bjorken x, the fraction of the momentum of the incoming nucleon taken by the struck quark (in the quark-parton model), and y which measures the energy transfer between the lepton and hadron systems. The differential cross-section for the NC process is given in terms of the structure functions by

$$\frac{d^2\sigma(e^{\pm}p)}{dxdQ^2} = \frac{2\pi\alpha^2}{Q^4 x}\left[Y_+\,F_2(x,Q^2) - y^2\,F_L(x,Q^2) \mp Y_-\,xF_3(x,Q^2)\right], \tag{3}$$

where $Y_{\pm} = 1 \pm (1-y)^2$. The structure functions F_2 and xF_3 are directly related to quark distributions, and their Q^2 dependence, or scaling violation, is predicted by pQCD. At $Q^2 \leq 1000$ GeV^2 F_2 dominates the charged lepton-hadron cross-section and for $x \leq 10^{-2}$, F_2 itself is sea quark dominated but its Q^2 evolution is controlled by the gluon contribution, such that HERA data provide crucial information on low-x sea-quark and gluon distributions. At high Q^2, the structure function xF_3 becomes increasingly important, and gives information on valence quark distributions. The CC interactions enable us to separate the flavour of the valence distributions at high-x, since their (LO) cross-sections are given by,

$$\frac{d^2\sigma(e^+p)}{dxdQ^2} = \frac{G_F^2 M_W^4}{(Q^2+M_W^2)^2 2\pi x} x\left[(\bar{u}+\bar{c}) + (1-y)^2(d+s)\right],$$

$$\frac{d^2\sigma(e^-p)}{dxdQ^2} = \frac{G_F^2 M_W^4}{(Q^2+M_W^2)^2 2\pi x} x\left[(u+c) + (1-y)^2(\bar{d}+\bar{s})\right].$$

For both HERA analyses the QCD predictions for the structure functions are obtained by solving the DGLAP evolution equations [12, 13, 14, 15] at NLO in the $\overline{\mathrm{MS}}$ scheme with the renormalisation and factorization scales chosen to be Q^2. These equations yield the PDFs at all values of Q^2 provided they are input as functions of x at some input scale Q_0^2. The resulting PDFs are then convoluted with coefficient functions, to give the structure functions which enter into the expressions for the cross-sections. For a full explanation of the relationships between DIS cross-sections, structure functions, PDFs and the QCD improved parton model see ref. [16].

The HERA data are all in a kinematic region where there is no sensitivity to target mass and higher twist contributions but a minimum Q^2 cut must be imposed to remain in the kinematic region where perturbative QCD should be applicable. For ZEUS this is $Q^2 > 2.5$ GeV^2, and for H1 it is $Q^2 > 3.5$ GeV^2. Both collaborations have included the sensitivity to this cut as part of their model errors.

In the ZEUS analysis, the PDFs for u valence, $xu_v(x)$, d valence, $xd_v(x)$, total sea, $xS(x)$, the gluon, $xg(x)$, and the difference between the d and u contributions to the sea, $x(\bar{d} - \bar{u})$, are each parametrized by the form

$$p_1 x^{p_2} (1-x)^{p_3} P(x), \tag{4}$$

where $P(x) = 1 + p_4 x$, at $Q_0^2 = 7\mathrm{GeV}^2$. The total sea $xS = 2x(\bar{u} + \bar{d} + \bar{s} + \bar{c} + \bar{b})$, where $\bar{q} = q_{sea}$ for each flavour, $u = u_v + u_{sea}, d = d_v + d_{sea}$ and $q = q_{sea}$ for all other flavours. The flavour structure of the light quark sea allows for the violation of the Gottfried sum rule. However, there is no information on the shape of the $\bar{d} - \bar{u}$ distribution in a fit to HERA data alone and so this distribution has its shape fixed consistent with the Drell-Yan data and its normalisation consistent with the size of the Gottfried sum-rule violation. A suppression of the strange sea with respect to the non-strange sea of a factor of 2 at Q_0^2, is also imposed consistent with neutrino induced dimuon data from CCFR. Parameters are further restricted as follows. The normalisation parameters, p_1, for the d and u valence and for the gluon are constrained to impose the number sum-rules and momentum sum-rule. The p_2 parameter which constrains the low-x behaviour of the u and d valence distributions is set equal, since there is no information to constrain any difference. When fitting to HERA data alone it is also necessary to constrain the high-x sea and gluon shapes, because HERA-I data do not have high statistics at large-x, in the region where these distributions are small. The sea shape has been restricted by setting $p_4 = 0$ for the sea, but the gluon shape is constrained by including data on jet production in the PDF fit. Finally the ZEUS analysis has 11 free PDF parameters. ZEUS have included reasonable variations of these assumptions about the input parametrization in their analysis of model uncertainties. The strong coupling constant was fixed to $\alpha_s(M_Z^2) = 0.118$ [17]. Full account has been taken of correlated experimental systematic errors by the Offset Method, as described in ref [9, 18].

For the H1 analysis, the value of $Q_0^2 = 4\mathrm{GeV}^2$, and the choice of quark distributions which are parametrized is different. The quarks are considered as u-type and d-type with different parametrizations for, $xU = x(u_v + u_{sea} + c)$, $xD = x(d_v + d_{sea} + s)$, $x\bar{U} = x(\bar{u} + \bar{c})$ and $x\bar{D} = x(\bar{d} + \bar{s})$, with $q_{sea} = \bar{q}$, as usual, and the the form of the quark and gluon parametrizations given by Eq. 4. For $x\bar{D}$ and $x\bar{U}$ the polynomial, $P(x) = 1.0$, for the gluon and xD, $P(x) = (1+p_4 x)$, and for xU, $P(x) = (1+p_4 x + p_5 x^3)$. The parametrization is then further restricted as follows. Since the valence distributions must vanish as $x \to 0$, the low-x parameters, p_1 and p_2 are set equal for xU and $x\bar{U}$, and for xD and $x\bar{D}$. Since there is no information on the flavour structure of the sea it is also necessary to set p_2 equal for $x\bar{U}$ and $x\bar{D}$. The normalisation, p_1, of the gluon is determined from the momentum sum-rule and the p_4 parameters for xU and xD are determined from the valence number sum-rules. Assuming that the strange and charm quark distributions can be expressed as x independent fractions, f_s and f_c, of the d and u type sea, gives the further constraint $p_1(\bar{U}) = p_1(\bar{D})(1-f_s)/(1-f_c)$. Finally there are 10 free parameters. H1 has also included reasonable variations of these assumptions in their analysis of model uncertainties. The strong coupling constant was fixed to $\alpha_s(M_Z^2) = 0.1185$ and this is sufficiently similar to the ZEUS choice that we can rule it out as a cause of any significant difference. Full account has been taken of correlated

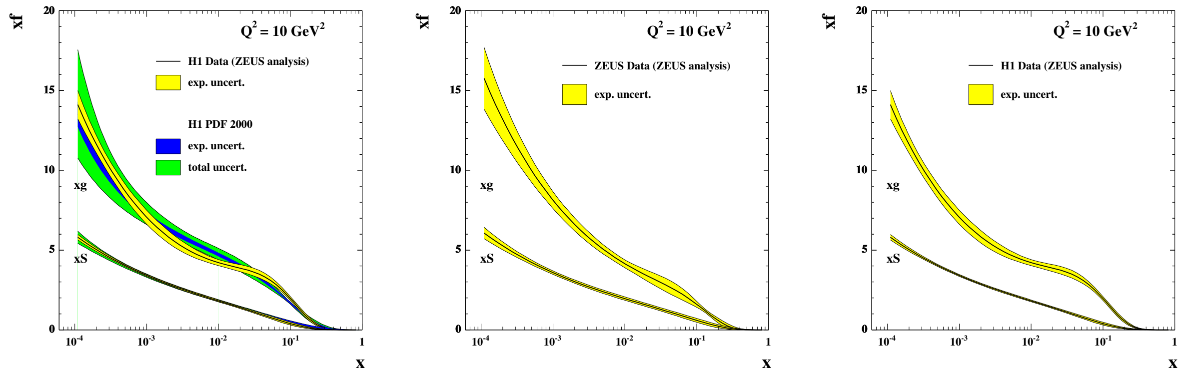


Fig. 5: Sea and gluon distributions at $Q^2 = 10\mathrm{GeV}^2$ extracted from different data sets and different analyses. Left plot: H1 data put through both ZEUS and H1 analyses. Middle plot: ZEUS data put through ZEUS analysis. Right plot: H1 data put through ZEUS analysis.

experimental systematic errors by the Hessian Method, see ref. [18].

For the ZEUS analysis, the heavy quark production scheme used is the general mass variable flavour number scheme of Roberts and Thorne [19]. For the H1 analysis, the zero mass variable flavour number scheme is used. It is well known that these choices have a small effect on the steepness of the gluon at very small-x, such that the zero-mass choice produces a slightly less steep gluon. However, there is no effect on the more striking differences in the gluon shapes at larger x.

There are two differences in the analyses which are worth further investigation. The different choices for the form of the PDF parametrization at Q_0^2 and the different treatment of the correlated experimental uncertainties.

3.2 Comparing different PDF analyses of the same data set and comparing different data sets using the same PDF analysis.

So far we have compared the results of putting two different data sets into two different analyses. Because there are many differences in the assumptions going into these analyses it is instructive to consider:(i) putting both data sets through the same analysis and (ii) putting one of the data sets through both analyses. For these comparisons, the ZEUS analysis does NOT include the jet data, so that the data sets are more directly comparable, involving just the inclusive double differential cross-section data. Fig. 5 compares the sea and gluon PDFs, at $Q^2 = 10\mathrm{GeV}^2$, extracted from H1 data using the H1 PDF analysis with those extracted from H1 data using the ZEUS PDF analysis. These alternative analyses of the same data set give results which are compatible within the model dependence error bands. Fig. 5 also compares the sea and gluon PDFs extracted from ZEUS data using the ZEUS analysis with those extracted from H1 data using the ZEUS analysis. From this comparison we can see that the different data sets lead to somewhat different gluon shapes even when put through exactly the same analysis. Hence the most of the difference in shape of the ZEUS and H1 PDF analyses can be traced back to a difference at the level of the data sets.

3.3 Comparing the Offset and Hessian method of assessing correlated experimental uncertainties

Before going further it is useful to discuss the treatment of correlated systematic errors in the ZEUS and H1 analyses. A full discussion of the treatment of correlated systematic errors in PDF analyses is given in ref [16], only salient details are recapped here. Traditionally, experimental collaborations have evaluated an overall systematic uncertainty on each data point and these have been treated as uncorrelated, such that they are simply added to the statistical uncertainties in quadrature when evaluating χ^2. However, modern deep inelastic scattering experiments have very small statistical uncertainties, so that the contribution of systematic uncertainties becomes dominant and consideration of point to point correlations between systematic uncertainties is essential.

For both ZEUS and H1 analyses the formulation of the χ^2 including correlated systematic uncertainties is constructed as follows. The correlated uncertainties are included in the theoretical prediction, $F_i(p,s)$, such that

$$F_i(p,s) = F_i^{\mathrm{NLOQCD}}(p) + \sum_\lambda s_\lambda \Delta_{i\lambda}^{\mathrm{sys}}$$

where, $F_i^{\mathrm{NLOQCD}}(p)$, represents the prediction from NLO QCD in terms of the theoretical parameters p, and the parameters s_λ represent independent variables for each source of systematic uncertainty. They have zero mean and unit variance by construction. The symbol $\Delta_{i\lambda}^{\mathrm{sys}}$ represents the one standard deviation correlated systematic error on data point i due to correlated error source λ. The χ^2 is then formulated as

$$\chi^2 = \sum_i \frac{[F_i(p,s) - F_i(\mathrm{meas})]^2}{\sigma_i^2} + \sum_\lambda s_\lambda^2 \tag{5}$$

where, $F_i(\mathrm{meas})$, represents a measured data point and the symbol σ_i represents the one standard deviation uncorrelated error on data point i, from both statistical and systematic sources. The experiments use this χ^2 in different ways. ZEUS uses the Offset method and H1 uses the Hessian method.

Traditionally, experimentalists have used ‘Offset’ methods to account for correlated systematic errors. The χ^2 is formluated without any terms due to correlated systematic errors ($s_\lambda = 0$ in Eq. 5) for evaluation of the central values of the fit parameters. However, the data points are then offset to account for each source of systematic error in turn (i.e. set $s_\lambda = +1$ and then $s_\lambda = -1$ for each source λ) and a new fit is performed for each of these variations. The resulting deviations of the theoretical parameters from their central values are added in quadrature. (Positive and negative deviations are added in quadrature separately.) This method does not assume that the systematic uncertainties are Gaussian distributed. An equivalent (and much more efficient) procedure to perform the Offset method has been given by Pascaud and Zomer [20], and this is what is actually used. The Offset method is a conservative method of error estimation as compared to the Hessian method. It gives fitted theoretical predictions which are as close as possible to the central values of the published data. It does not use the full statistical power of the fit to improve the estimates of s_λ, since it choses to mistrust the systematic error estimates, but it is correspondingly more robust.

The Hessian method is an alternative procedure in which the systematic uncertainty parameters s_λ are allowed to vary in the main fit when determining the values of the theoretical parameters. Effectively, the theoretical prediction is not fitted to the central values of the published experimental data, but these data points are allowed to move collectively, according to their correlated systematic uncertainties. The theoretical prediction determines the optimal settings for correlated systematic shifts of experimental data points such that the most consistent fit to all data sets is obtained. Thus, in a global fit, systematic shifts in one experiment are correlated to those in another experiment by the fit. In essence one is allowing the theory to calibrate the detectors. This requires great confidence in the theory, but more significantly, it requires confidence in the many model choices which go into setting the boundary conditions for the theory (such as the parametrization at Q_0^2).

The ZEUS analysis can be performed using the Hessian method as well as the Offset method and Fig. 6 compares the PDFs, and their uncertainties, extracted from ZEUS data using these two methods. The central values of the different methods are in good agreement but the use of the Hessian method results in smaller uncertainties, for a the standard set of model assumptions, since the input data can be shifted within their correlated systematic uncertainties to suit the theory better. However, model uncertainties are more significant for the Hessian method than for the Offset method. The experimental uncertainty band for any one set of model choices is set by the usual χ^2 tolerance, $\Delta\chi^2 = 1$, but the acceptability of a different set of choices is judged by the hypothesis testing criterion, such that the χ^2 should be approximately in the range $N \pm \sqrt{(2N)}$, where N is the number of degrees of freedom. The PDF parameters obtained for the different model choices can differ by much more than their experimental uncertainties, because each model choice can result in somewhat different values of the systematic

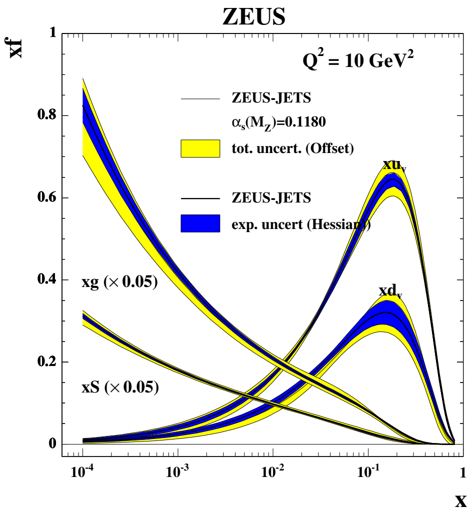


Fig. 6: PDFs at $Q^2 = 10\text{GeV}^2$, for the ZEUS analysis of ZEUS data performed by the Offset and the Hessian methods.

uncertainty parameters, s_λ, and thus a different estimate of the shifted positions of the data points. This results in a larger spread of model uncertainty than in the Offset method, for which the data points cannot move. Fig 4 illustrates the comparability of the ZEUS (Offset) total uncertainty estimate to the H1 (Hessian) experimental plus model uncertainty estimate.

Another issue which arises in relation to the Hessian method is that the data points should not be shifted far outside their one standard deviation systematic uncertainties. This can indicate inconsistencies between data sets, or parts of data sets, with respect to the rest of the data. The CTEQ collaboration have considered data inconsistencies in their most recent global fit [4]. They use the Hessian method but they increase the resulting uncertainty estimates, by increasing the χ^2 tolerance to $\Delta\chi^2 = 100$, to allow for both model uncertainties and data inconsistencies. In setting this tolerance they have considered the distances from the χ^2-minima of individual data sets to the global minimum for all data sets. These distances by far exceed the range allowed by the $\Delta\chi^2 = 1$ criterion. Strictly speaking such variations can indicate that data sets are inconsistent but the CTEQ collaboration take the view that all of the current world data sets must be considered acceptable and compatible at some level, even if strict statistical criteria are not met, since the conditions for the application of strict criteria, namely Gaussian error distributions, are also not met. It is not possible to simply drop "inconsistent" data sets, as then the partons in some regions would lose important constraints. On the other hand the level of "inconsistency" should be reflected in the uncertainties of the PDFs. This is achieved by raising the χ^2 tolerance. This results in uncertainty estimates which are comparable to those achieved by using the Offset method [18].

3.4 Using both H1 and ZEUS data in the same PDF analysis

Using data from a single experiment avoids questions of data consistency, but to get the most information from HERA it is necessary to put ZEUS and H1 data sets into the same analysis together, and then questions of consistency arise. Fig 7 compares the sea and gluon PDFs and the u and d valence PDFs extracted from the ZEUS PDF analysis of ZEUS data alone, to those extracted from the ZEUS PDF analysis of both H1 and ZEUS data. It is noticeable that, for the low-x sea and gluon PDFs, combining the data sets does not bring a reduction in uncertainty equivalent to doubling the statistics. This is because the data which determine these PDFs are systematics limited. In fact there is some degree of tension between the ZEUS and the H1 data sets, such that the χ^2 per degree of freedom rises for both data sets when they are fitted together. The Offset method of treating the systematic errors reflects this tension such that the overall uncertainty is not much improved when H1 data are added to ZEUS data. However, the uncertainty on the high-x valence distributions is reduced by the input of H1 data, since the data are still statistics limited at high x.

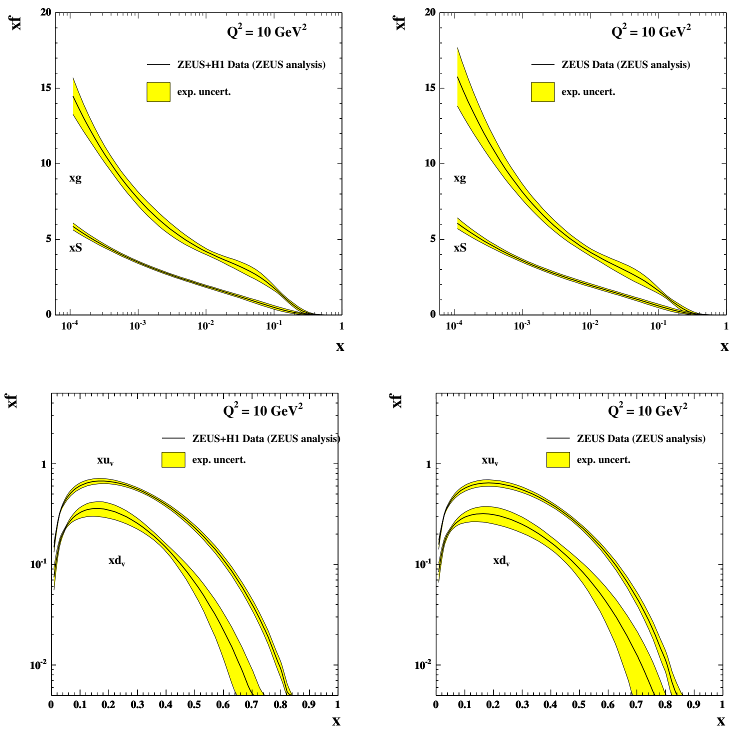


Fig. 7: Top plots: Sea and gluon distributions at $Q^2 = 10\mathrm{GeV}^2$ extracted from H1 and ZEUS data using the ZEUS analysis (left) compared to those extracted from ZEUS data alone using the ZEUS analysis (right). Bottom Plots: Valence distributions at $Q^2 = 10\mathrm{GeV}^2$, extracted from H1 and ZEUS data using the ZEUS analysis (left) compared to those extracted from ZEUS data alone using the ZEUS analysis (right).

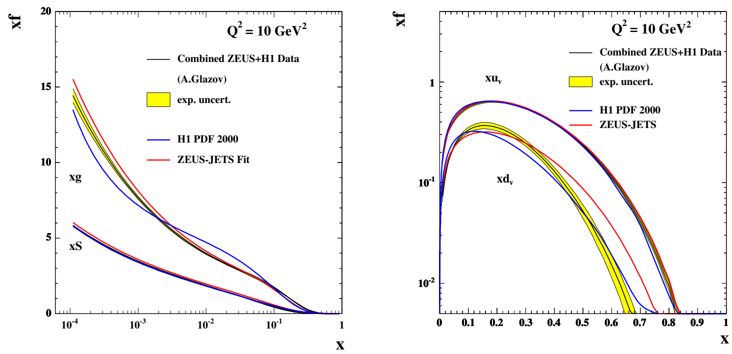


Fig. 8: Left plot: Sea and gluon distributions at $Q^2 = 10\text{GeV}^2$, extracted from the combined H1 and ZEUS data set using the ZEUS analysis. Right plot: Valence distributions at $Q^2 = 10\text{GeV}^2$, extracted from the combined H1 and ZEUS data set using the ZEUS analysis.

3.5 Combining the H1 and ZEUS data sets before PDF analysis

Thus there could be an advantage in combining ZEUS and H1 data in a PDF fit if the tension between the data sets could be resolved. It is in this context the question of combining these data into a single data set arises. The procedure for combination is detailed in the contribution of S. Glazov to these proceedings (section 4). Essentially, since ZEUS and H1 are measuring the same physics in the same kinematic region, one can try to combine them using a 'theory-free' Hessian fit in which the only assumption is that there is a true value of the cross-section, for each process, at each x, Q^2 point. The systematic uncertainty parameters, s_λ, of each experiment are fitted to determine the best fit to this assumption. Thus each experiment is calibrated to the other. This works well because the sources of systematic uncertainty in each experiment are rather different. Once the procedure has been performed the resulting systematic uncertainties on each of the combined data points are significantly smaller than the statistical errors. Thus one can legitimately make a fit to the combined data set in which these statistical and systematic uncertainties are simply combined in quadrature. The result of making such a fit, using the ZEUS analysis, is shown in Fig. 8. The central values of the ZEUS and H1 published analyses are also shown for comparison. Looking back to Fig. 7 one can see that there has been a dramatic reduction in the level of uncertainty compared to the ZEUS Offset method fit to the separate ZEUS and H1 data sets. This result is very promising. A preliminary study of model dependence, varying the form of the polynomial, $P(x)$, used in the PDF paremtrizations at Q_0^2, also indicates that model dependence is relatively small.

The tension between ZEUS and H1 data could have been resolved by putting them both into a PDF fit using the Hessian method to shift the data points. That is, rather than calibrating the two experiments to each other in the 'theory-free' fit, we could have used the theory of pQCD to calibrate each experiment. Fig. 9 shows the PDFs extracted when the ZEUS and H1 data sets are put through the ZEUS PDF analysis procedure using the Hessian method. The uncertainties on the resulting PDFs are comparable to those found for the fit to the combined data set, see Fig. 8. However, the central values of the resulting PDFs are rather different- particularly for the less well known gluon and d valence PDFs. For both of the fits shown in Figs. 8, 9 the values of the systematic error parameters, s_λ, for each experiment have been allowed to float so that the data points are shifted to give a better fit to our assumptions, but the values of the systematic error parameters chosen by the 'theory-free' fit and by the PDF fit are rather different. A representaive sample of these values is given in Table 3. These discrepancies might be somewhat alleviated by a full consideration of model errors in the PDF fit, or of appropriate χ^2 tolerance when combining the ZEUS and H1 experiments in a PDF fit, but these differences should make us wary about the uncritical use of the Hessian method.

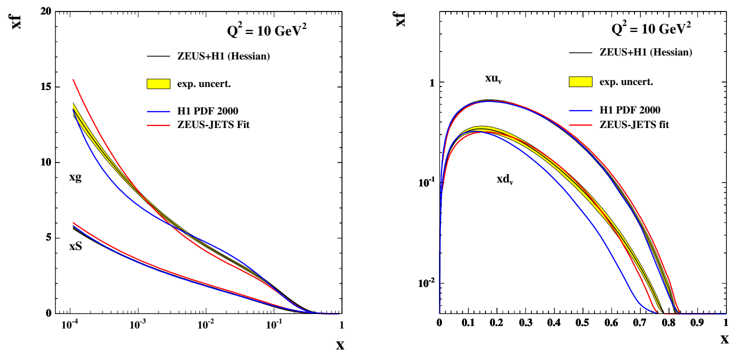


Fig. 9: Left plot: Sea and gluon distributions at $Q^2 = 10\mathrm{GeV}^2$, extracted from the H1 and ZEUS data sets using the ZEUS analysis done by Hessian method. Right plot: Valence distributions at $Q^2 = 10\mathrm{GeV}^2$, extracted from the H1 and ZEUS data sets using the ZEUS analysis done by Hessian method.

Syatematic uncertainty s_λ	in PDF fit	in Theory-free fit
ZEUS electron efficiency	1.68	0.31
ZEUS electron angle	-1.26	-0.11
ZEUS electron energy scale	-1.04	0.97
ZEUS hadron calorimeter energy scale	1.05	-0.58
H1 electron energy scale	-0.51	0.61
H1 hadron energy scale	-0.26	-0.98
H1 calorimeter noise	1.00	-0.63
H1 photoproduction background	-0.36	0.97

Table 3: Systematic shifts for ZEUS and H1 data as determine by a joint pQCD PDF fit, and as determined by the theory-free data combination fit

4 Averaging of DIS Cross Section Data [5]

The QCD fit procedures (Alekhin [6], CTEQ [4], MRST [5], H1 [11], ZEUS [9]) use data from a number of individual experiments directly to extract the parton distribution functions (PDF). All programs use both the central values of measured cross section data as well as information about the correlations among the experimental data points.

The direct extraction procedure has several shortcomings. The number of input datasets is large containing several individual publications. The data points are correlated because of common systematic uncertainties, within and also across the publications. Handling of the experimental data without additional expert knowledge becomes difficult. Additionally, as it is discussed in Sec. 3, the treatment of the correlations produced by the systematic errors is not unique. In the Lagrange Multiplier method [20] each systematic error is treated as a parameter and thus fitted to QCD. Error propogation is then used to estimate resulting uncertainties on PDFs. In the so-called "offset" method (see e.g. [9]) the datasets are shifted in turn by each systematic error before fitting. The resulting fits are used to form an envelope function to estimate the PDF uncertainty. Each method has its own advantages and shortcomings, and it is difficult to select the standard one. Finally, some global QCD analyses use non-statistical criteria to estimate the PDF uncertainties ($\Delta\chi^2 \gg 1$). This is driven by the apparent discrepancy between different experiments which is often difficult to quantify. Without a model independent consistency check of the data it might be the only safe procedure.

[5] Contributing author: A. Glazov

These drawbacks can be significantly reduced by averaging of the input structure function data in a model independent way before performing a QCD analysis of that data. One combined dataset of deep inelastic scattering (DIS) cross section measurements is much easier to handle compared to a scattered set of individual experimental measurements, while retaining the full correlations between data points. The averaging method proposed here is unique and removes the drawback of the offset method, which fixes the size of the systematic uncertainties. In the averaging procedure the correlated systematic uncertainties are floated coherently allowing in some cases reduction of the uncertainty. In addition, study of a global χ^2/dof of the average and distribution of the pulls allows a model independent consistency check between the experiments. In case of discrepancy between the input datasets, localised enlargement of the uncertainties for the average can be performed.

A standard way to represent a cross section measurement of a single experiment is given in the case of the F_2 structure function by:

$$\chi^2_{exp}(\left\{F_2^{i,true}\right\}, \{\alpha_j\}) \;=\; \sum_i \frac{\left[F_2^{i,true} - \left(F_2^i + \sum_j \frac{\partial F_2^i}{\partial \alpha_j}\alpha_j\right)\right]^2}{\sigma_i^2} + \sum_j \frac{\alpha_j^2}{\sigma_{\alpha_j}^2}. \qquad (6)$$

Here F_2^i (σ_i^2) are the measured central values (statistical and uncorrelated systematic uncertainties) of the F_2 structure function[6], α_j are the correlated systematic uncertainty sources and $\partial F_2^i/\partial\alpha_j$ are the sensitivities of the measurements to these systematic sources. Eq. 6 corresponds to the correlated probability distribution functions for the structure function $F_2^{i,true}$ and for the systematic uncertainties α_j. Eq. 6 resembles Eq. 5 where the theoretical predictions for F_2 are substituted by $F_2^{i,true}$.

The χ^2 function Eq. 6 by construction has a minimum $\chi^2 = 0$ for $F_2^{i,true} = F_2^i$ and $\alpha_j = 0$. One can show that the total uncertainty for $F_2^{i,true}$ determined from the formal minimisation of Eq. 6 is equal to the sum in quadrature of the statistical and systematic uncertainties. The reduced covariance matrix $cov(F_2^{i,true}, F_2^{j,true})$ quantifies the correlation between experimental points.

In the analysis of data from more than one experiment, the χ^2_{tot} function is taken as a sum of the χ^2 functions Eq. 6 for each experiment. The QCD fit is then performed in terms of parton density functions which are used to calculate predictions for $F_2^{i,true}$.

Before performing the QCD fit, the χ^2_{tot} function can be minimised with respect to $F_2^{i,true}$ and α_j. If none of correlated sources is present, this minimisation is equivalent to taking an average of the structure function measurements. If the systematic sources are included, the minimisation corresponds to a generalisation of the averaging procedure which contains correlations among the measurements.

Being a sum of positive definite quadratic functions, χ^2_{tot} is also a positive definite quadratic and thus has a unique minimum which can be found as a solution of a system of linear equations. Although this system of the equations has a large dimension it has a simple structure allowing fast and precise solution.

A dedicated program has been developed to perform this averaging of the DIS cross section data (`http://www.desy.de/~glazov/f2av.tar.gz`). This program can calculate the simultaneous averages for neutral current (NC) and charged current (CC) electron- and positron-proton scattering cross section data including correlated systematic sources. The output of the program includes the central values and uncorrelated uncertainties of the average cross section data. The correlated systematic uncertainties can be represented in terms of (i) covariance matrix, (ii) dependence of the average cross section on the original systematic sources together with the correlation matrix for the systematic sources, (iii) and finally the correlation matrix of the systematic sources can be diagonalised, in this case the form of χ^2 for the average data is identical to Eq. 6 but the original systematic sources are not preserved.

[6]The structure function is measured for different Q^2 (four momentum transfer squared) and Bjorken-x values which are omitted here for simplicity.

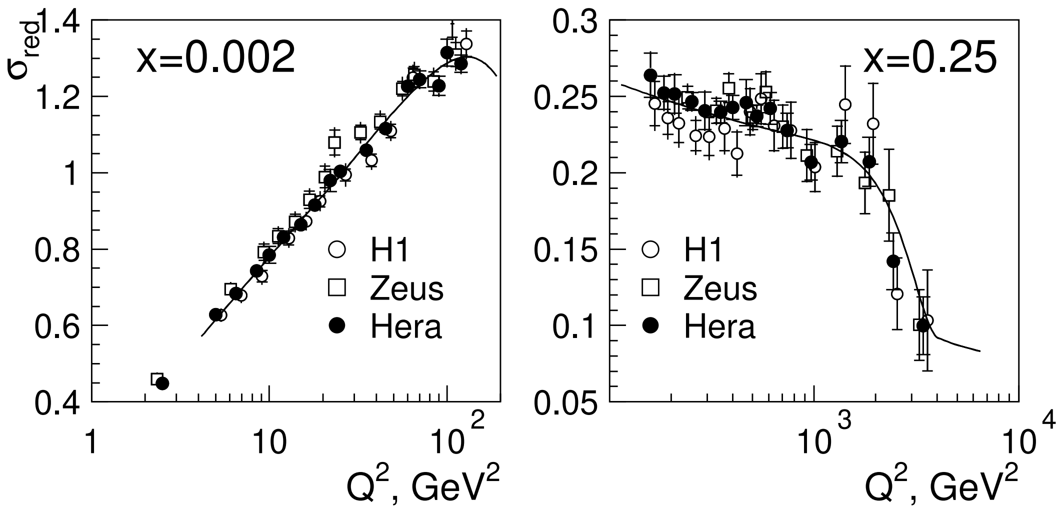


Fig. 10: Q^2 dependence of the NC reduced cross section for $x = 0.002$ and $x = 0.25$ bins. H1 data is shown as open circles, ZEUS data is shown as open squares and the average of H1 and ZEUS data is shown as filled circles. The line represents the expectation from the H1 PDF 2000 QCD fit.

The first application of the averaging program has been a determination of the average of the published H1 and ZEUS data [11, 3, 21, 22, 23, 24, 25, 26, 27, 28]. Nine individual NC and CC cross section measurements are included from H1 and seven are included from ZEUS. Several sources of systematic uncertainties are correlated between datasets, the correlations among H1 and ZEUS datasets are taken from [11] and [10], respectively. No correlations are assumed between H1 and ZEUS systematic uncertainties apart from a common 0.5% luminosity measurement uncertainty. The total number of data points is 1153 (552 unique points) and the number of correlated systematic sources, including normalisation uncertainties, is 43.

The averaging can take place only if most of the data from the experiments are quoted at the same Q^2 and x values. Therefore, before the averaging the data points are interpolated to a common Q^2, x grid. This interpolation is based on the H1 PDF 2000 QCD fit [11]. The interpolation of data points in principle introduces a model dependency. For H1 and ZEUS structure function data both experiments employ rather similar Q^2, x grids. About 20% of the input points are interpolated, for most of the cases the correction factors are small (few percent) and stable if different QCD fit parametrizations [4, 5] are used.

The cross section data have also been corrected to a fixed center of mass energy squared $S = 101570$ GeV^2. This has introduced a small correction for the data taken at $S = 90530$ GeV^2. The correction is based on H1-2000 PDFs, it is only significant for high inelasticity $y > 0.6$ and does not exceed 6%.

The HERA data sets agree very well: χ^2/dof for the average is $521/601$. The distribution of pulls does not show any significant tensions across the kinematic plane. Some systematic trends can be observed at low $Q^2 < 50$ GeV^2, where ZEUS NC data lie systematically higher than the H1 data, although this difference is within the normalisation uncertainty. An example of the resulting average DIS cross section is shown in Fig. 10, where the data points are displaced in Q^2 for clarity.

A remarkable side feature of the averaging is a significant reduction of the correlated systematic uncertainties. For example the uncertainty on the scattered electron energy measurement in the H1 backward calorimeter is reduced by a factor of three. The reduction of the correlated systematic uncertainties thus leads to a significant reduction of the total errors, especially for low $Q^2 < 100$ GeV^2, where systematic uncertainties limit the measurement accuracy. For this domain the total errors are often reduced by a factor two compared to the total errors of the individual H1 and ZEUS measurements.

The reduction of the correlated systematic uncertainties is achieved since the dependence of the measured cross section on the systematic sources is significantly different between H1 and ZEUS experiments. This difference is due mostly to the difference in the kinematic reconstruction methods used by the two collaborations, and to a lesser extent to the individual features of the H1 and ZEUS detectors. For example, the cross section dependence on the scattered electron energy scale has a very particular behaviour for H1 data which relies on kinematic reconstruction using only the scattered electron in one region of phase space. ZEUS uses the double angle reconstruction method where the pattern of this dependence is completely different leading to a measurement constraint.

In summary, a generalised averaging procedure to include point-to-point correlations caused by the systematic uncertainties has been developed. This averaging procedure has been applied to H1 and ZEUS DIS cross section data. The data show good consistency. The averaging of H1 and ZEUS data leads to a significant reduction of the correlated systematic uncertainties and thus a large improvement in precision for low Q^2 measurements. The goal of the averaging procedure is to obtain HERA DIS cross section set which takes into account all correlations among the experiments.

5 The longitudinal structure function F_L [7]

5.1 Introduction

At low x the sea quarks are determined by the accurate data on $F_2(x, Q^2)$. The charm contribution to F_2 is directly measured while there is no separation of up and down quarks at low x which are assumed to have the same momentum distribution, see [2]. Within this assumption, and setting the strange sea to be a fraction of the up/down sea, the proton quark content at low x is determined. The gluon distribution $xg(x, Q^2)$, however, is determined only by the derivative $\partial F_2/\partial \ln Q^2$ which is not well measured [3]. It is thus not surprising that rather different gluon distributions are obtained in global NLO analyses, as is illustrated in Figure 11. The figure displays the result of recent fits by MRST and CTEQ on the gluon distribution at low and high Q^2. It can be seen that there are striking differences at the initial scale, $Q^2 = 5\,\mathrm{GeV}^2$, which at high Q^2 get much reduced due to the evolution mechanism. The ratio of these distributions, however, exhibits differences at lower x at the level of 10% even in the LHC Higgs and W production kinematic range, see Figure 12. One also observes a striking problem at large x which is beyond the scope of this note, however. In a recent QCD analysis it was observed [3] that the dependence of the gluon distribution at low x, $xg \propto x^{b_G}$, is correlated to the value of $\alpha_s(M_Z^2)$, see Figure 13.

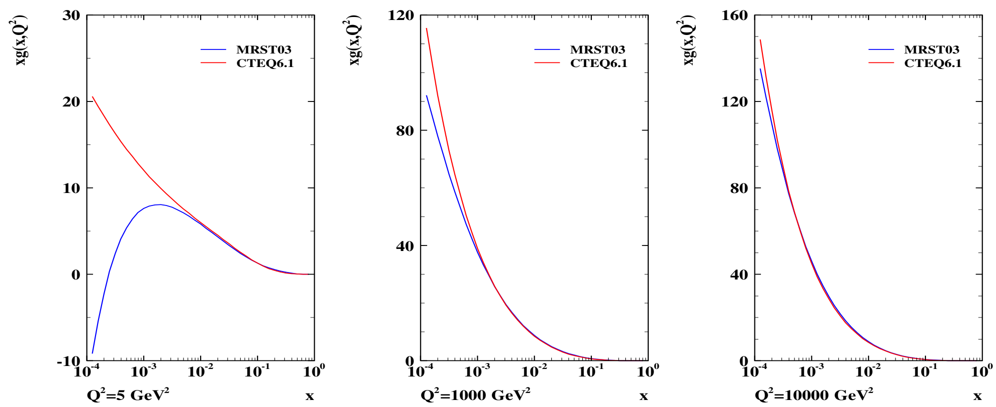


Fig. 11: Gluon momentum distributions determined by MRST and CTEQ in NLO QCD, as a function of x for $Q^2 = 5\,\mathrm{GeV}^2$, close to the initial scale of the fits, and at higher Q^2 as the result of the DGLAP evolution.

In the Quark-Parton Model the longitudinal structure function $F_L(x, Q^2)$ is zero [29]. In DGLAP

[7] Contributing authors: J. Feltesse, M. Klein

QCD, to lowest order, F_L is given by [30]

$$F_L(x,Q^2) = \frac{\alpha_s}{4\pi}x^2\int_x^1\frac{dz}{z^3}\cdot\left[\frac{16}{3}F_2(z,Q^2)+8\sum e_q^2\left(1-\frac{x}{z}\right)zg(z,Q^2)\right] \tag{7}$$

with contributions from quarks and from gluons. Approximately this equation can be solved [31] and the gluon distribution appears as a measurable quantity,

$$xg(x) = 1.8[\frac{3\pi}{2\alpha_s}F_L(0.4x) - F_2(0.8x] \simeq \frac{8.3}{\alpha_s}F_L, \tag{8}$$

determined by measurements of F_2 and F_L . Since F_L , at low x, is not much smaller than F_2 , to a good approximation F_L is a direct measure for the gluon distribution at low x.

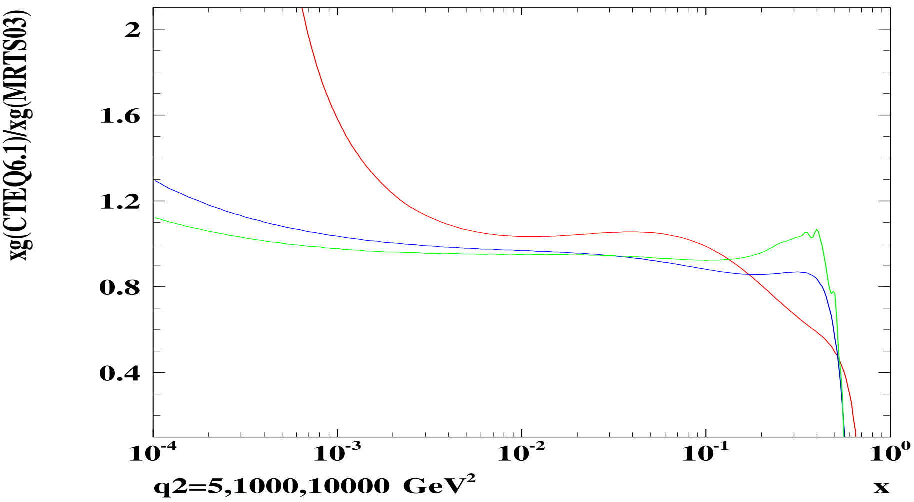


Fig. 12: Ratio of the gluon distributions of CTEQ to MRST as a function of x for low and large Q^2.

Apart from providing a very useful constraint to the determination of the gluon distribution, see also Sect. 7, a measurement of $F_L(x,Q^2)$ is of principal theoretical interest. It provides a crucial test of QCD to high orders. A significant departure of an F_L measurement from the prediction which is based on the measurement of $F_2(x,Q^2)$ and $\partial F_2/\partial \ln Q^2$ only, would require theory to be modified. There are known reasons as to why the theoretical description of gluon radiation at low x may differ from conventional DGLAP evolution: the neglect of $\ln(1/x)$, in contrast to BFKL evolution, or the importance of NLL resummation effects on the gluon splitting function (see [32]). Furthermore recent calculations of deep inelastic scattering to NNLO predict very large effects from the highest order on F_L contrary to F_2 [33].

Within the framework of the colour dipole model there exists a testable prediction for $F_L(x,Q^2)$, and the longitudinal structure function, unlike F_2 , may be subject to large higher twist effects [34].

5.2 Indirect Determinations of F_L at Low x

So far first estimates on $F_L(x,Q^2)$ at low x have been obtained by the H1 Collaboration. These result from data on the inclusive $ep \to eX$ scattering cross section

$$\frac{Q^4x}{2\pi\alpha^2Y_+}\cdot\frac{d^2\sigma}{dxdQ^2} = [F_2(x,Q^2)-f(y)\cdot F_L(x,Q^2)] = \sigma_r \tag{9}$$

obtained at fixed, large energy, $s = 4E_eE_p$. The cross section is defined by the two proton structure functions, F_2 and F_L , with $Y_+ = 1+(1-y)^2$ and $f(y) = y^2/Y_+$. At fixed s the inelasticity y is

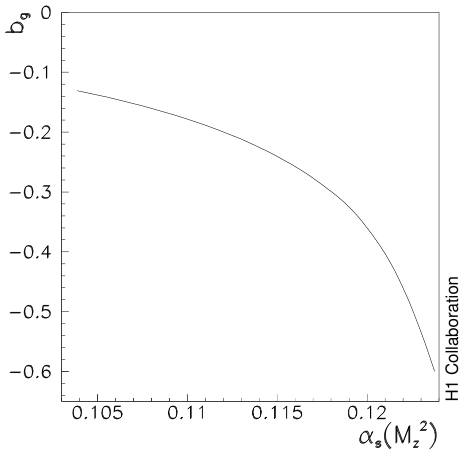


Fig. 13: Correlation of the low x behaviour of the gluon distribution, characterised by the power x^{-b_g}, with the strong coupling constant α_s as obtained in the H1 NLO QCD fit to H1 and BCDMS data.

fixed by x and Q^2 as $y = Q^2/sx$. Thus one can only measure a combination $F_2 - f(y)F_L$. Since HERA accesses a large range of y, and $f(y)$ is large only at large $y > 0.4$, assumptions have been made on F_L to extract F_2 at larger y. Since the cross section measurement accuracy has reached the few per cent level [3], the effect of the F_L assumption on F_2 at lowest x has been non-negligible. The determination of $F_2(x, Q^2)$ has thus been restricted to a region in which $y < 0.6$. The proton structure function $F_2(x, Q^2)$ is known over a few orders of magnitude in x rather well, from HERA and at largest x from fixed target data. Thus H1 did interpret the cross section at higher y as a determination of $F_L(x, Q^2)$ imposing assumptions about the behaviour of $F_2(x, Q^2)$ at lowest x. These were derived from QCD fits to the H1 data [35] or at lower Q^2, where QCD could not be trusted, from the derivative of F_2 [36]. Recently, with the established x behaviour [37] of $F_2(x, Q^2) = c(Q^2)x^{-\lambda(Q^2)}$, a new method [36] has been used to determine F_L . This "shape method" is based on the observation that the shape of σ_r, Eq. 9, at high y is driven by $f \propto y^2$ and sensitivity to F_L is restricted to a very narrow range of x corresponding to $y = 0.3 - 0.9$. Assuming that $F_L(x, Q^2)$ in this range, for each bin in Q^2, does not depend on x, one obtains a simple relation, $\sigma_r = cx^{-\lambda} - fF_L$. which has been used to determine $F_L(x, Q^2)$. Figure 14 shows the existing, preliminary data on $F_L(x, Q^2)$ at low Q^2 from the

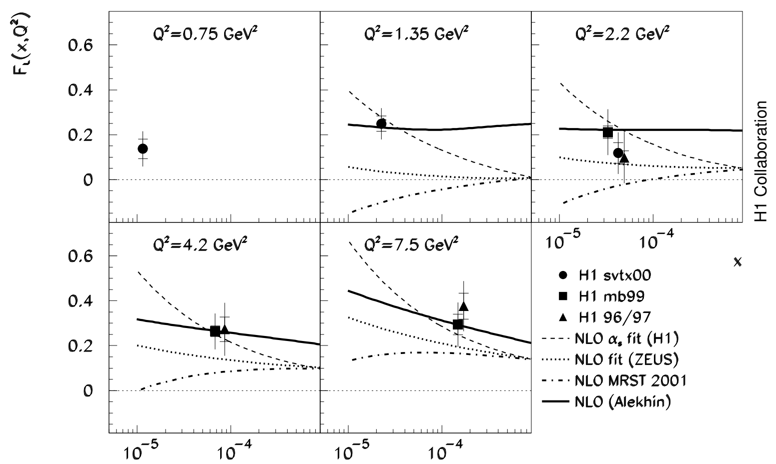


Fig. 14: Data on the longitudinal structure function obtained using assumptions on the behaviour of the other structure function F_2 in comparison with NLO QCD fit predictions. The data labeled svtx00 and mb99 data are preliminary.

H1 Collaboration in comparison with predictions from NLO DGLAP QCD fits to HERA and further cross section data. One can see that the accuracy and the x range of these $F_L(x, Q^2)$ determinations are rather limited although the data have some discriminative power already.

5.3 Backgrounds and Accuracy

The longitudinal structure function contribution to σ_r represents a small correction of the cross section in a small part of the kinematic range only. The demands for the F_L measurement are extremely high: the cross section needs to be measured at the per cent level and the scattered electron be uniquely identified up to high y. The method of unfolding F_2 and F_L consists in a measurement of σ_r at fixed x and Q^2 with varying s. This allows both structure functions to be determined from a straight line variation of σ_r as a function of $f(y)$, see [38].

At large y, corrresponding to low x, and low Q^2 the scattering kinematics at HERA resembles that of a fixed target scattering experiment: the electron scattered off quarks at very low x ("at rest") is going in the backward detector region, i.e. in the direction of the electron beam. The scattered electron is accompanied by part of the hadronic final state which is related to the struck quark. High inelasticities $y \simeq 1 - E'_e/E_e$ demand to identify scattered electrons down to a few GeV of energy E'_e. Thus a considerable background is to be isolated and removed which stems from hadrons or photons, from the $\pi_0 \rightarrow \gamma\gamma$ decay. These particles may originate both from a genuine DIS event but to a larger extent stem from photoproduction processes, in which the scattered electron escapes mostly non recognised in electron beam direction. Removal of this background in H1 is possible by requiring a track associated to the Spacal cluster, which rejects photons, and by measuring its charge which on a statistical basis removes the remaining part of the background as was demonstrated before [3, 36].

The scattered electron kinematics, E'_e and θ_e, can be accurately reconstructed using the high resolution Spacal calorimeter energy determination and the track measurements in the Backward Silicon Tracker (BST) and the Central Jet Drift Chamber (CJC). Reconstruction of the hadronic final state allows the energy momentum constraint to be imposed, using the "$E - p_z$" cut, which removes radiative corrections, and the Spacal energy scale to be calibrated at large E'_e using the double angle method. At low energies E'_e the Spacal energy scale can be calibrated to a few $\%$ using the π_0 mass constraint and be cross checked with the BST momentum measurement and with QED Compton events. The luminosity is measured to 1-2$\%$. Any common normalisation uncertainty may be removed, or further constrained, by comparing cross section data at very low y where the contribution of F_L is negligible.

Subsequently two case studies are presented which illustrate the potential of measuring F_L directly in unfolding it from the large F_2 contribution to the cross section, a study using a set of 3 low proton beam energies and a simulation for just one low Ep data set combined with standard 920 GeV data. Both studies use essentially the same correlated systematic errors and differ slightly in the assumptions on the background and efficiency uncertainties which regard the errors on cross section ratios. The following assumptions on the correlated systematics are used: $\delta E'_e/E'_e = 0.003$ at large E_e linearly rising to 0.03 at 3 GeV; $\delta\theta_e = 0.2$ mrad in the BST acceptance region and 1 mrad at larger angles; $\delta E_h/E_h = 0.02$. These and further assumed systematic uncertainties represent about the state of analysis reached so far in inclusive low Q^2 cross section measurements of H1.

5.4 Simulation Results

A simulation has been performed for $E_e = 27.6$ GeV and for four different proton beam energies, $E_p = 920,\ 575,\ 465$ and 400 GeV assuming luminosities of 10, 5, 3 and 2 pb^{-1} , respectively. The beam energies are chosen such that the cross section data are equidistant in $f(y)$. If the luminosity scales as expected as E_p^2, the low E_p luminosities are equivalent to 35 pb^{-1} at standard HERA settings. Further systematic errors regard the residual radiative corrections, assumed to be 0.5%, and the photoproduction background, 1-2% depending on y. This assumption on the background demands an improvement by a factor of about two at high y which can be expected from a high statistics subtraction of background using the charge assignment of the electron scattering candidate. An extra uncorrelated efficiency correction is assumed of 0.5%. The resulting cross section measurements are accurate to 1-2%. For each Q^2 and x point this choice provides up to four cross section measurements. The two structure functions are then obtained from a fit to $\sigma_r = F_2 + f(y)F_L$ taking into account the correlated systematics. This separation provides also accurate data of F_2, independently of F_L . The simulated data on F_L span nearly one order of magnitude in x and are shown in Figure 15. For the chosen luminosity the statistical and systematic errors on F_L are of similar size. The overall accuracy on $F_L(x, Q^2)$, which may be obtained according to the assumed experimental uncertainties, is thus estimated to be of the order of 10-20%.

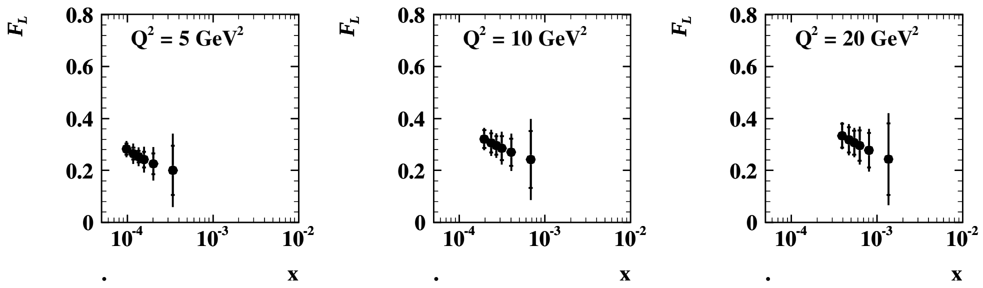


Fig. 15: Simulated measurement of the longitudinal structure function $F_L(x, Q^2)$ using the H1 backward apparatus to reconstruct the scattered electron up to maximum inelasticities of $y = 0.9$ corresponding to a mimimum electron energy of E'_e of about 3 GeV. The inner error bar is the statistical error. The full error bar denotes the statistical and systematic uncertainty added in quadrature.

Based on recent information about aspects of the machine conditions in a low proton beam energy mode, a further case study was performed [39] for only one reduced proton beam energy. In this simulation, for the standard electron beam energy of $E_e = 27.6$ GeV, proton beam energies of $E_p = 920$ and 460 GeV were chosen with luminosities of 30 and 10 pb^{-1}, respectively. According to [40] it would take about three weeks to change the configuration of the machine and to tune the luminosity plus 10 weeks to record 10 pb^{-1} of good data with High Voltage of trackers on. Uncertainties besides the correlated errors specified above are assumed for photo-production background subtraction varying from 0% at y=0.65 to 4% at y = 0.9, and of 0.5% for the residual radiative corrections. An overall uncertainty of 1% is assumed on the measurement of the cross section at low beam energy settings, which covers relative uncertainties on electron identification, trigger efficiency, vertex efficiency, and relative luminosity.

To evaluate the errors two independent methods have been considered an analytic calculation and a fast Monte-Carlo simulation technique. The two methods provide statistical and systematic errors which are in excellent agreement. The overall result of this simulation of F_L is displayed in Figure 16. In many bins the overall precision on $F_L(x, Q^2)$ is around or below 20%. It is remarkable that the overall precision would stay below 25% even if the statistical error or the larger source of systematic uncertainty would turn out to be twice larger than assumed to be in this study.

5.5 Summary

It has been demonstrated with two detailed studies that a direct measurement of the longitudinal structure function $F_L(x, Q^2)$ may be performed at HERA at the five sigma level of accuracy, in the x range from 10^{-4} to 10^{-3} in four bins of Q^2. This measurement requires about three months of running and tuning time at reduced proton beam energy. In addition it would provide the first measurement of the diffractive longitudinal structure function at the three sigma level (see the contribution of P. Newman in the summary of Working Group 4). The exact choice of the parameters of such a measurement are subject to further studies. In conclusion an accurate measurement of $F_L(x, Q^2)$ is feasible, it requires efficient detectors, dedicated beam time and analysis skills. It would be the right sign of completion to have measured

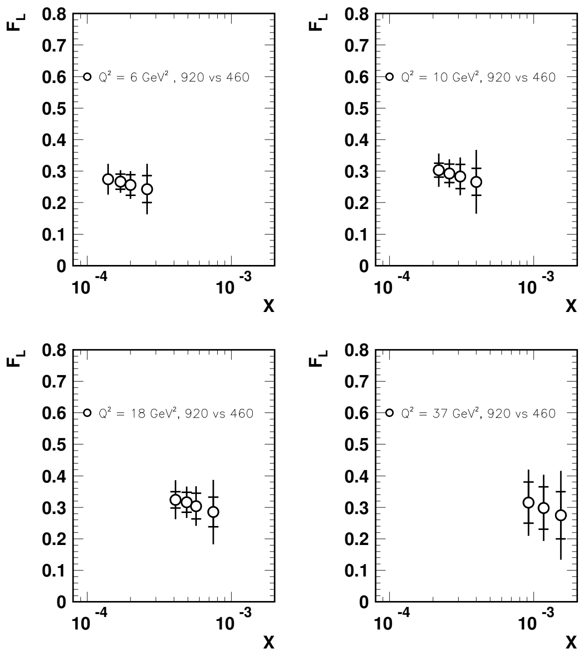


Fig. 16: Simulated measurement of the longitudinal structure function $F_L(x, Q^2)$ for data at 920 GeV (30 pb^{-1}) and 460 GeV (10 pb^{-1}). The inner error bar is the statistical error. The full error bar denotes the statistical and systematic uncertainty added in quadrature.

F_2 first, in 1992 and onwards, and to conclude the HERA data taking with a dedicated measurement of the second important structure function $F_L(x, Q^2)$, which is related to the gluon density in the low x range of the LHC.

6 Determination of the Light Quark Momentum Distributions at Low x at HERA [8]

Based on the data taken in the first phase of HERA's operation (1993-2000), the HERA collider experiments have measured a complete set of neutral (NC) and charged (CC) current double differential $e^{\pm}p$ inclusive scattering cross sections, based on about 120 pb^{-1} of positron-proton and 15 pb^{-1} of electron-proton data. The NC and CC deep inelastic scattering (DIS) cross sections for unpolarised $e^{\pm}p$ scattering are determined by structure functions and quark momentum distributions in the proton as follows:

$$\sigma^{\pm}_{NC} \sim Y_+ F_2 \mp Y_- xF_3, \tag{10}$$

$$F_2 \simeq e_u^2 x(U + \overline{U}) + e_d^2 x(D + \overline{D}), \tag{11}$$

$$xF_3 \simeq 2x[a_u e_u (U - \overline{U}) + a_d e_d (D - \overline{D})], \tag{12}$$

$$\sigma^{+}_{CC} \sim x\overline{U} + (1-y)^2 xD, \tag{13}$$

$$\sigma^{-}_{CC} \sim xU + (1-y)^2 x\overline{D}. \tag{14}$$

Here $y = Q^2/sx$ is the inelasticity, $s = 4E_eE_p$ and $Y_{\pm} = 1 \pm (1-y)^2$. The parton distribution $U = u + c + b$ is the sum of the momentum distributions of the up-type quarks with charge $e_u = 2/3$ and axial vector coupling $a_u = 1/2$, while $D = d + s$ is the sum of the momentum distributions of the down type quarks with charge $e_d = -1/3$, $a_d = -1/2$. Similar relationships hold for the anti-quark distributions $\overline{U}$ and $\overline{D}$.

As is illustrated in Fig. 17 the H1 experiment [11] has determined all four quark distributions and the gluon distribution xg. The accuracy achieved so far by H1, for $x = 0.01, 0.4$ and 0.65, is

[8]Contributing authors: M. Klein, B. Reisert

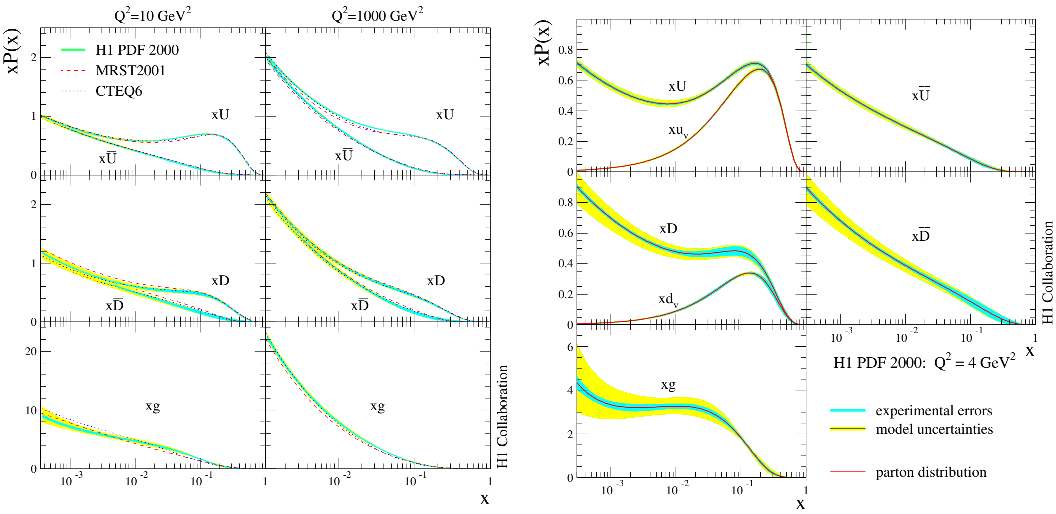


Fig. 17: Determination of the sum of up, anti-up, down and anti-downquark distributions and of the gluon distribution in the proton based on the H1 neutral and charged current cross section data. Left: for Q^2 of 10 and 1000 GeV^2 compared with results from MRST and CTEQ; Right: the parton distributions with their experimental and model uncertainties as determined by H1 at the starting scale $Q_0^2 = 4\,\mathrm{GeV}^2$.

$1\%,\ 3\%, 7\%$ for the sum of up quark distributions and $2\%,\ 10\%, 30\%$ for the sum of down quark distributions, respectively. The extracted parton distributions are in reasonable agreement with the results obtained in global fits by the MRST [5] and CTEQ [4] collaborations. The H1 result is also consistent with the pdfs determined by the ZEUS Collaboration [10] which uses jet data to improve the accuracy for the gluon distribution and imports a $\overline{d} - \overline{u}$ asymmetry fit result from MRST. New data which are being taken (HERA II) will improve the accuracy of these determinations further. At the time this is written, the available data per experiment have been grown to roughly 150 pb^{-1} for both e^+p and e^-p scattering, and more is still to come. These data will be particularly important to improve the accuracy at large x, which at HERA is related to high Q^2.

As is clear from the above equations, the NC and CC cross section data are sensitive directly to only these four quark distribution combinations. Disentangling the individual quark flavours (up, down, strange, charm and beauty) requires additional assumptions. While informations on the c and b densities are being obtained from measurements of $F_2^{c\overline{c}}$ and $F_2^{b\overline{b}}$ of improving accuracy, the determination of the strange quark density at HERA is less straightforward and may rest on $sW^+ \rightarrow c$ and strange (Φ) particle production [41]. The relative contributions from the heavy quarks become increasingly important with Q^2, as is illustrated in Fig. 18.

The larger x domain is dominated by the valence quarks. At HERA the valence quark distributions are not directly determined but extracted from the differences $u_v = U - \overline{U}$ and $d_v = D - \overline{D}$. Note that this implies the assumption that sea and anti-quarks are equal which in non-perturbative QCD models may not hold. A perhaps more striking assumption is inherent in these fits and regards the sea quark asymmetries at low x which is the main subject of the subsequent discussion.

Fig. 19 shows the difference $x\overline{d} - x\overline{u}$ as determined in the H1 PDF 2000 fit based on the H1 data alone (left) and using in addition the BCDMS proton and deuteron data (right). One observes a trend of these fits to reproduce the asymmetry near $x \sim 0.1$ which in the MRST and CTEQ fits, shown in Fig. 19, is due to fitting the Drell Yan data from the E866/NuSea experiment [42]. While this enhancement is not very stable in the H1 fit [43] and not significant either, with the BCDMS data an asymmetry is observed which reflects the violation of the Gottfried sum rule.

In the H1 fit [11] the parton distributions at the initial scale $Q^2 = 4\,\mathrm{GeV}^2$ are parameterised as

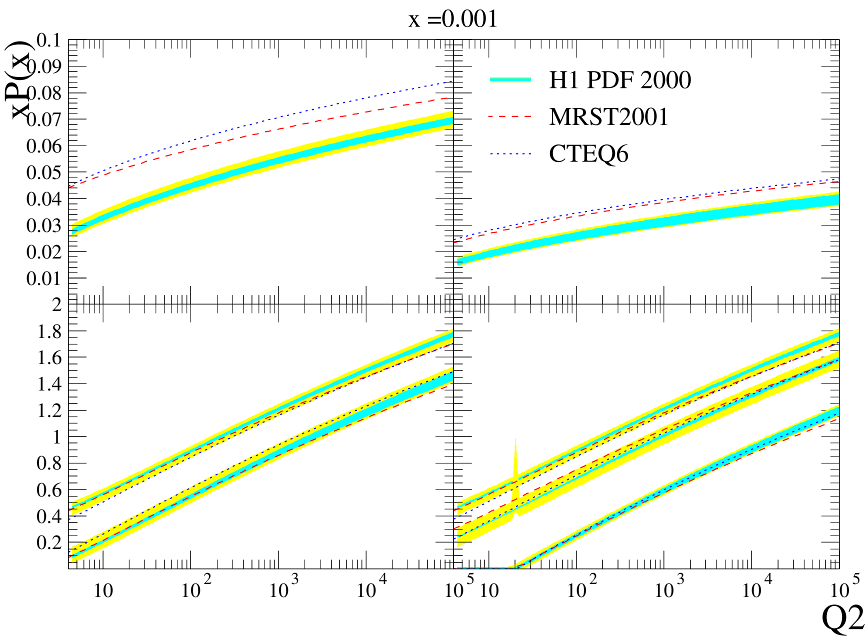


Fig. 18: Parton distributions and their uncertainties as determined by H1 extrapolated to the region of the LHC, for $x = 0.001$ near to the rapidity plateau. Top left: u valence; top right: d valence; bottom left: $\overline{u}$ and below c; bottom right, in decreasing order: $\overline{d}$, s, b. The results are compared with recent fits to global data by MRST and CTEQ. Note that at such small x the valence quark distributions are very small. With increasing Q^2 the relative importance of the heavy quarks compared to the light quarks increases while the absolute difference of the quark distributions is observed to be rather independent of Q^2. The beauty contribution to the cross section thus becomes sizeable, amounting to about 5% for $pp \rightarrow HW$.

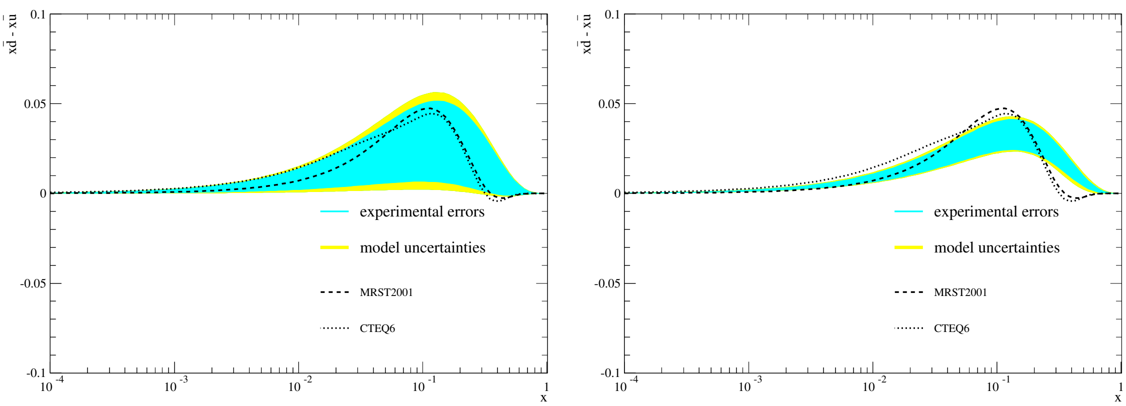


Fig. 19: Determination of the difference $x(\overline{d} - \overline{u})$ in the H1 PDF 2000 fit to the H1 data (left) and the H1 and the BCDMS μp and μD data (right). The sea quark difference is enforced to tend to zero at low x. The global fit results of MRST and CTEQ include Drell Yan data which suggest a sea quark asymmetry at $x \sim 0.1$.

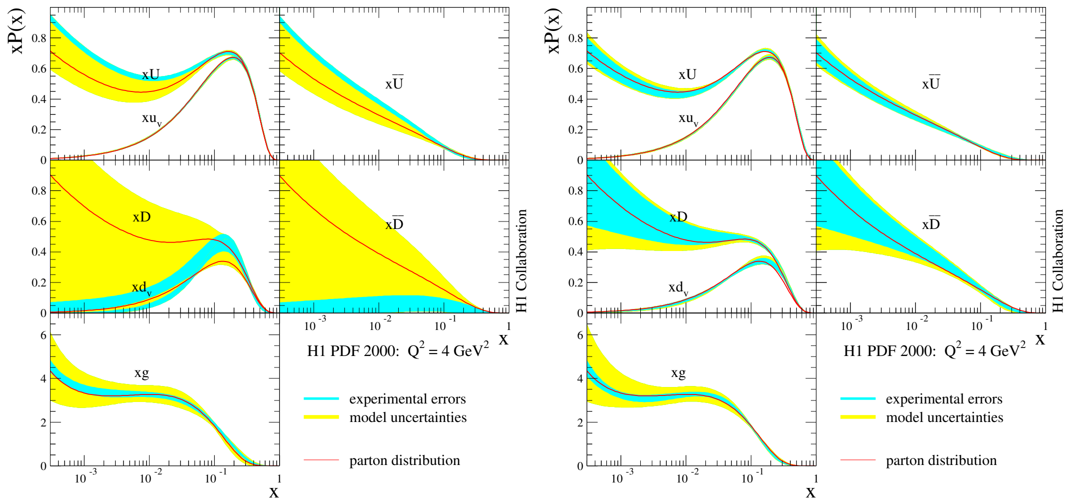


Fig. 20: Determinations of the quark and gluon momentum distributions releasing the constraint $x\overline{d} = x\overline{u}$ at low x, from the H1 NC and CC data alone (left) and from the H1 ep and the BCDMS μp and μD data (right). Since at low $x < 0.01$ there is no further constraint than that given from F_2 the uncertainties of $\overline{U}$ and in particular of $\overline{D}$ become sizeable.

$xP = A_p x^{B_P}(1-x)^{C_P} \cdot f_P(x)$. The function f_P is a polynomial in x which is determined by requiring "χ^2 saturation" of the fits, i.e. starting from $f_P = 1$ additional terms $D_P x$, $E_P x^2$ etc. are added and only considered if they cause a significant improvement in χ^2, half integer powers were considered in [43]. The result for fitting the H1 data has been as follows: $f_g = (1 + D_g x)$, $f_U = (1 + D_U x + F_U x^3)$, $f_D = (1 + D_D x)$ and $f_{\overline{U}} = f_{\overline{D}} = 1$. The parton distributions at low x are thus parameterised as $xP \rightarrow A_P x^{B_P}$. The strange (charm) anti-quark distribution is coupled to the total amount of down (up) anti-quarks as $\overline{s} = f_c\overline{D}$ ($\overline{c} = f_c\overline{U}$). Two assumptions have been made on the behaviour of the quark and anti-quark distributions at low x. It has been assumed that quark and anti-quark distributions are equal and, moreover, that the sea is flavour symmetric. This implies that the slopes B of all four quark distributions are set equal $B_U = B_D = B_{\overline{U}} = B_{\overline{D}}$. Moreover, the nomalisations of up and down quarks are the same, i.e. $A_{\overline{U}}(1 - f_c) = A_{\overline{D}}(1 - f_s)$, which ensures that $\overline{d}/\overline{u} \rightarrow 1$ as x tends to zero. The consequence of this assumption is illustrated in Fig. 19. While the DIS data suggest some asymmetry at larger x, the up-down quark asymmetry is enforced to vanish at lower x. This results in a rather fake high accuracy in the determination of the four quark distributions at low x, despite the fact that at low x there is only one combination of them measured, which is $F_2 = x[4(U + \overline{U}) + (D + \overline{D})]/9$. If one relaxes both the conditions on the slopes and normalisations, the fit to the H1 data decides to completely remove the down quark contributions as is seen in Fig. 20 (left plot).

In DIS the up and down quark asymmetry can be constrained using deuteron data because the nucleon structure function determines a different linear combination according to $F_2^N = 5x(U + \overline{U} + D + \overline{D})/18 + x(c + \overline{c} - s - \overline{s})/6$ with $N = (p+n)/2$. Unfortunately, there are only data at rather large x available. The effect of including the BCDMS data on the low x behaviour of the parton distributions is illustrated in Fig. 20 (right plot). It restores some amount of down quarks at low x , the errors, however, in particular of the down quarks, are still very large. The result is a large sea quark asymmetry uncertainty, which is shown in Fig. 21. At HERA a proposal had been made [44] to operate the machine in electron-deuteron mode. Measuring the behaviour at low x would not require high luminosity. Such data would constrain [9] a possible sea quark asymmetry with very high accuracy, as is also shown in Fig. 21.

[9]Constraints on the sea quark distributions may also be obtained from W^+/W^- production at the TeVatron. However, the sensitivity is limited to larger $x \geq 0.1$ [45] since $W's$ produced in collisions involving sea quarks of smaller x will be boosted so strongly, that their decay products are not within the acceptance of the collider detectors. W^+ and W^- production at the LHC has been discussed in [46].

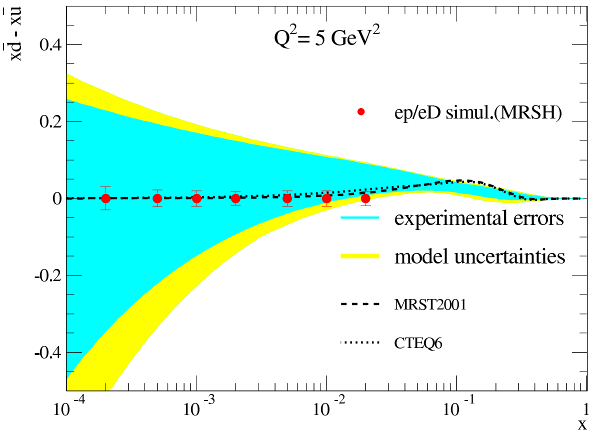


Fig. 21: Simulation of the difference of sea quark distributions, here assumed to be zero, at low x based on additional 20 pb^{-1} of electron-deuteron data at HERA. The error band represents the uncertainty of the H1 NLO QCD fit to the H1 ep and the BCDMS μp and μd data without the constraint $\overline{d} = \overline{u}$ at low x. The dashed curves represent calculations using recent global fits by MRST and by CTEQ.

Deuterons at HERA would require a new source and modest modifications to the preaccelerators. The H1 apparatus could be used in its standard mode with a forward proton detector added to take data at half the beam energy. Tagging the spectator protons with high accuracy at HERA, for the first time in DIS, one could reconstruct the electron-neutron scattering kinematics essentially free of nuclear corrections [44]. Since the forward scattering amplitude is related to diffraction one would also be able to constrain shadowing to the per cent level [47]. The low x measurements would require small luminosity amounts, of less than 50 pb^{-1}. Long awaited constraints of the d/u ratio at large x and Q^2 would require extended running, as would CC data. Besides determining the parton distributions better, the measurement of the singlet F_2^N structure function would give important constraints on the evolution and theory at low x [48]. It would also result in an estimated factor of two improvement on the measurement of α_s at HERA [49]. For the development of QCD, of low x physics in particular, but as well for understanding physics at the LHC and also for superhigh energy neutrino astrophysics, HERA eD data remain to be important.

7 Impact of future HERA data on the determination of proton PDFs using the ZEUS NLO QCD fit [10]

7.1 PDF fits to HERA data

Recently, the ZEUS Collaboration have performed a combined NLO QCD fit to inclusive neutral and charged current DIS data [23, 25, 27, 24, 26, 28] as well as high precision jet data in DIS [50] and γp scattering [51]. This is called the ZEUS-JETS PDF fit [10]. The use of only HERA data eliminates the uncertainties from heavy-target corrections and removes the need for isospin symmetry assumptions. It also avoids the difficulties that can sometimes arise from combining data-sets from several different experiments, thereby allowing a rigorous statistical treatment of the PDF uncertainties. Furthermore, PDF uncertainties from current global fits are, in general, limited by (irreducible) experimental systematics. In contrast, those from fits to HERA data alone, are largely limited by the statistical precision of existing measurements. Therefore, the impact of future data from HERA is likely to be most significant in fits to only HERA data.

[10]Contributing authors: C. Gwenlan, A. Cooper-Sarkar, C. Targett-Adams.

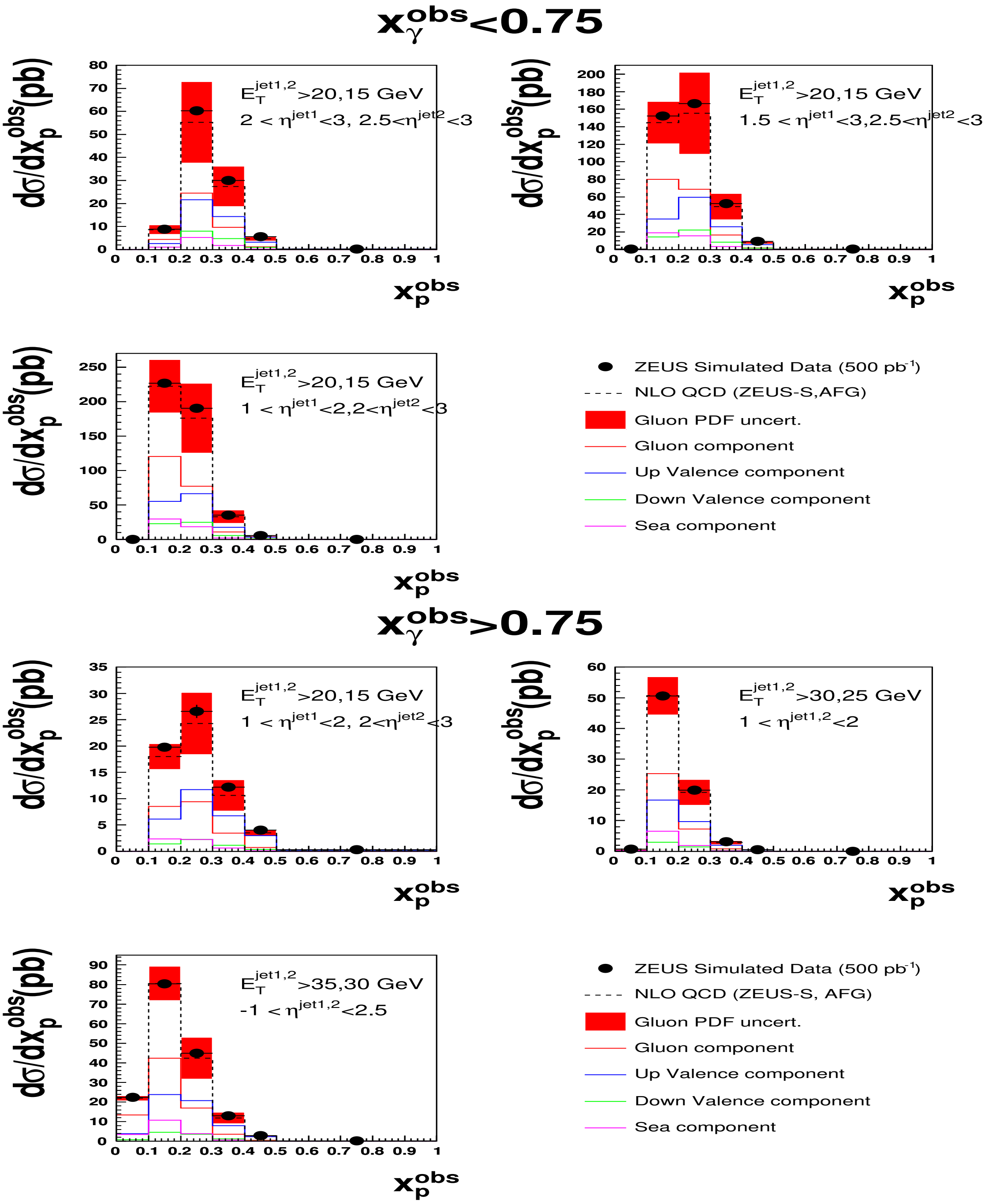


Fig. 22: The optimised jet cross sections included in the HERA-II projected fit. The solid points show the simulated data generated using the NLO QCD programme of Frixione-Ridolfi, using the CTEQ5M1 proton and the AFG photon PDFs. The error bars show the statistical uncertainties, which correspond to 500 pb^{-1} of HERA data. Systematic uncertainties have been neglected. The dashed line shows the NLO QCD prediction using the ZEUS-S proton and AFG photon PDFs. The shaded band shows the contribution to the cross section uncertainty arising from the uncertainty in the gluon distribution in the proton.

data sample	kinematic coverage	**HERA-I** $\mathcal{L}$ (pb^{-1})	**HERA-II** $\mathcal{L}$ (pb^{-1}) **(assumed)**
96-97 NC e^+p [23]	$2.7 < Q^2 < 30000$ GeV^2; $6.3 \cdot 10^{-5} < x < 0.65$	30	30
94-97 CC e^+p [24]	$280 < Q^2 < 17000$ GeV^2; $6.3 \cdot 10^{-5} < x < 0.65$	48	48
98-99 NC e^-p [25]	$200 < Q^2 < 30000$ GeV^2; $0.005 < x < 0.65$	16	350
98-99 CC e^-p [26]	$280 < Q^2 < 17000$ GeV^2; $0.015 < x < 0.42$	16	350
99-00 NC e^+p [27]	$200 < Q^2 < 30000$ GeV^2; $0.005 < x < 0.65$	63	350
99-00 CC e^+p [28]	$280 < Q^2 < 17000$ GeV^2; $0.008 < x < 0.42$	61	350
96-97 inc. DIS jets [50]	$125 < Q^2 < 30000$ GeV^2; $E_{\mathrm{T}}^{Breit} > 8$ GeV	37	500
96-97 dijets in γp [51]	$Q^2 \lesssim 1$ GeV^2; $E_{\mathrm{T}}^{jet1,2} > 14, 11$ GeV	37	500
optimised jets [52]	$Q^2 \lesssim 1$ GeV^2; $E_{\mathrm{T}}^{jet1,2} > 20, 15$ GeV	-	500

Table 4: The data-sets included in the ZEUS-JETS and HERA-II projected PDF fits. The first column lists the type of data and the second gives the kinematic coverage. The third column gives the integrated luminosities of the HERA-I measurements included in the ZEUS-JETS fit. The fourth column gives the luminosities assumed in the HERA-II projection. Note that the 96-97 NC and the 94-97 CC measurements have not had their luminosity scaled for the HERA-II projection.

7.2 The ZEUS NLO QCD fit

The ZEUS-JETS PDF fit has been used as the basis for all results shown in this contribution. The most important details of the fit are summarised here. A full description may be found elsewhere [10]. The fit includes the full set of ZEUS inclusive neutral and charged current $e^{\pm}p$ data from HERA-I (1994-2000), as well as two sets of high precision jet data in e^+p DIS ($Q^2 >> 1$ GeV^2) and γp ($Q^2 \sim 0$) scattering. The inclusive data used in the fit, span the kinematic range $6.3 \times 10^{-5} < x < 0.65$ and $2.7 < Q^2 < 30000$ GeV^2.

The PDFs are obtained by solving the NLO DGLAP equations within the $\overline{\mathrm{MS}}$ scheme. These equations yield the PDFs at all values of Q^2 provided they are input as functions of x at some starting scale Q_0^2. The resulting PDFs are convoluted with coefficient functions to give predictions for structure functions and, hence, cross sections. In the ZEUS fit, the $xu_v(x)$ (u-valence), $xd_v(x)$ (d-valence), $xS(x)$ (total sea-quark), $xg(x)$ (gluon) and $x(\bar{d}(x)-\bar{u}(x))$ PDFs are parameterised at a starting scale of $Q_0^2 = 7$ GeV^2 by the form,

$$xf(x) = p_1 x^{p_2}(1-x)^{p_3} P(x), \tag{15}$$

where $P(x) = (1 + p_4 x)$. No advantage in the χ^2 results from using more complex polynomial forms. The normalisation parameters, $p_1(u_v)$ and $p_1(d_v)$, are constrained by quark number sum rules while $p_1(g)$ is constrained by the momentum sum rule. Since there is no information to constrain any difference in the low-x behaviour of the u- and d-valence quarks, $p_2(u_v)$ has been set equal to $p_2(d_v)$. The data from HERA are currently less precise than the fixed target data in the high-x regime. Therefore, the high-x sea and gluon distributions are not well constrained in current fits to HERA data alone. To account for this, the sea shape has been restricted by setting $p_4(S) = 0$. The high-x gluon shape is constrained by the inclusion of HERA jet data. In fits to only HERA data, there is no information on the shape of $\bar{d}-\bar{u}$. Therefore, this distribution has its shape fixed consistent with Drell-Yan data and its normalisation set consistent with the size of the Gottfried sum rule violation. A suppression of the strange sea with respect to the non-strange sea of a factor of 2 at Q_0^2 is also imposed, consistent with neutrino induced dimuon data from CCFR. The value of the strong coupling has been fixed to $\alpha_s(M_Z) = 0.1180$. After all constraints, the ZEUS-JETS fit has 11 free parameters. Heavy quarks were treated in the variable flavour number scheme of Thorne & Roberts [19]. Full account was taken of correlated experimental systematic uncertainties, using the Offset Method [9, 18].

The results of two separate studies are presented. The first study provides an estimate of how

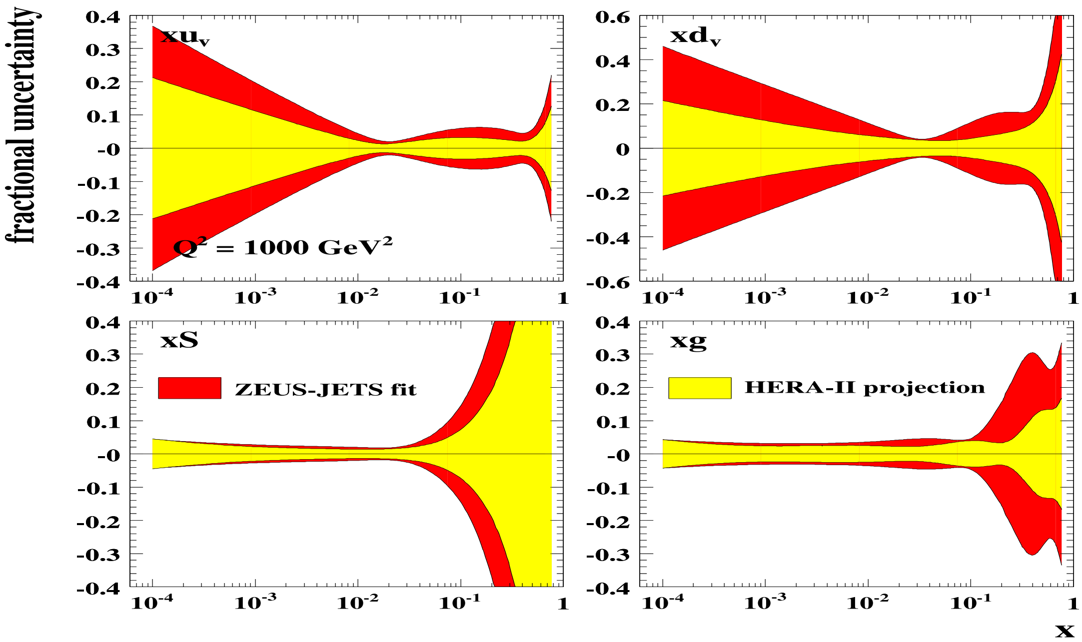


Fig. 23: The fractional PDF uncertainties, as a function of x, for the u-valence, d-valence, sea-quark and gluon distributions at $Q^2 = 1000$ GeV^2. The red shaded bands show the results of the ZEUS-JETS fit and the yellow shaded bands show the results of the HERA-II projected fit.

well the PDF uncertainties may be known by the end of HERA-II, within the currently planned running scenario, while the second study investigates the impact of a future HERA measurement of F_L on the gluon distribution. All results presented, are based on the recent ZEUS-JETS PDF analysis [10].

7.3 PDF uncertainty estimates for the end of HERA running

The data from HERA-I are already very precise and cover a wide kinematic region. However, HERA-II is now running efficiently and is expected to provide a substantial increase in luminosity. Current estimates suggest that, by the end of HERA running (in mid-2007), an integrated luminosity of 700 pb^{-1} should be achievable. This will allow more precise measurements of cross sections that are curently statistically limited: in particular, the high-Q^2 NC and CC data, as well as high-Q^2 and/or high-E_T jet data. In addition to the simple increase in luminosity, recent studies [52] have shown that future jet cross section measurements, in kinematic regions optimised for sensitivity to PDFs, should have a significant impact on the gluon uncertainties. In this contribution, the effect on the PDF uncertainties, of both the higher precision expected from HERA-II and the possibility of optimised jet cross section measurements, has been estimated in a new QCD fit. This fit will be referred to as the “HERA-II projection”.

In the HERA-II projected fit, the statistical uncertainties on the currently available HERA-I data have been reduced. For the high-Q^2 inclusive data, a total integrated luminosity of 700 pb^{-1} was assumed, equally divided between e^+ and e^-. For the jet data, an integrated luminosity of 500 pb^{-1} was assumed. The central values and systematic uncertainties were taken from the published data in each case. In addition to the assumed increase in precision of the measurements, a set of optimised jet cross sections were also included, for forward dijets in γp collisions, as defined in a recent study [52]. Since no real data are yet available, simulated points were generated using the NLO QCD program of Frixione-Ridolfi [53], using the CTEQ5M1 [4] proton and AFG [54] photon PDFs. The statistical uncertainties were taken to correspond to 500 pb^{-1}. For this study, systematic uncertainties on the optimised jet cross sections were ignored. The simulated optimised jet cross section points, compared to the predictions of NLO QCD using the ZEUS-S proton PDF [9], are shown in Fig. 22.

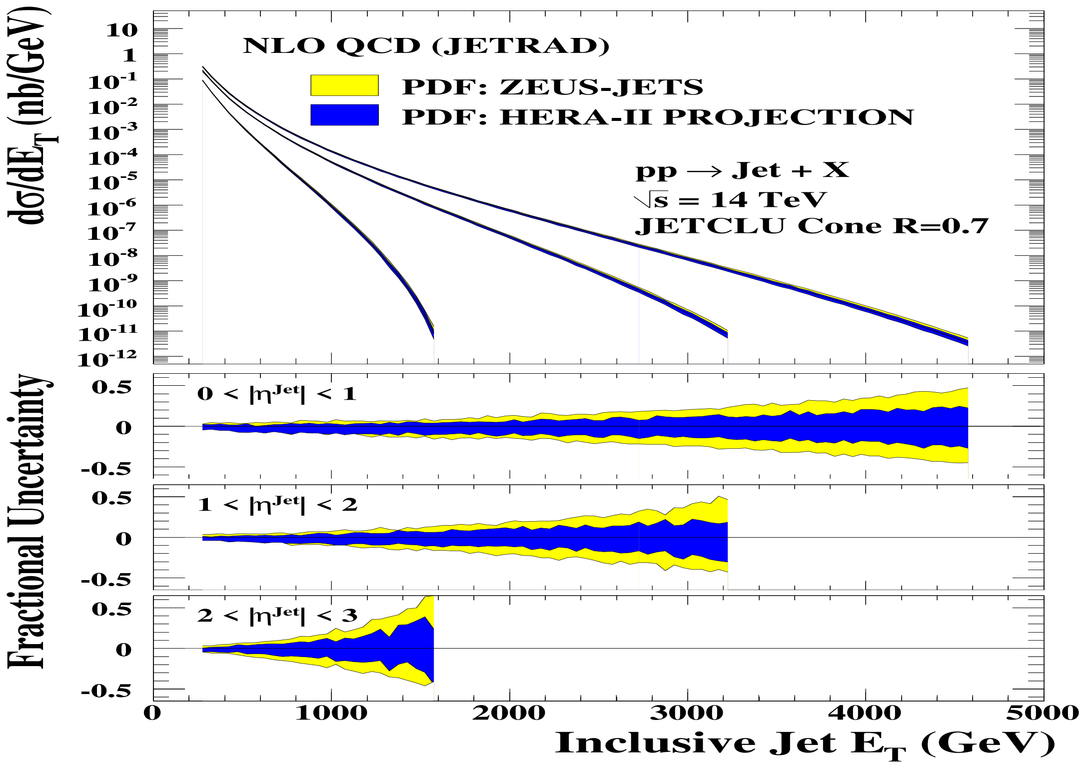


Fig. 24: NLO QCD inclusive jet cross section predictions at $\sqrt{s}$=14 TeV in three regions of pseudo-rapidity. The yellow and blue bands show the PDF uncertainties from the ZEUS-JETS and HERA-II projected fits, respectively.

Table 4 lists the data-sets included in the ZEUS-JETS and HERA-II projected fits. The luminosities of the (real) HERA-I measurements and those assumed for the HERA-II projection are also given.

The results are summarised in Fig. 23, which shows the fractional PDF uncertainties, for the u- and d-valence, sea-quark and gluon distributions, at $Q^2 = 1000$ GeV^2. The yellow bands show the results of the ZEUS-JETS fit while the red bands show those for the HERA-II projection. Note that the same general features are observed for all values of Q^2. In fits to only HERA data, the information on the valence quarks comes from the high-Q^2 NC and CC cross sections. The increased statistical precision of the high-Q^2 data, as assumed in the HERA-II projected fit, gives a significant improvement in the valence uncertainties over the whole range of x. For the sea quarks, a significant improvement in the uncertainties at high-x is also observed. In contrast, the low-x uncertainties are not visibly reduced. This is due to the fact that the data constraining the low-x region tends to be at lower-Q^2, which are already systematically limited. This is also the reason why the low-x gluon uncertainties are not significantly reduced. However, the mid-to-high-x gluon, which is constrained by the jet data, is much improved in the HERA-II projected fit. Note that about half of the observed reduction in the gluon uncertainties is due to the inclusion of the simulated optimised jet cross sections.

Inclusive jet cross sections at the LHC The improvement to the high-x partons, observed in the HERA-II projection compared to the ZEUS-JETS fit, will be particularly relevant for high-scale physics at the LHC. This is illustrated in Fig. 24, which shows NLO QCD predictions from the JETRAD [55] programme for inclusive jet production at $\sqrt{s} = 14$ TeV. The results are shown for both the ZEUS-JETS and the HERA-II projected PDFs. The uncertainties on the cross sections, resulting from the PDFs, have been calculated using the LHAPDF interface [56]. For the ZEUS-JETS PDF, the uncertainty reaches $\sim 50\%$ at central pseudo-rapidities, for the highest jet transverse energies shown. The prediction using the HERA-II projected PDF shows a marked improvement at high jet tranverse energy.

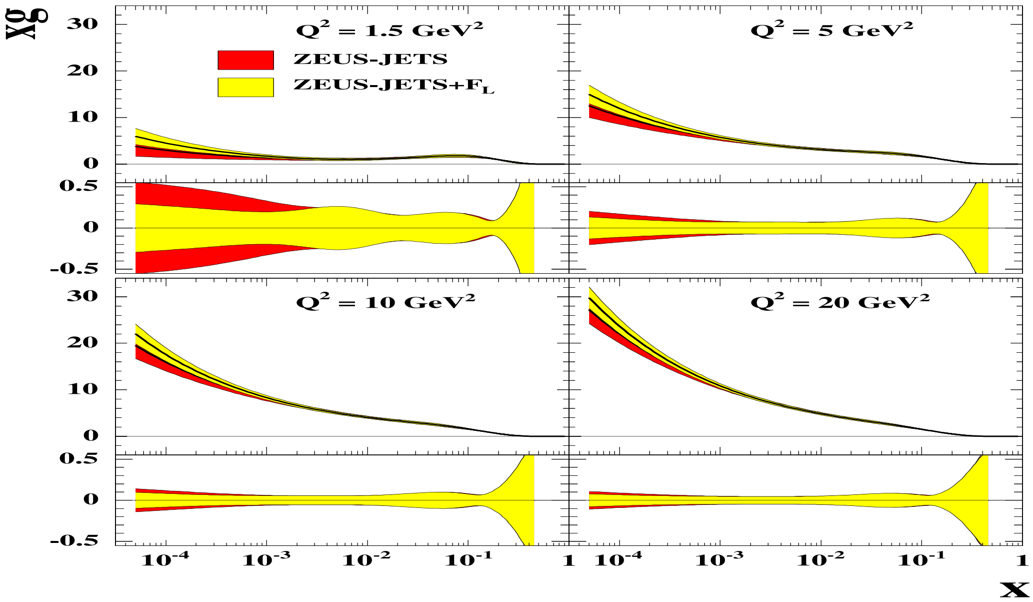


Fig. 25: The gluon PDFs, showing also the fractional uncertainty, for fits with and without inclusion of the simulated F_L data, for $Q^2 =$ 1.5, 5, 10 and 20 GeV^2. The red shaded bands show the results of the ZEUS-JETS fit and the yellow shaded band show the results of the ZEUS-JETS+F_L fit.

7.4 Impact of a future HERA measurement of F_L on the gluon PDF

The longitudinal structure function, F_L, is directly related to the gluon density in the proton. In principle, F_L can be extracted by measuring the NC DIS cross section at fixed x and Q^2, for different values of y (see Eqn. 3). A precision measurement could be achieved by varying the centre-of-mass energy, since $s = Q^2/xy \approx 4E_eE_p$, where E_e and E_p are the electron and proton beam energies, respectively. Studies [38] (Sec. 5) have shown that this would be most efficiently achieved by changing the proton beam energy. However, such a measurement has not yet been performed at HERA.

There are several reasons why a measurement of F_L at low-x could be important. The gluon density is not well known at low-x and so different PDF parameterisations can give quite different predictions for F_L at low-x. Therefore, a precise measurement of the longitudinal sturcture function could both pin down the gluon PDF and reduce its uncertainties. Furthermore, predictions of F_L also depend upon the nature of the underlying theory (e.g. order in QCD, resummed calculation etc). Therefore, a measurement of F_L could also help to discriminate between different theoretical models.

Impact on the gluon PDF uncertainties The impact of a possible future HERA measurement of F_L on the gluon PDF uncertainties has been investigated, using a set of simulated F_L data-points [38]. (see Sec. 5). The simulation was performed using the GRV94 [57] proton PDF for the central values, and assuming $E_e = 27.6$ GeV and $E_p = 920, 575, 465$ and 400 GeV, with luminosities of 10, 5, 3 and 2 pb^{-1}, respectively. Assuming that the luminosity scales simply as E_p^2, this scenario would nominally cost $35\ \mathrm{pb}^{-1}$ of luminosity under standard HERA conditions. However, this estimate takes no account of time taken for optimisation of the machine with each change in E_p, which could be considerable. The systematic uncertainties on the simulated data-points were calculated assuming a $\sim 2\%$ precision on the inclusive NC cross section measurement. A more comprehensive description of the simulated data is given in contribution for this proceedings, see Sec. 5.

The simulated data were included in the ZEUS-JETS fit. Figure 25 shows the gluon distribution and fractional uncertainties for fits with and without inclusion of the simulated F_L data. The results

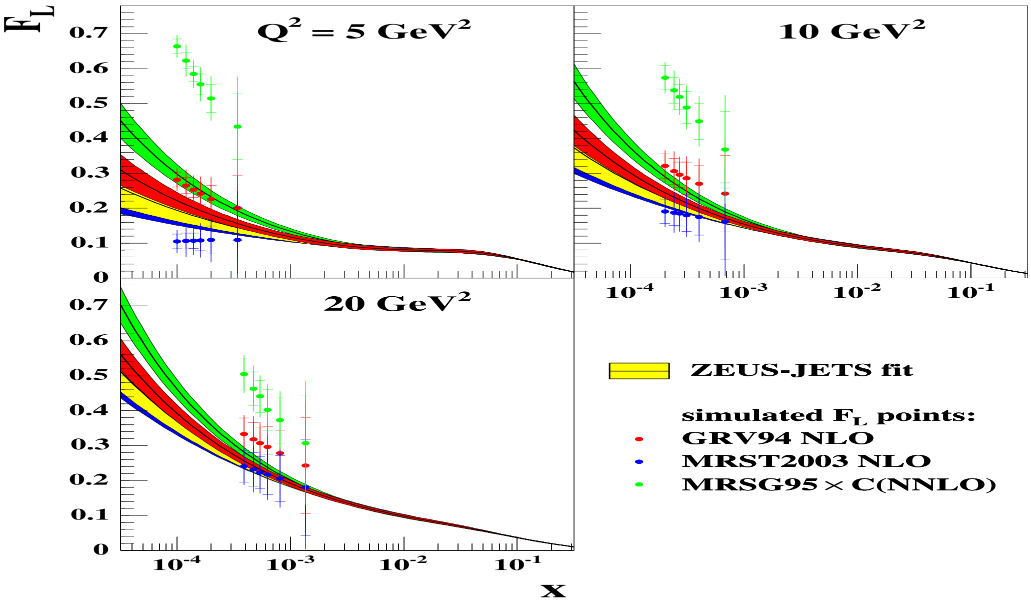


Fig. 26: The distribution of the longitudinal structure function F_L at Q^2 =5, 10 and 20 GeV^2. The blue, red and green points show the simulated F_L data-points, respectively labelled maximum, middle and minimum in Table 5. The blue, red and green shaded bands show the NLO QCD predictions, in the case where the data-points of the corresponding colour have been included in the fit. For comparison, the yellow shaded band shows the prediction of the ZEUS-JETS fit.

indicate that the gluon uncertainties are reduced at low-x, but the improvement is only significant at relatively low $Q^2 \lesssim 20$ GeV^2.

Discrimination between theoretical models In order to assess whether a HERA measurement of F_L could discriminate between theoretical models, two more sets of F_L data-points have been simulated [58], using different theoretical assumptions. The first of the two sets was generated using the MRSG95 [59] proton PDF, which has a large gluon density. The PDFs were then convoluted with the NNLO order coefficient functions, which are large and positive. This gives the “maximum” set of F_L data-points. In contrast, the second set has been generated using the MRST2003 [60] proton PDF, which has a negative gluon at low-x and low-Q^2, thus providing a “minimum” set of F_L data. The original set of F_L points described in the previous subsection lies between these two extremes. The details of all three sets are summarised in Tab. 5.

Figure 26 shows the results of including, individually, each set of simulated F_L data into the ZEUS NLO QCD fit. The results show that the NLO fit is relatively stable to the inclusion of the extreme sets of data. This indicates that a measurement of F_L could discriminate between certain theoretical models. However, it should be noted that the maximum and minimum models studied here were chosen specifically to give the widest possible variation in F_L. There are many other alternatives that would lie between these extremes and the ability of an F_L measurement to discriminate between them would depend both on the experimental precision of the measurement itself, as well as the theoretical uncertainties on the models being tested.

	PDF	QCD order of coefficient functions
Maximum F_L	MRSG95	NNLO
Middle F_L	GRV94	NLO
Minimum F_L	MRST2003	NLO

Table 5: Summary of the PDFs used to generate the simulated F_L data-points. The extreme maximum F_L points were generated using the MRSG95 PDF, and convoluted with NNLO coefficient functions. The middle points were generated using the GRV94 PDF, and the extreme minimum points were generated using the MRST2003 PDF, which has a negative gluon at low-x.

8 A Method to Include Final State Cross-sections Measured in Proton-Proton Collisions to Global NLO QCD Analysis [11]

The Large Hadron Collider (LHC), currently under construction at CERN, will collide protons on protons with an energy of 7 TeV. Together with its high collision rate the high available centre-of-mass energy will make it possible to test new interactions at very short distances that might be revealed in the production cross-sections of Standard Model (SM) particles at very high transverse momentum (P_T) as deviation from the SM theory.

The sensitivity to new physics crucially depends on experimental uncertainties in the measurements and on theoretical uncertainties in the SM predictions. It is therefore important to work out a strategy to minimize both the experimental and theoretical uncertainties from LHC data. For instance, one could use single inclusive jet or Drell-Yan cross-sections at low P_T to constrain the PDF uncertainties at high P_T. Typical residual renormalisation and factorisation scale uncertainties in next-to-leading order (NLO) calculations for single inclusive jet-cross-section are about $5 - 10\%$ and should hopefully be reduced as NNLO calculations become available. The impact of PDF uncertainties on the other hand can be substantially larger in some regions, especially at large P_T, and for example at $P_T = 2000$ GeV dominate the overall uncertainty of 20%. If a suitable combination of data measured at the Tevatron and LHC can be included in global NLO QCD analyses, the PDF uncertainties can be constrained.

The aim of this contribution is to propose a method for consistently including final-state observables in global QCD analyses.

For inclusive data like the proton structure function F_2 in deep-inelastic scattering (DIS) the perturbative coefficients are known analytically. During the fit the cross-section can therefore be quickly calculated from the strong coupling (α_s) and the PDFs and can be compared to the measurements. However, final state observables, where detector acceptances or jet algorithms are involved in the definition of the perturbative coefficients (called “weights” in the following), have to be calculated using NLO Monte Carlo programs. Typically such programs need about one day of CPU time to calculate accurately the cross-section. It is therefore necessary to find a way to calculate the perturbative coefficients with high precision in a long run and to include α_s and the PDFs “a posteriori”.

To solve this problem many methods have been proposed in the past [61, 62, 63, 3, 64, 10]. In principle the highest efficiencies can be obtained by taking moments with respect to Bjorken-x [61, 62], because this converts convolutions into multiplications. This can have notable advantages with respect to memory consumption, especially in cases with two incoming hadrons. On the other hand, there are complications such as the need for PDFs in moment space and the associated inverse Mellin transforms.

Methods in x-space have traditionally been somewhat less efficient, both in terms of speed (in the ‘a posteriori’ steps — not a major issue here) and in terms of memory consumption. They are, however, somewhat more transparent since they provide direct information on the x values of relevance. Furthermore they can be used with any PDF. The use of x-space methods can be further improved by using methods developed originally for PDF evolution [65, 66].

[11]Contributing authors: T. Carli, G. Salam, F. Siegert.

8.1 PDF-independent representation of cross-sections

Representing the PDF on a grid We make the assumption that PDFs can be accurately represented by storing their values on a two-dimensional grid of points and using $n^{\rm th}$-order interpolations between those points. Instead of using the parton momentum fraction x and the factorisation scale Q^2, we use a variable transformation that provides good coverage of the full x and Q^2 range with uniformly spaced grid points:[12]

$$y(x) = \ln \frac{1}{x} \quad \text{and} \quad \tau(Q^2) = \ln \ln \frac{Q^2}{\Lambda^2}. \tag{16}$$

The parameter Λ is to be chosen of the order of $\Lambda_{\rm QCD}$, but not necessarily identical. The PDF $q(x, Q^2)$ is then represented by its values q_{i_y,i_τ} at the 2-dimensional grid point $(i_y\, \delta y, i_\tau\, \delta\tau)$, where δy and $\delta\tau$ denote the grid spacings, and obtained elsewhere by interpolation:

$$q(x, Q^2) = \sum_{i=0}^{n} \sum_{\iota=0}^{n'} q_{k+i,\kappa+\iota}\, I_i^{(n)}\left(\frac{y(x)}{\delta y} - k\right) I_\iota^{(n')}\left(\frac{\tau(Q^2)}{\delta\tau} - \kappa\right), \tag{17}$$

where n, n' are the interpolation orders. The interpolation function $I_i^{(n)}(u)$ is 1 for $u = i$ and otherwise is given by:

$$I_i^{(n)}(u) = \frac{(-1)^{n-i}}{i!(n-i)!} \frac{u(u-1)\dots(u-n)}{u-i}. \tag{18}$$

Defining $\mathrm{int}(u)$ to be the largest integer such that $\mathrm{int}(u) \le u$, k and κ are defined as:

$$k(x) = \ \mathrm{int}\left(\tfrac{y(x)}{\delta y} - \tfrac{n-1}{2}\right), \quad \kappa(x) = \mathrm{int}\left(\frac{\tau(Q^2)}{\delta\tau} - \frac{n'-1}{2}\right). \tag{19}$$

Given finite grids whose vertex indices range from $0 \dots N_y - 1$ for the y grid and $0 \dots N_\tau - 1$ for the τ grid, one should additionally require that eq. (17) only uses available grid points. This can be achieved by remapping $k \to \max(0, \min(N_y - 1 - n, k))$ and $\kappa \to \max(0, \min(N_\tau - 1 - n', \kappa))$.

Representing the final state cross-section weights on a grid (DIS case) Suppose that we have an NLO Monte Carlo program that produces events $m = 1 \dots N$. Each event m has an x value, x_m, a Q^2 value, Q^2_m, as well as a weight, w_m, and a corresponding order in $\alpha_{\rm s}$, p_m. Normally one would obtain the final result W of the Monte Carlo integration from:[13]

$$W = \sum_{m=1}^{N} w_m \left(\frac{\alpha_s(Q^2_m)}{2\pi}\right)^{p_m} q(x_m, Q^2_m). \tag{20}$$

Instead one introduces a weight grid $W^{(p)}_{i_y,i_\tau}$ and then for each event updates a portion of the grid with:

$i = 0 \dots n,\ \iota = 0 \dots n' :$

$$W^{(p_m)}_{k+i,\kappa+\iota} \to W^{(p_m)}_{k+i,\kappa+\iota} + w_m\, I_i^{(n)}\left(\frac{y(x_m)}{\delta y} - k\right) I_\iota^{(n')}\left(\frac{\tau(Q^2_m)}{\delta\tau} - \kappa\right), \tag{21}$$

$$\text{where} \quad k \equiv k(x_m),\ \kappa \equiv \kappa(Q^2_m).$$

The final result for W, for an arbitrary PDF, can then be obtained *subsequent* to the Monte Carlo run:

$$W = \sum_p \sum_{i_y} \sum_{i_\tau} W^{(p)}_{i_y,i_\tau} \left(\frac{\alpha_s\left(Q^{2(i_\tau)}\right)}{2\pi}\right)^p q\left(x^{(i_y)}, Q^{2(i_\tau)}\right), \tag{22}$$

[12] An alternative for the x grid is to use $y = \ln 1/x + a(1-x)$ with a a parameter that serves to increase the density of points in the large x region.

[13] Here, and in the following, renormalisation and factorisation scales have been set equal for simplicity.

where the sums index with i_y and i_τ run over the number of grid points and we have have explicitly introduced $x^{(i_y)}$ and $Q^{2(i_\tau)}$ such that:

$$y(x^{(i_y)}) = i_y\,\delta y \quad \text{and} \quad \tau\left(Q^{2(i_\tau)}\right) = i_\tau\,\delta\tau. \tag{23}$$

Including renormalisation and factorisation scale dependence If one has the weight matrix $W^{(p)}_{i_y,i_\tau}$ determined separately order by order in $\alpha_{\rm s}$, it is straightforward to vary the renormalisation μ_R and factorisation μ_F scales a posteriori (we assume that they were kept equal in the original calculation).

It is helpful to introduce some notation relating to the DGLAP evolution equation:

$$\frac{dq(x,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}(P_0\otimes q)(x,Q^2) + \left(\frac{\alpha_s(Q^2)}{2\pi}\right)^2 (P_1\otimes q)(x,Q^2) + \ldots, \tag{24}$$

where the P_0 and P_1 are the LO and NLO matrices of DGLAP splitting functions that operate on vectors (in flavour space) q of PDFs. Let us now restrict our attention to the NLO case where we have just two values of p, $p_{\rm LO}$ and $p_{\rm NLO}$. Introducing ξ_R and ξ_F corresponding to the factors by which one varies μ_R and μ_F respectively, for arbitrary ξ_R and ξ_F we may then write:

$$\begin{aligned} W(\xi_R,\xi_F) = \sum_{i_y}\sum_{i_\tau} \left(\frac{\alpha_s\left(\xi_R^2 Q^{2(i_\tau)}\right)}{2\pi}\right)^{p_{\rm LO}} W^{(p_{\rm LO})}_{i_y,i_\tau}\, q\Big(x^{(i_y)},\xi_F^2 Q^{2(i_\tau)}\Big) + \\ \left(\frac{\alpha_s\left(\xi_R^2 Q^{2(i_\tau)}\right)}{2\pi}\right)^{p_{\rm NLO}} \Big[\Big(W^{(p_{\rm NLO})}_{i_y,i_\tau} + 2\pi\beta_0 p_{\rm LO}\ln\xi_R^2\, W^{(p_{\rm LO})}_{i_y,i_\tau}\Big)\, q\Big(x^{(i_y)},\xi_F^2 Q^{2(i_\tau)}\Big) \\ - \ln\xi_F^2\, W^{(p_{\rm LO})}_{i_y,i_\tau}(P_0\otimes q)\Big(x^{(i_y)},\xi_F^2 Q^{2(i_\tau)}\Big)\Big]\,, \end{aligned} \tag{25}$$

where $\beta_0 = (11N_c - 2n_f)/(12\pi)$ and N_c (n_f) is the number of colours (flavours). Though this formula is given for x-space based approach, a similar formula applies for moment-space approaches. Furthermore it is straightforward to extend it to higher perturbative orders.

Representing the weights in the case of two incoming hadrons In hadron-hadron scattering one can use analogous procedures with one more dimension. Besides Q^2, the weight grid depends on the momentum fraction of the first (x_1) and second (x_2) hadron.

In the case of jet production in proton-proton collisions the weights generated by the Monte Carlo program as well as the PDFs can be organised in seven possible initial state combinations of partons:

$$\begin{aligned}
\mathrm{gg}: \quad F^{(0)}(x_1,x_2;Q^2) &= G_1(x_1)G_2(x_2) && (26)\\
\mathrm{qg}: \quad F^{(1)}(x_1,x_2;Q^2) &= \left(Q_1(x_1)+\overline{Q}_1(x_1)\right)G_2(x_2) && (27)\\
\mathrm{gq}: \quad F^{(2)}(x_1,x_2;Q^2) &= G_1(x_1)\left(Q_2(x_2)+\overline{Q}_2(x_2)\right) && (28)\\
\mathrm{qr}: \quad F^{(3)}(x_1,x_2;Q^2) &= Q_1(x_1)Q_2(x_2)+\overline{Q}_1(x_1)\overline{Q}_2(x_2) - D(x_1,x_2) && (29)\\
\mathrm{qq}: \quad F^{(4)}(x_1,x_2;Q^2) &= D(x_1,x_2) && (30)\\
\mathrm{q\bar{q}}: \quad F^{(5)}(x_1,x_2;Q^2) &= \overline{D}(x_1,x_2) && (31)\\
\mathrm{q\bar{r}}: \quad F^{(6)}(x_1,x_2;Q^2) &= Q_1(x_1)\overline{Q}_2(x_2)+\overline{Q}_1(x_1)Q_2(x_2) - \overline{D}(x_1,x_2), && (32)
\end{aligned}$$

where g denotes gluons, q quarks and r quarks of different flavour $q \neq r$ and we have used the generalized

PDFs defined as:

$$G_H(x) = f_{0/H}(x, Q^2), \qquad Q_H(x) = \sum_{i=1}^{6} f_{i/H}(x, Q^2), \;\; \overline{Q}_H(x) = \sum_{i=-6}^{-1} f_{i/H}(x, Q^2),$$

$$\begin{aligned} D(x_1, x_2) &= \sum_{\substack{i=-6\\ i\neq 0}}^{6} f_{i/H_1}(x_1, Q^2) f_{i/H_2}(x_2, Q^2), \\ \overline{D}(x_1, x_2, \mu_F^2) &= \sum_{\substack{i=-6\\ i\neq 0}}^{6} f_{i/H_1}(x_1, Q^2) f_{-i/H_2}(x_2, Q^2), \end{aligned} \tag{33}$$

where $f_{i/H}$ is the PDF of flavour $i = -6 \ldots 6$ for hadron H and H_1 (H_2) denotes the first or second hadron[14].

The analogue of eq. 22 is then given by:

$$W = \sum_p \sum_{l=0}^{6} \sum_{i_{y_1}} \sum_{i_{y_2}} \sum_{i_\tau} W^{(p)(l)}_{i_{y_1}, i_{y_2}, i_\tau} \left(\frac{\alpha_s \left(Q^{2(i_\tau)} \right)}{2\pi} \right)^p F^{(l)} \left(x_1^{(i_{y_1})}, x_2^{(i_{y_1})}, Q^{2(i_\tau)} \right). \tag{34}$$

Including scale depedence in the case of two incoming hadrons It is again possible to choose arbitrary renormalisation and factorisation scales, specifically for NLO accuracy:

$$\begin{aligned} W(\xi_R, \xi_F) = &\sum_{l=0}^{6} \sum_{i_{y_1}} \sum_{i_{y_2}} \sum_{i_\tau} \left(\frac{\alpha_s \left(\xi_R^2 Q^{2(i_\tau)} \right)}{2\pi} \right)^{p_{\mathrm{LO}}} W^{(p_{\mathrm{LO}})(l)}_{i_{y_1}, i_{y_2}, i_\tau} F^{(l)} \left(x_1^{(i_{y_1})}, x_2^{(i_{y_1})}, \xi_F^2 Q^{2(i_\tau)} \right) + \\ &\left(\frac{\alpha_s \left(\xi_R^2 Q^{2(i_\tau)} \right)}{2\pi} \right)^{p_{\mathrm{NLO}}} \left[\left(W^{(p_{\mathrm{NLO}})(l)}_{i_{y_1}, i_{y_2}, i_\tau} + 2\pi\beta_0 p_{\mathrm{LO}} \ln \xi_R^2 \, W^{(p_{\mathrm{LO}})(l)}_{i_{y_1}, i_{y_2}, i_\tau} \right) F^{(l)} \left(x_1^{(i_{y_1})}, x_2^{(i_{y_1})}, \xi_F^2 Q^{2(i_\tau)} \right) \right. \\ &\left. - \ln \xi_F^2 \, W^{(p_{\mathrm{LO}})(l)}_{i_{y_1}, i_{y_2}, i_\tau} \left(F^{(l)}_{q1 \to P_0 \otimes q1} \left(x_1^{(i_{y_1})}, x_2^{(i_{y_1})}, \xi_F^2 Q^{2(i_\tau)} \right) + F^{(l)}_{q2 \to P_0 \otimes q2} \left(x_1^{(i_{y_1})}, x_2^{(i_{y_1})}, \xi_F^2 Q^{2(i_\tau)} \right) \right) \right], \end{aligned} \tag{35}$$

where $F^{(l)}_{q1 \to P_0 \otimes q1}$ is calculated as $F^{(l)}$, but with q_1 replaced wtih $P_0 \otimes q_1$, and analogously for $F^{(l)}_{q2 \to P_0 \otimes q2}$.

8.2 Technical implementation

To test the scheme discussed above we use the NLO Monte Carlo program NLOJET++ [67] and the CTEQ6 PDFs [4]. The grid $W^{(p)(l)}_{i_{y_1}, i_{y_2}, i_\tau}$ of eq. 34 is filled in a NLOJET++ user module. This module has access to the event weight and parton momenta and it is here that one specifies and calculates the physical observables that are being studied (e.g. jet algorithm).

Having filled the grid we construct the cross-section in a small standalone program which reads the weights from the grid and multiplies them with an arbitrary α_{s} and PDF according to eq. 34. This program runs very fast (in the order of seconds) and can be called in a PDF fit.

The connection between these two programs is accomplished via a C++ class, which provides methods e.g. for creating and optimising the grid, filling weight events and saving it to disk. The classes are general enough to be extendable for the use with other NLO calculations.

The complete code for the NLOJET++ module, the C++ class and the standalone job is available from the authors. It is still in a development, testing and tuning stage, but help and more ideas are welcome.

[14] In the above equation we follow the standard PDG Monte Carlo numbering scheme [17] where gluons are denoted as 0, quarks have values from 1-6 and anti-quarks have the corresponding negative values.

The C++ class The main data members of this class are the grids implemented as arrays of three-dimensional ROOT histograms, with each grid point at the bin centers[15]:

$$\mathrm{TH3D[p][l][iobs](x_1, x_2, Q^2)}, \tag{36}$$

where the l and p are explained in eq. 34 and $iobs$ denotes the observable bin, e.g. a given P_T range[16].

The C++ class initialises, stores and fills the grid using the following main methods:

- *Default constructor:* Given the pre-defined kinematic regions of interest, it initializes the grid.
- *Optimizing method:* Since in some bins the weights will be zero over a large kinematic region in x_1, x_2, Q^2, the optimising method implements an automated procedure to adapt the grid boundaries for each observable bin. These boundaries are calculated in a first (short) run. In the present implementation, the optimised grid has a fixed number of grid points. Other choices, like a fixed grid spacing, might be implemented in the future.
- *Loading method:* Reads the saved weight grid from a ROOT file
- *Saving method:* Saves the complete grid to a ROOT file, which will be automatically compressed.

The user module for NLOJET++ The user module has to be adapted specifically to the exact definition of the cross-section calculation. If a grid file already exists in the directory where NLOJET++ is started, the grid is not started with the default constructor, but with the optimizing method (see 8.2). In this way the grid boundaries are optimised for each observable bin. This is necessary to get very fine grid spacings without exceeding the computer memory. The grid is filled at the same place where the standard NLOJET++ histograms are filled. After a certain number of events, the grid is saved in a root-file and the calculation is continued.

The standalone program for constructing the cross-section The standalone program calculates the cross-section in the following way:

1. Load the weight grid from the ROOT file
2. Initialize the PDF interface[17], load $q(x, Q^2)$ on a helper PDF-grid (to increase the performance)
3. For each observable bin, loop over $i_{y_1}, i_{y_2}, i_\tau, l, p$ and calculate $F^l(x_1, x_2, Q^2)$ from the appropriate PDFs $q(x, Q^2)$, multiply α_s and the weights from the grid and sum over the initial state parton configuration l, according to eq. 34.

8.3 Results

We calculate the single inclusive jet cross-section as a function of the jet transverse momentum (P_T) for jets within a rapidity of $|y| < 0.5$. To define the jets we use the seedless cone jet algorithm as implemented in NLOJET++ using the four-vector recombination scheme and the midpoint algorithm. The cone radius has been put to $R = 0.7$, the overlap fraction was set to $f = 0.5$. We set the renormalisation and factorization scale to $Q^2 = P_{T,max}^2$, where $P_{T,max}$ is the P_T of the highest P_T jet in the required rapidity region[18].

[15]ROOT histograms are easy to implement, to represent and to manipulate. They are therefore ideal in an early development phase. An additional advantage is the automatic file compression to save space. The overhead of storing some empty bins is largely reduced by optimizing the x_1, x_2 and Q^2 grid boundaries using the NLOJET++ program before final filling. To avoid this residual overhead and to exploit certain symmetries in the grid, a special data class (e.g. a sparse matrix) might be constructed in the future.

[16]For the moment we construct a grid for each initial state parton configuration. It will be easy to merge the qg and the gq initial state parton configurations in one grid. In addition, the weights for some of the initial state parton configurations are symmetric in x_1 and x_2. This could be exploited in future applications to further reduce the grid size.

[17]We use the C++ wrapper of the LHAPDF interface [56].

[18]Note that beyond LO the $P_{T,max}$ will in general differ from the P_T of the other jets, so when binning an inclusive jet cross section, the P_T of a given jet may not correspond to the renormalisation scale chosen for the event as a whole. For this reason we shall need separate grid dimensions for the jet P_T and for the renormalisation scale. Only in certain moment-space approaches [62] has this requirement so far been efficiently circumvented.

In our test runs, to be independent from statistical fluctuations (which can be large in particular in the NLO case), we fill in addition to the grid a reference histogram in the standard way according to eq. 20.

The choice of the grid architecture depends on the required accuracy, on the exact cross-section definition and on the available computer resources. Here, we will just sketch the influence of the grid architecture and the interpolation method on the final result. We will investigate an example where we calculate the inclusive jet cross-section in $N_{\rm obs} = 100$ bins in the kinematic range $100 \leq P_T \leq 5000\,\mathrm{GeV}$. In future applications this can serve as guideline for a user to adapt the grid method to his/her specific problem. We believe that the code is transparent and flexible enough to adapt to many applications.

As reference for comparisons of different grid architectures and interpolation methods we use the following:

- *Grid spacing in $y(x)$:* $10^{-5} \leq x_1, x_2 \leq 1.0$ with $N_y = 30$
- *Grid spacing in $\tau(Q^2)$:* $100\,\mathrm{GeV} \leq Q \leq 5000\,\mathrm{GeV}$ with $N_\tau = 30$
- *Order of interpolation:* $n_y = 3$, $n_\tau = 3$

The grid boundaries correspond to the user setting for the first run which determines the grid boundaries for each observable bin. In the following we call this grid architecture 30^2x30x100$(3,3)$. Such a grid takes about 300 Mbyte of computer memory. The root-file where the grid is stored has about 50 Mbyte.

The result is shown in Fig. 27a). The reference cross-section is reproduced everywhere to within 0.05%. The typical precision is about 0.01%. At low and high P_T there is a positive bias of about 0.04%. Also shown in Fig. 27a) are the results obtained with different grid architectures. For a finer x grid (50^2x30x100$(3,3)$) the accuracy is further improved (within 0.005%) and there is no bias. A finer (30^2x60x100$(3,3)$) as well as a coarser (30^2x10x100$(3,3)$) binning in Q^2 does not improve the precision.

Fig. 27b) and Fig. 27c) show for the grid (30^2x30x100) different interpolation methods. With an interpolation of order $n = 5$ the precision is 0.01% and the bias at low and high P_T observed for the $n = 3$ interpolation disappears. The result is similar to the one obtained with finer x-points. Thus by increasing the interpolation order the grid can be kept smaller. An order $n = 1$ interpolation gives a systematic negative bias of about 1% becoming even larger towards high P_T.

Depending on the available computer resources and the specific problem, the user will have to choose a proper grid architecture. In this context, it is interesting that a very small grid 10^2x10x100$(5,5)$ that takes only about 10 Mbyte computer memory reaches still a precision of 0.5%, if an interpolation of order $n = 5$ is used (see Fig. 27d)).

We have developed a technique to store the perturbative coefficients calculated by an NLO Monte Carlo program on a grid allowing for a-posteriori inclusion of an arbitrary parton density function (PDF) set. We extended a technique that was already successfully used to analyse HERA data to the more demanding case of proton-proton collisions at LHC energies.

The technique can be used to constrain PDF uncertainties, e.g. at high momentum transfers, from data that will be measured at LHC and allows the consistent inclusion of final state observables in global QCD analyses. This will help increase the sensitivity of LHC to find new physics as deviations from the Standard Model predictions.

Even for the large kinematic range for the parton momentum fractions x_1 and x_2 and of the squared momentum transfer Q^2 accessible at LHC, grids of moderate size seem to be sufficient. The single inclusive jet cross-section in the central region $|y| < 0.5$ can be calculated with a precision of 0.01% in a realistic example with 100 bins in the transverse jet energy range $100 \leq P_T \leq 5000\,\mathrm{GeV}$. In this example, the grid occupies about 300 Mbyte computer memory. With smaller grids of order 10 Mbyte the reachable accuracy is still 0.5%. This is probably sufficient for all practical applications.

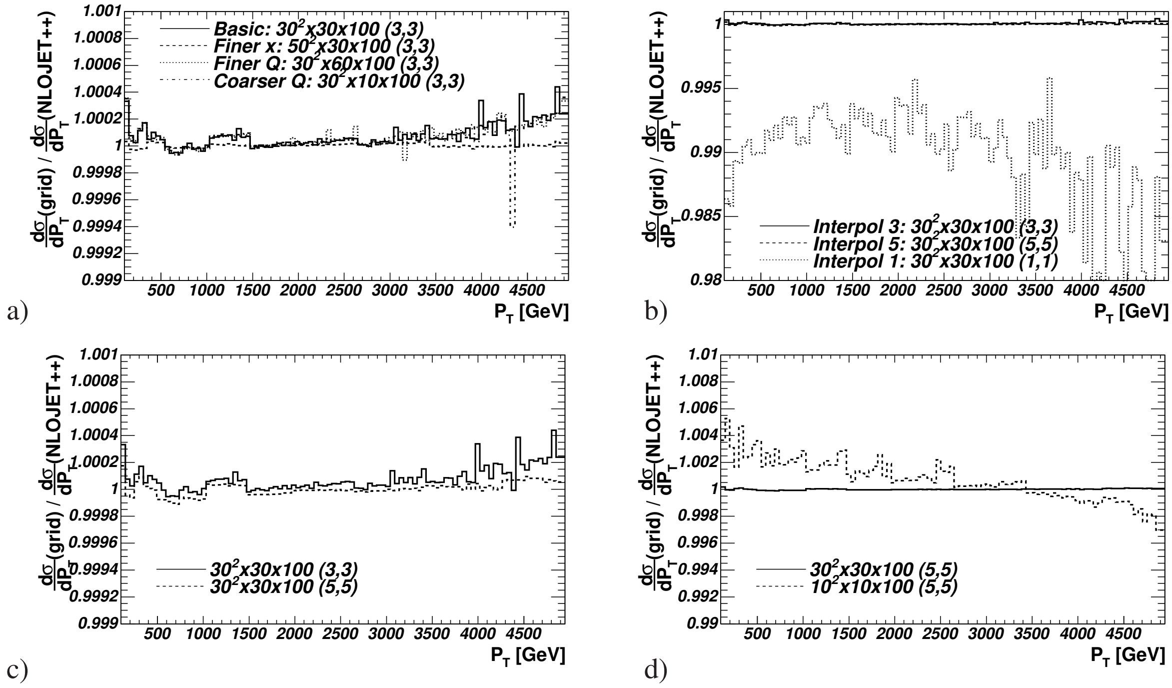


Fig. 27: Ratio between the single inclusive jet cross-section with 100 P_T bins calculated with the grid technique and the reference cross-section calculated in the standard way. Shown are the standard grid, grids with finer x and Q^2 sampling (a) with interpolation of order 1, 3 and 5 (b) (and on a finer scale in c)) and a small grid (d).

References

[1] G. Laštovička-Medin et al., *Precision limits for hera dis cross section measurement*. These proceedings.

[2] M. Klein, B. Reisert, *Determination of the light quark momentum distributions at low x at hera*. These proceedings.

[3] Adloff, C. et al., Eur. Phys. J. **C21**, 33 (2001).

[4] Pumplin, J. et al., JHEP **07**, 012 (2002).

[5] Martin, A. D. et al., Eur. Phys. J **C23**, 73 (2002).

[6] Alekhin, S., Phys. Rev. **D68**, 014002 (2003).

[7] Bassler, U. Thesis, Université Pierre et Marie Curie, Paris, 2003.

[8] Laštovička, T., Eur. Phys. J. **C24**, 529 (2002);
Laštovička, T., Acta Phys. Polon. **B33**, 2867 (2002).

[9] ZEUS Coll., Chekanov, S. et al., Phys. Rev. **D67**, 012007 (2003).

[10] ZEUS Coll., Chekanov, S. et al., Eur.Phys.J **C42**, 1 (2005).

[11] Adloff, C. et al., Eur. Phys. J. **C30**, 1 (2003).

[12] Gribov, V.N. and Lipatov, L.N., Sov.J.Nucl.Phys **15**, 438 (1972).

[13] Altarelli, G. and Parisi, G., Nucl.Phys. **B126**, 298 (1977).

[14] Lipatov, L. N., Sov.J.Nucl.Phys **20**, 94 (1975).

[15] Dokshitzer, Yu. L., JETP **46**, 641 (1977).

[16] Devenish, R. C. E. and Cooper-Sarkar, A. M., *Deep Inelastic Scattering*. Oxford Unviersity Press, Oxford, 2004.

[17] Eidelman, S. et al., Phys. Lett. **B592**, 1 (2004).

[18] Cooper-Sarkar, A. M., J. Phys. **G28**, 2669 (2002).

[19] Thorne, R.S. and Roberts, R.G., Phys. Rev **D57**, 6871 (1998).

[20] Pascaud, C. and Zomer, F., *Qcd analysis from the proton structure function f2 measurement: Issues on fitting, statistical and systematic errors*. LAL-95-05.

[21] Adloff, C. et al., Eur. Phys. J. **C13**, 609 (2000).

[22] Adloff, C. et al., Eur. Phys. J. **C19**, 269 (2001).

[23] Chekanov, S. et al., Eur. Phys. J. **C21**, 443 (2001).

[24] Breitweg, J. et al., Eur. Phys. J. **C12**, 411 (2000).

[25] Chekanov, S. et al., Eur. Phys. J. **C28**, 175 (2003).

[26] ZEUS Coll., Chekanov, S. et al., Phys. Lett. **B539**, 197 (2002).

[27] Chekanov, S. and others, Phys. Rev. **D70**, 052001 (2004).

[28] ZEUS Coll., Chekanov, S. et al., Eur. Phys. J. **C32**, 16 (2003).

[29] Callan, Curtis G. , Jr. and Gross, David J., Phys. Rev. Lett. **22**, 156 (1969).

[30] Altarelli, Guido and Martinelli, G., Phys. Lett. **B76**, 89 (1978).

[31] Cooper-Sarkar, Amanda M. and Ingelman, G. and Long, K. R. and Roberts, R. G. and Saxon, D. H., Z. Phys. **C39**, 281 (1988).

[32] G. Altarelli et al., *Resummation*. These proceedings.

[33] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Phys. Lett. **B606**, 123 (2005).

[34] Bartels, Jochen and Golec-Biernat, K. and Peters, K., Eur. Phys. J. **C17**, 121 (2000).

[35] Adloff, C. and others, Phys. Lett. **B393**, 452 (1997).

[36] Adloff, C. and others, *Determination of the longitudinal proton structure function f(l)(x,q**2) at low q**2*. Preprint CH1prelim-03-043. To appear in the proceedings of 32nd International Conference on High-Energy Physics (ICHEP 04), Beijing, China, 16-22 Aug 2004.

[37] Adloff, C. and others, Phys. Lett. **B520**, 183 (2001).

[38] Klein, M., *On the future measurement of the longitudinal structure function at low x at hera*. Prepared for 12th International Workshop on Deep Inelastic Scattering (DIS 2004), Strbske Pleso, Slovakia, 14-18 Apr 2004.

[39] Feltesse, J., *On a measurement of the longitudinal structure function f_l at hera*. Preprint in preparation. Talk at the Ringberg Workshop (October 2005), to be published in the proceedings.

[40] Willeke, J, *Prospects for operating hera with lower proton energy at hera*. Preprint unpublished memo (October 2005).

[41] Chekanov, S. et al., Phys. Lett. **B553**, 141 (2003).

[42] Towell, R. S. et al., Phys. Rev. **D64**, 052002 (2001).

[43] Portheault, B. Thesis, Univ. Paris XI Orsay, March 2005.

[44] Alexopoulos, T. and others, *Electron deuteron scattering with hera, a letter of intent for an experimental programme with the h1 detector*. DESY-03-194;
Abramovicz, H. et al., *A new experiment for hera*. 2003. MPP-2003-62;
Willeke, F. and Hoffstaetter, G. Talks at the Workshop on the Future of DIS, Durham 2001.

[45] Acosta, D. et al., Phys. Rev. **D71**, 051104 (2005);
Heinemann, B. Talk at this workshop.

[46] Stirling, J. Talk at the Binn workshop on LHC Physics, 2003.

[47] Strikman, M., private communication.

[48] Forte, S., private communication.

[49] Botje, M. and Klein, M. and Pascaud, C. Preprint hep-ph/9609489. HERA Physics Workshop 1996/97.

[50] Chekanov, S. et al., Phys. Lett. **B547**, 164 (2002).

[51] Chekanov S. et al., Eur. Phys. J. **C23**, 615 (2002).

[52] Targett-Adams, C., private communication.

[53] Frixione, S. and Ridolfi, G., Nucl. Phys., **B507**, 315 (1997).

[54] Aurenche, P. and Guillet, J. and Fontannaz, M., Z. Phys. **C64**, 621 (1994).

[55] Giele, W. T. and Glover, E. W. N. and Kosower, D. A., Nucl. Phys. **B403**, 633 (1993).

[56] Whalley, M. R. and Bourilkov, D. and Group, R. C., *The les houches accord pdfs (lhapdf) and lhaglue*. Preprint hep-ph/0508110, 2005.

[57] Glück, M. and Reya, E. and Vogt, A., Z. Phys. **C67**, 433 (1995).

[58] Thorne, R., private communication.

[59] Martin, A. D. and Stirling, W. J. and Roberts, R. G., Phys. Lett. **B354**, 155 (1995).

[60] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C35**, 2004 (2004).

[61] Graudenz, D. and Hampel, M. and Vogt, A. and Berger, C., Z. Phys. **C70**, 77 (1996).

[62] Kosower, D. A., Nucl. Phys. **B520**, 263 (1998).

[63] Stratmann, M. and Vogelsang, W., Phys. Rev. **D64**, 114007 (2001).

[64] Wobisch, M. Thesis, RWTH Aachen, PITHA 00/12 and DESY-THESIS-2000-049, December 2000.

[65] Ratcliffe, P. G., Phys. Rev. **D63**, 116004 (2001).

[66] Dasgupta, M. and Salam, G. P., Eur. Phys. J. **C24**, 213 (2002).

[67] Nagy, Z., Phys. Rev. **D68**, 094002 (2003);
Nagy, Z., Phys. Rev. Lett. **88**, 122003 (2002);
Nagy, Z. and Trocsanyi, Z., Phys. Rev. Lett. **87**, 082001 (2001).

DGLAP evolution and parton fits

S. I. Alekhin, J. Blümlein, H. Böttcher, L. Del Debbio, S. Forte, A. Glazov, A. Guffanti, J. Huston, G. Ingelman J. I. Latorre, S. Moch, A. Piccione, J. Pumplin, V. Ravindran, J. Rojo G.P. Salam, R.S. Thorne, J.A.M. Vermaseren, A. Vogt

1 DGLAP evolution and parton fits [1]

1.1 Introduction

The high-precision data from HERA and the anticipated data from LHC open the possibility for a precise determination of parton distributions. This, however, requires an improvement in the theoretical description of DIS and hard hadronic scattering processes, as well as an improvement of the techniques used to extract parton distributions from the data.

The determination of perturbative QCD corrections has undergone substantial progress recently. The key ingredient of a complete next-to-next-to-leading order (NNLO) prediction in perturbative QCD are the recently calculated three-loop splitting functions which govern the scale dependence of PDFs. Extensions in the accuracy of the perturbative predictions yet beyond NNLO are given by the three-loop coefficient functions for F_2, while the coefficient functions for F_L at this order are actually required to complete the NNLO predictions. Section 2 briefly discusses the recent results and their phenomenological implications. Certain mathematical aspects, which are important in the calculation of higher order corrections in massless QCD are presented in section 3. In particular, algebraic relations in Mellin-N space are pointed out, which are of importance for harmonic sums, harmonic polylogarithms and multiple ζ-values.

These calculation of the PDF evolution to NNLO in perturbative QCD are used in section 4 to provide an update and extension of a set of benchmark tables for the evolution of parton distributions of hadrons. These benchmark tables were first presented in the report of the QCD/SM working group at the 2001 Les Houches workshop, but based on approximate NNLO splitting functions, which are superseded by the exact results which are now available. In addition, section 4 now includes also reference tables for the case of polarized PDF evolution.

Whereas in principle the x-shapes of PDFs at low scales can be determined from first principles using non-perturbative methods, in practice at present this is only possible using models (briefly touched in in section 5). Therefore, an accurate determination of PDFs requires a global QCD fit to the data, which is the subject of sections 6–8.

Section 6 discusses in particular the impact on parton fits of NNLO corrections on the one hand, and of the inclusion of Drell-Yan data and future LHC data on the other hand. It then presents values for a benchmark fit together with a table of correlation coefficients for the parameter obtained in the fit. This benchmark fit is then re-examined in sec. 7, along with a comparison between PDFs and the associated uncertainty obtained using the approaches of Alekhin and the MRST group. The differences between these benchmark partons and the actual global fit partons are also discussed, and used to explore complications inherent in extracting PDFs with uncertainties. Finally, in section 8 the stability of PDF determinations in NLO global analyses is re-investigated and the results of the CTEQ PDF group on this issue are summarized.

An alternative approach to a completely bias-free parameterization of PDFs is presented in section 9. There, a neural network approach to global fits of parton distribution functions is introduced and work on unbiased parameterizations of deep-inelastic structure functions with faithful estimation of

[1] Subsection coordinators: A. Glazov, S. Moch

their uncertainties is reviewed together with a summary of the current status of neural network parton distribution fits.

2 Precision Predictions for Deep-Inelastic Scattering [2]

With high-precision data from HERA and in view of the outstanding importance of hard scattering processes at the LHC, a quantitative understanding of deep-inelastic processes is indispensable, necessitating calculations beyond the standard next-to-leading order of perturbative QCD.

In this contribution we review recent results for the complete next-to-next-to-leading order (NNLO, N^{2}LO) approximation of massless perturbative QCD for the structure functions F_1, F_2, F_3 and F_L in DIS. These are based on the second-order coefficient functions [1, 2, 3, 4, 5], the three-loop splitting functions which govern the evolution of unpolarized parton distributions of hadrons [6, 7] and the three-loop coefficient functions for $F_L = F_2 - 2xF_1$ in electromagnetic (photon-exchange) DIS [8, 9]. Moreover we discuss partial N^{3}LO results for F_2, based on the corresponding three-loop coefficient functions also presented in Ref. [9]. For the splitting functions P and coefficient functions C we employ the convention

$$P(\alpha_s) = \sum_{n=0} \left(\frac{\alpha_s}{4\pi}\right)^{n+1} P^{(n)}, \qquad C(\alpha_s) = \sum_{n=0} \left(\frac{\alpha_s}{4\pi}\right)^{n} C^{(n)} \tag{1}$$

for the expansion in the running coupling constant α_s. For the longitudinal structure function F_L the third-order corrections are required to complete the NNLO predictions, since the leading contribution to the coefficient function C_L is of first order in the strong coupling constant α_s.

In the following we briefly display selected results to demonstrate the quality of precision predictions for DIS and their effect on the evolution. The exact (analytical) results to third order for the quantities in Eq. (1) are too lengthy, about $\mathcal{O}(100)$ pages in normalsize fonts and will not be reproduced here. Also the method of calculation is well documented in the literature [5, 10, 6, 7, 9, 11]. In particular, it proceeds via the Mellin transforms of the functions of the Bjorken variable x,

$$A(N) = \int_0^1 dx\, x^{N-1} A(x)\,. \tag{2}$$

Selected mathematical aspects of Mellin transforms are discussed in section 3.

2.1 Parton evolution

The well-known $2n_f - 1$ scalar non-singlet and 2×2 singlet evolution equations for n_f flavors read

$$\frac{d}{d\ln\mu_f^2}\, q_{\rm ns}^i = P_{\rm ns}^i \otimes q_{\rm ns}^i\,, \qquad i = \pm, {\rm v}\,, \tag{3}$$

for the quark flavor asymmetries $q_{\rm ns}^{\pm}$ and the valence distribution $q_{\rm ns}^{\rm v}$, and

$$\frac{d}{d\ln\mu_f^2} \begin{pmatrix} q_{\rm s} \\ g \end{pmatrix} = \begin{pmatrix} P_{\rm qq} & P_{\rm qg} \\ P_{\rm gq} & P_{\rm gg} \end{pmatrix} \otimes \begin{pmatrix} q_{\rm s} \\ g \end{pmatrix} \tag{4}$$

for the singlet quark distribution q_s and the gluon distribution g, respectively. Eqs. (3) and (4) are governed by three independent types of non-singlet splitting functions, and by the 2×2 matrix of singlet splitting functions. Here $\otimes$ stands for the Mellin convolution. We note that benchmark numerical solutions to NNLO accuracy of Eqs. (3) and (4) for a specific set of input distributions are given in section 4. Phenomenological QCD fits of parton distributions in data analyses are extensively discussed in sections 6–8. An approach based on neural networks is described in section 9.

[2]Contributing authors: S. Moch, J.A.M. Vermaseren, A. Vogt

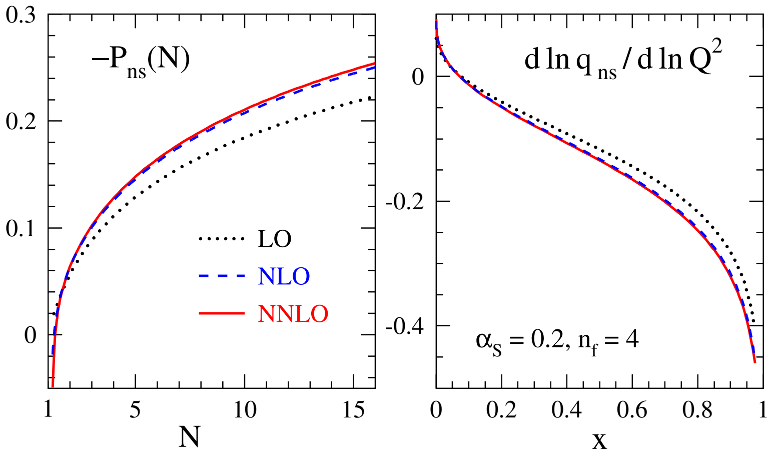


Fig. 1: On the left we show the perturbative expansion of $P_{\rm ns}^{\rm v}(N)$, and on the right the resulting perturbative expansion of the logarithmic scale derivative $d \ln q_{\rm ns}/d \ln \mu_f^2$ is displayed for a model input. See the text for details.

Let us start the illustration of the precision predictions by looking at the parton evolution and at large Mellin-N (large Bjorken-x) behavior. Fig. 1 shows the stability of the perturbative expansion which is very benign and indicates, for $\alpha_{\rm s} \lesssim 0.2$, corrections of less than 1% beyond NNLO. On the left we show the results for the perturbative expansion of $P_{\rm ns}$ in Mellin space, cf. Eqs. (1), (2). We employ four active flavors, $n_f = 4$, and an order-independent value for the strong coupling constant,

$$\alpha_{\rm s}(\mu_0^2) \;=\; 0.2 \;, \tag{5}$$

which corresponds to $\mu_0^2 \simeq 25 \ldots 50\,\text{GeV}^2$ for $\alpha_{\rm s}(M_Z^2) = 0.114 \ldots 0.120$ beyond the leading order. On the right of Fig. 1 the perturbative expansion of the logarithmic derivative, cf. Eqs. (1), (3), is illustrated at the standard choice $\mu_r = \mu_f$ of the renormalization scale. We use the schematic, but characteristic model distribution,

$$x q_{\rm ns}(x, \mu_0^2) \;=\; x^{\,0.5}(1-x)^3 \,. \tag{6}$$

The normalization of $q_{\rm ns}$ is irrelevant at this point, as we consider the logarithmic scale derivative only.

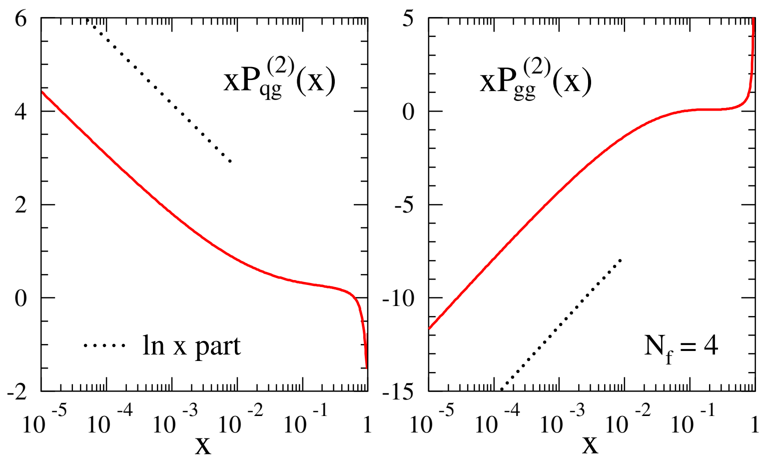


Fig. 2: The three-loop gluon-quark (left) and gluon-gluon (right) splitting functions together with the leading small-x contribution (dotted line).

Next, let us focus on the three-loop splitting functions at small momentum fractions x, where the splitting functions $P_{\rm ig}$ in the lower row of the 2×2 matrix in Eq. (4), representing ${\rm g} \to {\rm i}$ splittings, are most important. In Fig. 2 we show, again for $n_f = 4$, the three-loop splitting functions $P^{(2)}_{\rm qg}$ and $P^{(2)}_{\rm gg}$ together with the leading small-x term indicated separately for $x < 0.01$. In the present singlet case the leading logarithmic small-x limits $\sim x^{-1} \ln x$ of Refs. [12, 13] are confirmed together with the general structure of the BFKL limit [14, 15, 16]. The same holds for the leading small-x terms $\ln^4 x$ in the non-singlet sector [17, 18], with the qualification that a new, unpredicted leading logarithmic contribution is found for the color factor $d^{abc}d_{abc}$ entering at three loops for the first time.

It is obvious from Fig. 2 (see also Refs. [5, 6, 7, 11]) that the leading $x \to 0$-terms alone are insufficient for collider phenomenology at HERA or the LHC as they do not provide good approximations of the full results at experimentally relevant small values of x. Resummation of the small-x terms and various phenomenological improvements are discussed in detail in [19].

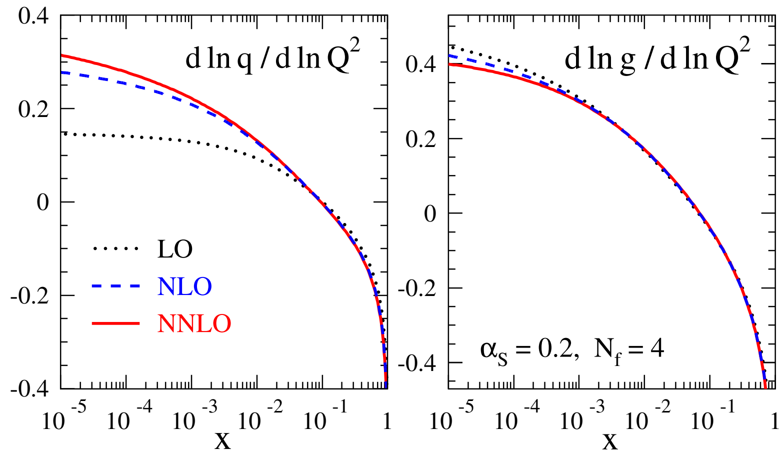


Fig. 3: The perturbative expansion of the scale derivatives (4) of the singlet distributions (7).

In the same limit of small x, it is instructive to look at the evolution of parton distributions. Again, we choose the reference scale of Eq. (5), $n_f = 4$ and the sufficiently realistic model distributions

$$
\begin{aligned}
xq_{\rm s}(x,\mu_0^2) &= 0.6\, x^{-0.3}(1-x)^{3.5}\,(1+5.0\, x^{0.8}) \\
xg(x,\mu_0^2) &= 1.6\, x^{-0.3}(1-x)^{4.5}\,(1-0.6\, x^{0.3})
\end{aligned} \tag{7}
$$

irrespective of the order of the expansion to facilitate direct comparisons of the various contributions. Of course, this order-independence does not hold for actual data-fitted parton distributions like those in sections 6–8. In Fig. 3 we display the perturbative expansion of the scale derivative for the singlet quark and gluon densities at $\mu_f^2 = \mu_0^2$ for the initial conditions specified in Eqs. (5) and (7). For the singlet quark distribution the total NNLO corrections, while reaching 10% at $x = 10^{-4}$, remain smaller than the NLO results by a factor of eight or more over the full x-range. For the gluon distribution already the NLO corrections are small and the NNLO contribution amounts to only 3% for x as low as 10^{-4}. Thus, we see in Fig. 3 that the perturbative expansion is very stable. It appears to converge rapidly at $x > 10^{-3}$, while relatively large third-order corrections are found for very small momenta $x \lesssim 10^{-4}$.

2.2 Coefficient functions

While the previous considerations were addressing the evolution of parton distributions, we now turn to the further improvements of precision predictions due to the full third-order coefficient functions for the structure functions F_2 and F_L in electromagnetic DIS [8, 9]. The results for F_L complete the NNLO description of unpolarized electromagnetic DIS, and the third-order coefficient functions for F_2 form, at not too small values of the Bjorken variable x, the dominant part of the next-to-next-to-next-to-leading order (N^{3}LO) corrections. Thus, they facilitate improved determinations of the strong coupling α_s from scaling violations.

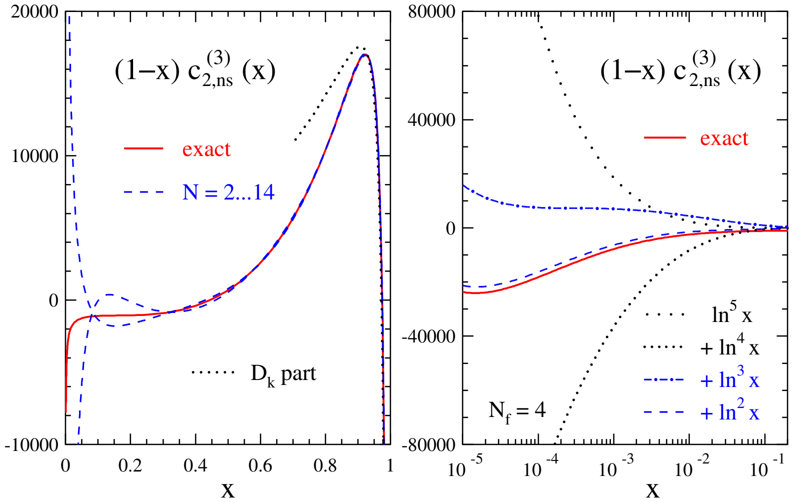


Fig. 4: The three-loop non-singlet coefficient function $c^{(3)}_{2,\mathrm{ns}}(x)$ in the large-x (left) and the small-x (right) region, multiplied by $(1-x)$ for display purposes.

Let us start with the three-loop coefficient functions for F_2 in the non-singlet case. In Fig. 4 we display the three-loop non-singlet coefficient function $c^{(3)}_{2,\mathrm{ns}}(x)$ for $n_f = 4$ flavors. We also show the soft-gluon enhanced terms $\mathcal{D}_k$ dominating the large-x limit,

$$\mathcal{D}_k \quad = \quad \frac{\ln^{2k-1}(1-x)}{(1-x)_+}\,, \tag{8}$$

and the small-x approximations obtained by successively including enhanced logarithms $\ln^k x$. However the latter are insufficient for an accurate description of the exact result. The dashed band in Fig. 4 shows the uncertainty of previous estimates [20] mainly based on the calculation of fixed Mellin moments [21, 22, 23]. For a detailed discussion of the soft-gluon resummation of the the $\mathcal{D}_k$ terms, we refer to [19].

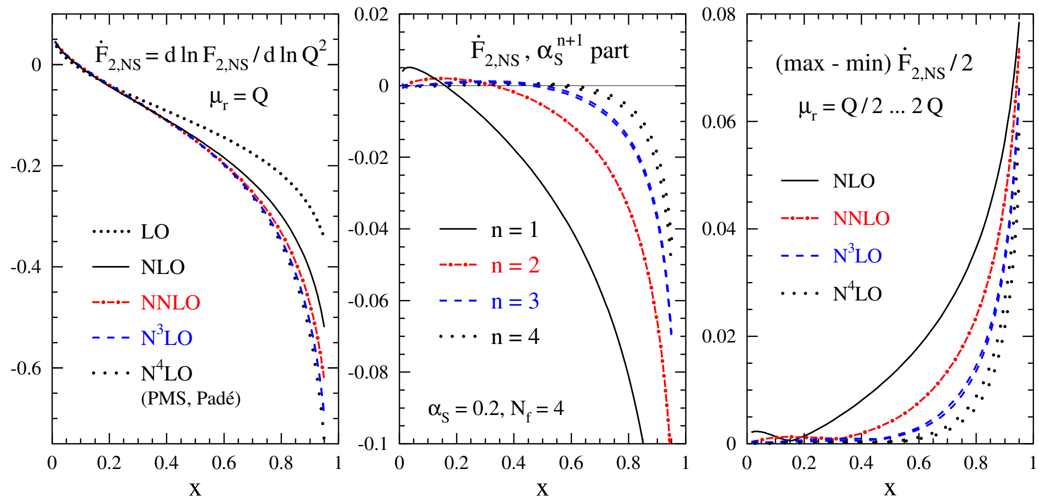


Fig. 5: The perturbative expansion of the logarithmic scale derivative of the non-singlet structure function $F_{2,\mathrm{ns}}$. The results up to NNLO are exact, while those at N^{3}LO are very good approximations. The N^{4}LO corrections have been estimated by various methods.

Building on the coefficient functions, it is interesting to study the perturbative expansion of the logarithmic scale derivative for the non-singlet structure function $F_{2,\mathrm{ns}}$. To that end we use in Fig. 5

again the input shape Eq. (6) (this time for $F_{2,\mathrm{ns}}$ itself) irrespective of the order of the expansion, $n_f = 4$ flavors and the reference scale of Eq. (5). The N^{4}LO approximation based on Padé summations of the perturbation series can be expected to correctly indicate at least the rough size of the four-loop corrections, see Ref. [9] for details. From Fig. 5 we see that the three-loop results for F_2 can be employed to effectively extend the main part of DIS analyses to the N^{3}LO at $x > 10^{-2}$ where the effect of the unknown fourth-order splitting functions is expected to be very small. This has, for example, the potential for a 'gold-plated' determination of $\alpha_\mathrm{s}(M_Z)$ with an error of less than 1% from the truncation of the perturbation series. On the right hand side of Fig. 5 the scale uncertainty which is conventionally estimated by

$$\Delta \dot{f} \quad \equiv \quad \frac{1}{2}\left(\max\,[\dot{f}(x,\mu_r^2)] - \min\,[\dot{f}(x,\mu_r^2)]\right) \,, \tag{9}$$

is plotted, where the scale varies $\mu_r \in [Q/2, 2Q]$.

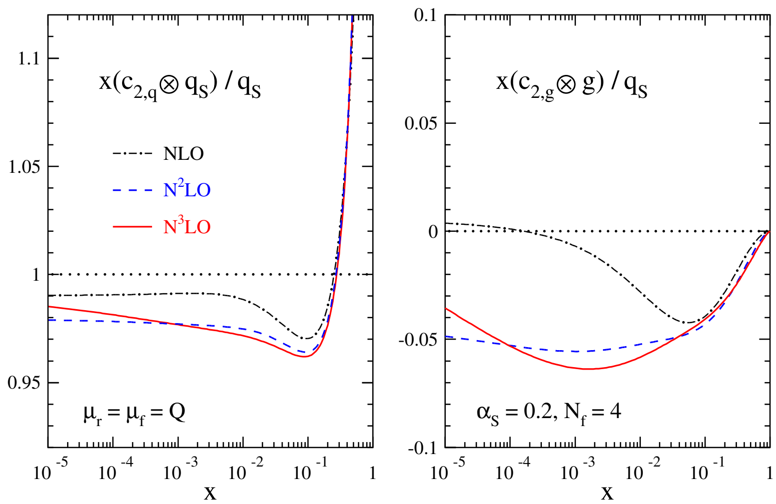


Fig. 6: The perturbative expansion up to three loops (N^{3}LO) of the quark (left) and gluon (right) contributions to singlet structure function F_2.

In the singlet case, we can study the quark and gluon contributions to the structure function F_2. In Fig. 6 we plot the perturbative expansion up to N^{3}LO of the quark and gluon contributions to structure function $F_{2,\mathrm{s}}$ at the scale (5) using the distributions (7). All curves have been normalized to the leading-order result $F_{2,\mathrm{s}}^{\mathrm{LO}} = \langle e^2 \rangle \, q_\mathrm{s}$. Fig. 6 nicely illustrates the perturbative stability of the structure function F_2.

Finally, we address the longitudinal structure function F_L at three loops. In the left part of Fig. 7 we plot the singlet-quark and gluon coefficient functions $c_{L,\mathrm{q}}$ and $c_{L,\mathrm{g}}$ for F_L up to the third order for four flavors and the α_s-value of Eq. (5). The curves have been divided by $a_\mathrm{s} = \alpha_s/(4\pi)$ to account for the leading contribution being actually of first order in the strong coupling constant α_s. Both the second-order and the third-order contributions are rather large over almost the whole x-range. Most striking, however, is the behavior at very small values of x, where the anomalously small one-loop parts are negligible against the (negative) constant two-loop terms, which in turn are completely overwhelmed by the (positive) new three-loop corrections $xc_{L,a}^{(3)} \sim \ln x + \mathit{const}$, which we have indicated in Fig. 7.

To assess the effect for longitudinal structure function F_L, we convolute in Fig. 7 on the right the coefficient functions with the input shapes Eq. (7) for $n_f = 4$ flavors and the reference scale of Eq. (5). A comparison of the left and right plots in Fig. 7 clearly reveals the smoothening effect of the Mellin convolutions. For the chosen input conditions, the (mostly positive) NNLO corrections to the flavor-singlet F_L amount to less than 20% for $5 \cdot 10^{-5} < x < 0.3$. In data fits we expect that the parton

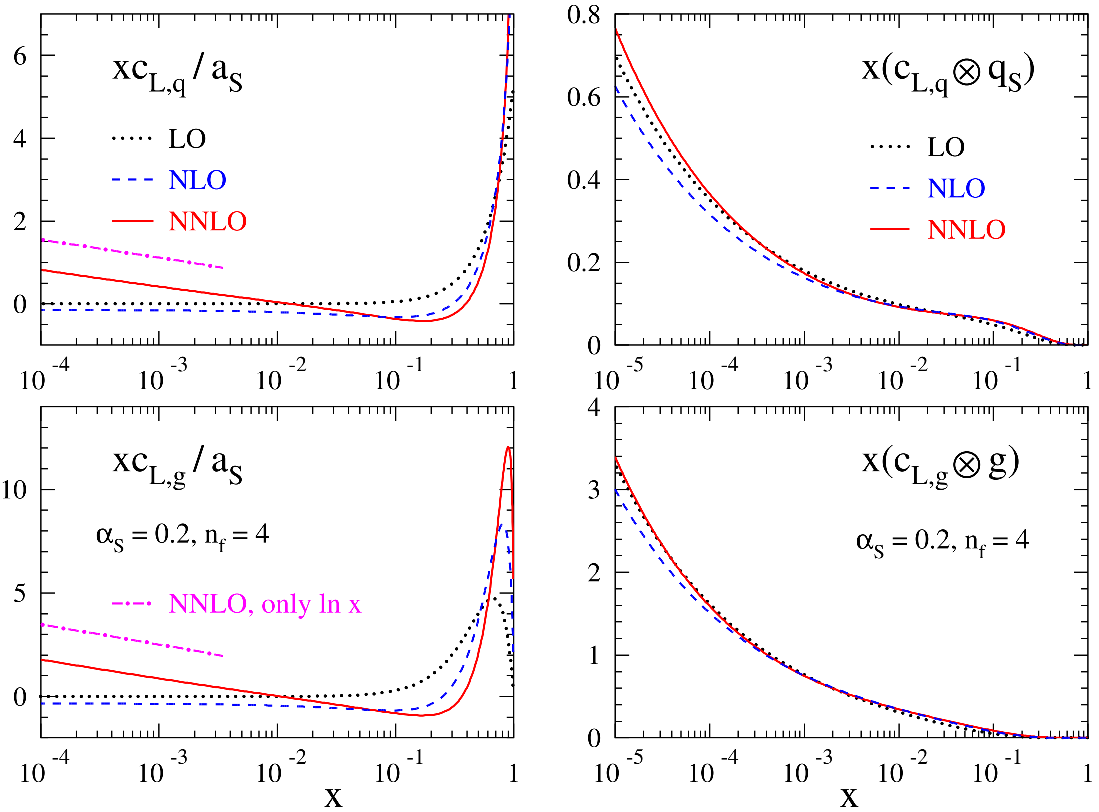


Fig. 7: The perturbative expansion to N^{2}LO of the longitudinal singlet-quark and gluon coefficient functions to third order multiplied by x for display purposes (left) and of the quark and gluon contributions to singlet structure function F_L (right).

distributions, in particular the gluon distribution, will further stabilize the overall NNLO/NLO ratio. Thus, at not too small scales, F_L is a quantity of good perturbative stability, for the x-values accessible at HERA, see Ref. [8] for more details.

3 Mathematical Structure of Higher Order Corrections [3]

The QCD anomalous dimensions and Wilson coefficients for structure functions are single scale quantities and may be expressed in simple form in Mellin space in terms of polynomials of harmonic sums and ration functions of the Mellin variable. Unlike the case in various calculations using representations in momentum-fraction (z-) space the use of multiple nested harmonic sums leads to a synchronization in language. Furthermore, significant simplifications w.r.t. the number of functions needed can be achieved. This is due to algebraic [24, 25] relations between these quantities, which in a similar way are also present between harmonic polylogarithms [26] and multiple ζ-values [27]. These relations result from the the specific index pattern of the objects considered and their multiplication relation and do not refer to further more specific properties. In Table 1 we illustrate the level of complexity which one meets in case of harmonic sums. To three–loop order weight w=6 harmonic sums occur. The algebraic relations for the whole class of harmonic sums lead to a reduction by a factor of ~ 4 (column 3). As it turns out, physical pseudo-observables, as anomalous dimensions and Wilson-coefficients in the $\overline{\text{MS}}$ scheme, to 2– resp. 3–loop order depend on harmonic sums only, in which the index $\{-1\}$ never occurs. The algebraic reduction for this class is illustrated in column 5. We also compare the complexity of only non-alternating harmonic sums and their algebraic reduction, which is much lower. This class of sums is, however, not wide enough to describe the above physical quantities. In addition to the algebraic relations of harmonic

[3]Contributing authors: J. Blümlein, H. Böttcher, A. Guffanti, V. Ravindran

sums structural relations exist, which reduces the basis further [28].

Weight	Number of					
	Sums	a-basic sums	Sums $\neg\{-1\}$	a-basic sums	Sums $i > 0$	a-basic sums
1	2	2	1	1	1	1
2	6	3	3	2	2	1
3	18	8	7	4	4	2
4	54	18	17	7	8	3
5	162	48	41	16	16	6
6	486	116	99	30	32	9
7	1458	312	239	68	64	18

Table 1: Number of alternating and non-alternating harmonic sums in dependence of their weight, [28].

Using all these relations one finds that 5 basic functions are sufficient to describe all 2–loop Wilson coefficients for deep–inelastic scattering [29] and further 8 [30] for the 3–loop anomalous dimensions. Their analytic continuations to complex values of the Mellin variable are given in [31, 32]. These functions are the (regularized) Mellin transforms of :

$$\begin{array}{llll} \dfrac{\ln(1+x)}{1+x}, & \dfrac{\mathrm{Li}_2(x)}{1\pm x}, & \dfrac{S_{1,2}(x)}{1\pm x}, & \dfrac{\mathrm{Li}_4(x)}{x\pm 1}, \\ \dfrac{S_{1,3}(x)}{1+x}, & \dfrac{S_{2,2}(x)}{x\pm 1}, & \dfrac{\mathrm{Li}_2^2(x)}{1+x}, & \dfrac{S_{2,2}(-x)-\mathrm{Li}_2^2(-x)/2}{x\pm 1}\,. \end{array} \tag{10}$$

It is remarkable, that the numerator–functions in (10) are Nielsen integrals [33] and polynomials thereof, although one might expect harmonic polylogarithms [26] outside this class in general. The representation of the Wilson coefficients and anomalous dimensions in the way described allows for compact expressions and very fast and precise numerical evaluation well suited for fitting procedures to experimental data.

3.1 Two–loop Processes at LHC in Mellin Space

Similar to the case of the Wilson coefficients in section 3 one may consider the Wilson coefficients for inclusive hard processes at hadron colliders, as the Drell–Yan process to $O(\alpha_s^2)$ [34, 35, 36], scalar or pseudoscalar Higgs–boson production to $O(\alpha_s^3)$ in the heavy–mass limit [37, 38, 39, 40, 41, 42], and the 2–loop time–like Wilson coefficients for fragmentation [43, 44, 45]. These quantities have been analyzed in [46, 47] w.r.t. their general structure in Mellin space. The cross section for the Drell–Yan process and Higgs production is given by

$$\sigma\left(\frac{\hat{s}}{s}, Q^2\right) = \int_x^1 \frac{dx_1}{x_1} \int_{x/x_1}^1 \frac{dx_2}{x_2} f_a(x_1,\mu^2) f_b(x_2,\mu^2) \hat{\sigma}\left(\frac{x}{x_1 x_2}, \frac{Q^2}{\mu^2}\right) , \tag{3.11}$$

with $x = \hat{s}/s$. Here, $f_c(x,\mu^2)$ are the initial state parton densities and μ^2 denotes the factorization scale. The Wilson coefficient of the process is $\hat{\sigma}$ and Q^2 is the time–like virtuality of the s–channel boson. Likewise, for the fragmentation process of final state partons into hadrons in *pp*–scattering one considers the double differential final state distribution

$$\frac{d^2\sigma^H}{dx d\cos\theta} = \frac{3}{8}(1+\cos^2\theta)\frac{d\sigma_T^H}{dx} + \frac{3}{4}\sin^2\theta \frac{d\sigma_L^H}{dx} \,. \tag{3.12}$$

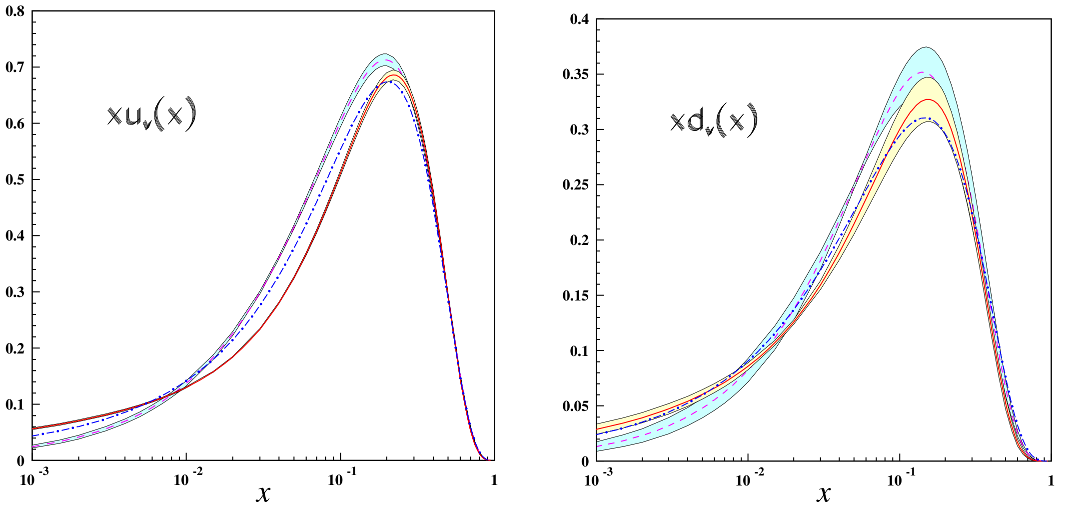


Fig. 8: xu_v and xd_v at $Q_0^2 = 4\text{GeV}^2$ (full lines) [48]; dashed lines [50]; dash-dotted lines [51].

Here,

$$\frac{d\sigma_k^H}{dx} = \int_x^1 \frac{dz}{z}\left[\sigma_{\rm tot}^{(0)}\left\{D_S^H\left(\frac{x}{z},M^2\right)C_{k,q}^{\rm S}(z,Q^2/M^2) + D_g^H\left(\frac{x}{z},M^2\right)C_{k,q}^{\rm S}(z,Q^2/M^2)\right\}\right.$$
$$\left. + \sum_{p=1}^{N_f}\sigma_p^{(0)}D_{{\rm NS},p}^H\left(\frac{x}{z},M^2\right)C_{k,q}^{\rm NS}(z,Q^2/M^2)\right]. \quad (3.13)$$

In the subsystem cross-sections σ the initial state parton distributions are included. D_k^H denote the non–perturbative fragmentation functions and $C_{k,i}^{\rm S,NS}(z,Q^2/M^2)$ the respective time–like Wilson coefficients describing the fragmentaion process for a parton i into the hadron H.

Although these Wilson coefficients are not directly related to the 2–loop Wilson coefficients for deeply inelastic scattering, one finds for these functions at most the same set of basic functions as given above. Again one obtains very fast and concise numerical programs also for these processes working in Mellin space, which will be well suited for inclusive analyses of experimental collider data at LHC in the future.

3.2 Non–Singlet Parton Densities at $O(\alpha_s^3)$

The precision determination of the QCD–scale $\Lambda_{\rm QCD}$ and of the idividual parton densities is an important issue for the whole physics programme at LHC since all measurements rely on the detailed knwoledge of this parameter and distribution functions. In Ref. [48] first results were reported of a world data analysis for charged lepton-$p(d)$ scattering w.r.t. the flavor non–singlet sector at $O(\alpha_s^3)$ accuracy. The flavor non–singlet distributions $xu_v(x,Q^2)$ and $xd_v(x,Q^2)$ were determined along with fully correlated error bands giving parameterizations both for the values and errors of these distributions for a wide range in x and Q^2. In Figure 8 these distributions including their error are shown. The value of the strong coupling constant $\alpha_s(M_Z^2)$ was determined as $0.1135 + 0.0023 - 0.0026$ (exp.) The full analysis is given in [49], including the determination of higher twist contributions in the large x region both for $F_2^p(x,Q^2)$ and $F_2^d(x,Q^2)$.

3.3 Scheme-invariant evolution for unpolarzed DIS structure functions

The final HERA-II data on unpolarized DIS structure functions, combined with the present world data from other experiments, will allow to reduce the experimental error on the strong coupling constant,

$\alpha_s(M_Z^2)$, to the level of 1% [52]. On the theoretical side the NLO analyzes have intrinsic limitations which allow no better than 5% accuracy in the determination of α_s [53]. In order to match the expected experimental accuracy, analyzes of DIS structure functions need then to be carried out at the NNLO–level. To perform a full NNLO analysis the knowledge of the 3-loop β-function coefficient, β_2, the 2– resp. 3–loop Wilson coefficients and the 3-loop anomalous dimensions is required. With the calculation of the latter [6, 7], the whole scheme–independent set of quantities is known, thus allowing a complete NNLO study of DIS structure functions.

Besides the standard approach solving the QCD evolution equations for parton densities in the $\overline{\mathrm{MS}}$ scheme it appears appealing to study scheme–invariant evolution equations [54]. Within this approach the input distributions at a scale Q_0^2 are measured experimentally. The only parameter to be determined by a fit to data is the QCD–scale Λ_{QCD}. To perform an analysis in the whole kinematic region the non–singlet [48] contribution has to be separated from the singlet terms of two measured observables. In practice these can be chosen to be $F_2(x,Q^2)$ and $\partial F_2(x,Q^2)/\partial\ln(Q^2)$ or $F_2(x,Q^2)$ and $F_L(x,Q^2)$ if the latter structure function is measured well enough. Either $\partial F_2(x,Q^2)/\partial\ln(Q^2)$ or $F_L(x,Q^2)$ play a role synonymous to the gluon distribution while $F_2(x,Q^2)$ takes the role of the singlet–quark distribution compared to the standard analysis. These equations do no longer describe the evolution of universal quantities depending on the choice of a scheme but of process–dependent quantities which are observables and thus factorization scheme–indedependent. Since the respective evolution kernels are calculated in perturbation theory the dependence on the renormalization scale remains and becomes smaller with the order in the coupling constant included.

Physical evolution kernels have been studied before in [55, 56, 57]. The 3–loop scheme–invariant evolution equations were solved in the massless case in [54]. This analysis is extended including the heavy flavor contributions at present [49]. The large complexity of the evolution kernels can only be handeled in Mellin space since in z–space various inverse and direct Mellin convolutions would be required numerically, causing significant accuracy and run–time problems. The inclusion of the heavy flavor contributions is possible using the parameterizations [58].

In Fig. 8 we present the scheme invariant evolution for the structure functions F_2 and $\partial F_2/\partial t$ to NNLO with $t = -2/\beta_0 \ln(\alpha_s(Q^2)/\alpha_s(Q_0^2))$. The input distribution at the reference scale are not extracted from data, but rather built up as a convolution of Wilson coefficients and PDFs, the latter being parametrised according to [59].

Scheme–invariant evolution equations allow a widely un–biased approach to determine the initial conditions for QCD evolution, which in general is a source of systematic effects which are difficult to control. On the other hand, their use requires to consider all correlations of the input measurements in a detailed manner experimentally. At any scale Q^2 mappings are available to project the observables evolved onto the quark–singlet and the gluon density in whatever scheme. In this way the question whether sign changes in the unpolarized gluon distribution in the $\overline{\mathrm{MS}}$ scheme do occur or do not occur in the small x region can be answered uniquely. As in foregoing analyses [60, 48] correlated error propagation throughout the evolution is being performed.

4 Updated reference results for the evolution of parton distributions [4]

In this contribution we update and extend our benchmark tables, first presented in the report of the QCD/SM working group at the 2001 Les Houches workshop [59], for the evolution of parton distributions of hadrons in perturbative QCD. Since then the complete next-to-next-to-leading order (NNLO) splitting functions have been computed [6, 7], see also section 2. Thus we can now replace the NNLO results of 2001 which were based on the approximate splitting functions of Ref. [61]. Furthermore we now include reference tables for the polarized case treated in neither Ref. [59] nor the earlier study during the 1995/6 HERA workshop [62]. Since the spin-dependent NNLO splitting functions are still unknown, we

[4]Contributing authors: G.P. Salam, A. Vogt

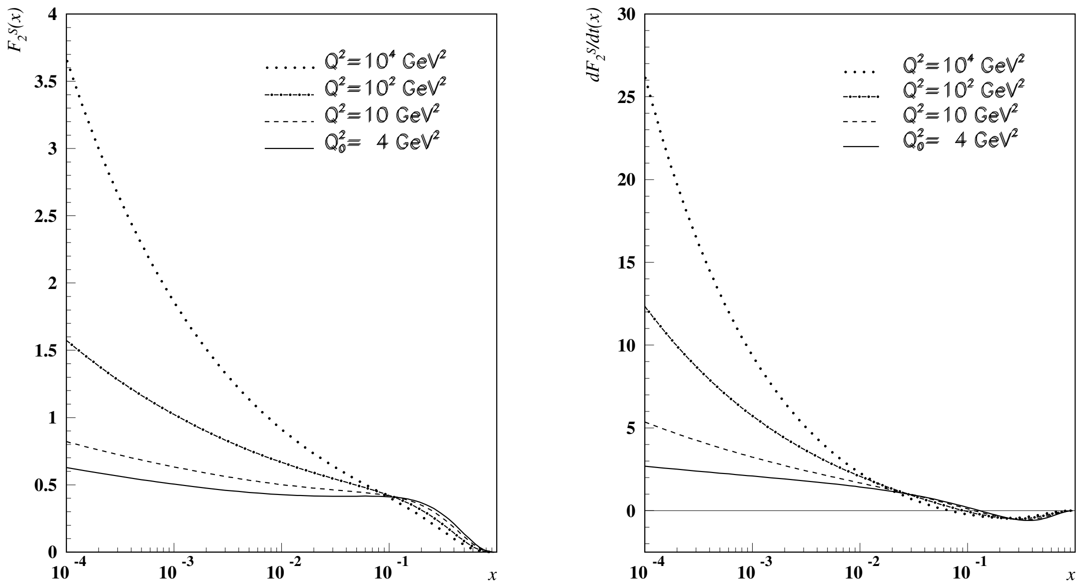


Fig. 9: NNLO scheme invariant evolution for the singlet part of the structure function F_2 and its slope $\partial F_2/\partial t$ for four massless flavours, [54].

have to restrict ourselves to the polarized leading-order (LO) and next-to-leading-order (NLO) evolution.

As in Ref. [59], we employ two entirely independent and conceptually different FORTRAN programs. At this point, the x-space code of G.S. is available from the author upon request, while the Mellin-space program of A.V. has been published in Ref. [63]. The results presented below correspond to a direct iterative solution of the N$^{\rm m}$LO evolution equations for the parton distributions $f_p(x,\mu_{\rm f}^2) \equiv p(x,\mu_{\rm f}^2)$, where $p = q_i,\, \bar{q}_i\,, g$ with $i = 1,\, \ldots, N_{\rm f}$,

$$\frac{d f_p(x,\mu_{\rm f}^2)}{d\ln\mu_{\rm f}^2} \;=\; \sum_{l=0}^{m} a_{\rm s}^{l+1}(\mu_{\rm r}^2) \int_x^1 \frac{dy}{y} \sum_{p'} P_{pp'}^{(l)}\left(\frac{x}{y}, \frac{\mu_{\rm f}^2}{\mu_{\rm r}^2}\right) f_{p'}(y,\mu_{\rm f}^2) \tag{4.14}$$

with the strong coupling, normalized as $a_{\rm s} \equiv \alpha_{\rm s}/(4\pi)$, given in terms of

$$\frac{d\, a_{\rm s}}{d\ln\mu_r^2} \;=\; \beta_{\rm N^mLO}(a_{\rm s}) \;=\; -\sum_{l=0}^{m} a_{\rm s}^{l+2}\, \beta_l \tag{4.15}$$

with $\beta_0 = 11 - 2/3\, N_{\rm f}$ etc. $\mu_{\rm r}$ and $\mu_{\rm f}$ represent the renormalization and mass-factorization scales in the $\overline{\rm MS}$ scheme. The reader is referred to Refs. [59, 63] for the scale dependence of the splitting functions $P^{(l)}$ and a further discussion of our solutions of Eqs. (4.14) and (4.15).

For the unpolarized case we retain the initial conditions as set up at the Les Houches meeting: The evolution is started at

$$\mu_{\rm f,0}^2 \;=\; 2\,{\rm GeV}^2 \ . \tag{4.16}$$

Roughly along the lines of the CTEQ5M parametrization [64], the input distributions are chosen as

$$\begin{aligned}
xu_{\rm v}(x,\mu_{\rm f,0}^2) &= 5.107200\, x^{0.8}\,(1-x)^3 \\
xd_{\rm v}(x,\mu_{\rm f,0}^2) &= 3.064320\, x^{0.8}\,(1-x)^4 \\
xg\,(x,\mu_{\rm f,0}^2) &= 1.700000\, x^{-0.1}(1-x)^5 \\
x\bar{d}\,(x,\mu_{\rm f,0}^2) &= .1939875\, x^{-0.1}(1-x)^6 \\
x\bar{u}\,(x,\mu_{\rm f,0}^2) &= (1-x)\, x\bar{d}\,(x,\mu_{\rm f,0}^2) \\
xs\,(x,\mu_{\rm f,0}^2) &= x\bar{s}\,(x,\mu_{\rm f,0}^2) \;=\; 0.2\, x(\bar{u}+\bar{d})(x,\mu_{\rm f,0}^2)
\end{aligned} \tag{4.17}$$

where, as usual, $q_{i,\mathrm{v}} \equiv q_i - \bar{q}_i$. The running couplings are specified by Eq. (4.15) and

$$\alpha_{\mathrm{s}}(\mu_{\mathrm{r}}^2 = 2\ \mathrm{GeV}^2) \ = \ 0.35 \ . \tag{4.18}$$

For simplicity initial conditions (4.17) and (4.18) are employed regardless of the order of the evolution and the (fixed) ratio of the renormalization and factorization scales.

For the evolution with a fixed number $N_{\mathrm{f}} > 3$ of quark flavours the quark distributions not specified in Eq. (4.17) are assumed to vanish at $\mu_{\mathrm{f},0}^2$, and Eq. (4.18) is understood to refer to the chosen value of N_{f}. For the evolution with a variable $N_{\mathrm{f}} = 3 \ldots 6$, Eqs. (4.16) and (4.17) always refer to three flavours. N_{f} is then increased by one unit at the heavy-quark pole masses taken as

$$m_{\mathrm{c}} \ = \ \mu_{\mathrm{f},0} \ , \quad m_{\mathrm{b}} \ = \ 4.5\ \mathrm{GeV}^2 \ , \quad m_{\mathrm{t}} \ = \ 175\ \mathrm{GeV}^2 \ , \tag{4.19}$$

i.e., Eqs. (4.14) and (4.15) are solved for a fixed number of flavours between these thresholds, and the respective matching conditions are invoked at $\mu_{\mathrm{f}}^2 = m_{\mathrm{h}}^2,\ h = c,\, b,\, t$. The matching conditions for the unpolarized parton distributions have been derived at NNLO in Ref. [65], and were first implemented in an evolution program in Ref. [66]. Note that, while the parton distributions are continuous up to NLO due to our choice of the matching scales, α_{s} is discontinuous at these flavour thresholds already at this order for $\mu_{\mathrm{r}} \neq \mu_{\mathrm{f}}$, see Refs. [67, 68]. Again the reader is referred to Refs. [59, 63] for more details.

Since the exact NNLO splitting functions $P^{(2)}$ are rather lengthy and not directly suitable for use in a Mellin-space program (see, however, Ref. [32]), the reference tables shown below have been computed using the parametrizations (4.22) – (4.24) of Ref. [6] and (4.32) – (4.35) of Ref. [7]. Likewise, the operator matrix element $\widetilde{A}_{\mathrm{hg}}^{\mathrm{S},2}$ entering the NNLO flavour matching is taken from Eq. (3.5) of Ref. [63]. The relative error made by using the parametrized splitting functions is illustrated in Fig. 10. It is generally well below 10^{-4}, except for the very small sea quark distributions at very large x.

Eqs. (4.16), (4.18) and (4.19) are used for the (longitudinally) polarized case as well, where Eq. (4.17) replaced by the sufficiently realistic toy input [63]

$$\begin{aligned}
xu_{\mathrm{v}} \ &= \ +1.3\, x^{0.7}\, (1-x)^3\, (1+3x) \\
xd_{\mathrm{v}} \ &= \ -0.5\, x^{0.7}\, (1-x)^4\, (1+4x) \\
xg \ &= \ +1.5\, x^{0.5}\, (1-x)^5 \\
x\bar{d} \ &= \ x\bar{u} \ = \ -0.05\, x^{0.3}\, (1-x)^7 \\
xs \ &= \ x\bar{s} \ = \ +0.5\, x\bar{d} \ .
\end{aligned} \tag{4.20}$$

As Eq. (4.17) in the unpolarized case, this input is employed regardless of the order of the evolution.

As in Ref. [59], we have compared the results of our two evolution programs, under the conditions specified above, at 500 x-μ_{f}^2 points covering the range $10^{-8} \leq x \leq 0.9$ and $2\ \mathrm{GeV}^2 \leq \mu_{\mathrm{f}}^2 \leq 10^6\ \mathrm{GeV}^2$. A representative subset of our results at $\mu_{\mathrm{f}}^2 = 10^4\ \mathrm{GeV}^4$, a scale relevant to high-E_T jets and close to m_{W}^2, m_{Z}^2 and, possibly, m_{Higgs}^2, is presented in Tables 2 – 6. These results are given in terms of the valence distributions, defined below Eq. (4.17), $L_\pm \equiv \bar{d} \pm \bar{u}$, and the quark-antiquark sums $q_+ \equiv q - \bar{q}$ for $q = s,\, c$ and, for the variable-N_{f} case, b.

For compactness an abbreviated notation is employed throughout the tables, i.e., all numbers $a \cdot 10^b$ are written as a^b. In the vast majority of the x-μ_{f}^2 points our results are found to agree to all five figures displayed, except for the tiny NLO and NNLO sea-quark distributions at $x = 0.9$, in the tables. Entries where the residual offsets between our programs lead to a different fifth digit after rounding are indicated by the subscript '$*$'. In these cases the number with the smaller modulus is given in the tables.

The approximate splitting functions [61], as mentioned above employed in the previous version [59] of our reference tables, have been used in (global) NNLO fits of the unpolarized parton distributions [51, 69], which in turn have been widely employed for obtaining NNLO cross sections, in particular for

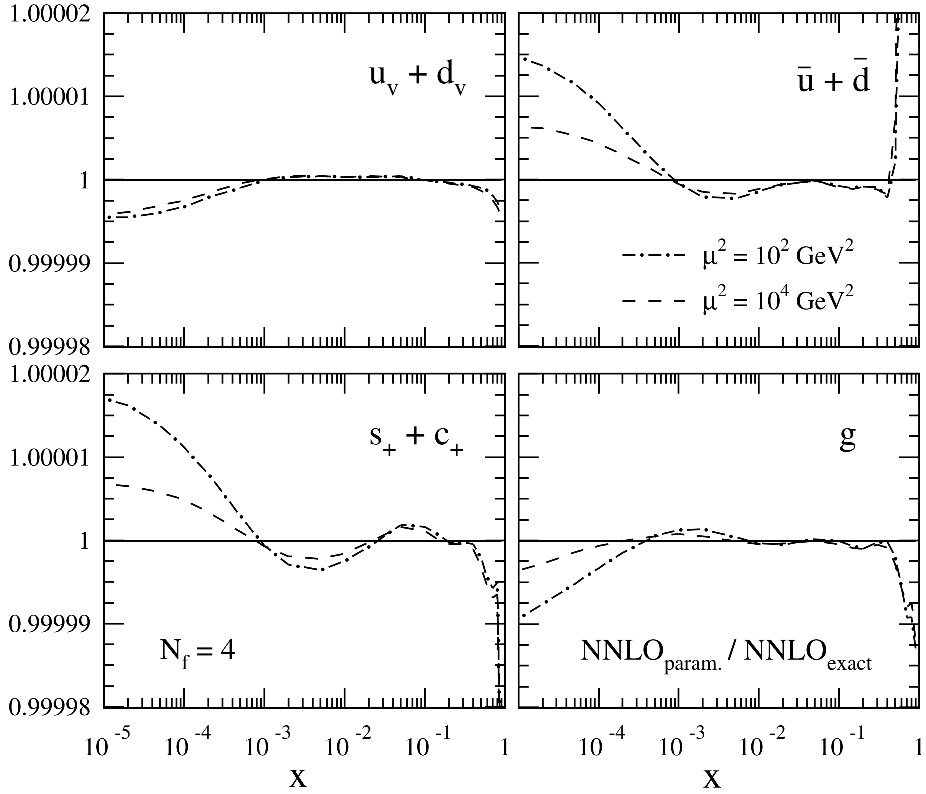


Fig. 10: Relative effects of using the parametrized three-loop splitting functions of Refs. [6, 7], instead of the exact expressions from the same source, on the NNLO evolution for the input (4.16) – (4.18) at two representative values of $\mu = \mu_r = \mu_f$.

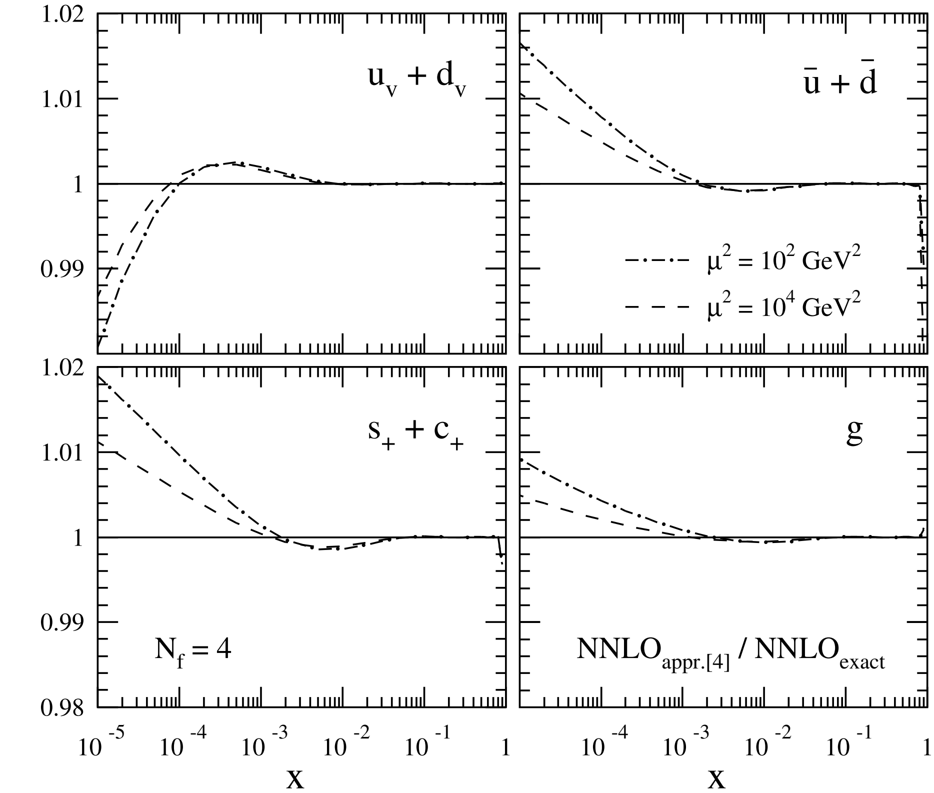


Fig. 11: Relative errors made by using the previous average approximations [61] for the three-loop splitting functions (used, e.g., in Refs. [51, 69]) instead of the full results [6, 7], on the NNLO evolution of the input (4.16) – (4.18) at $\mu_r = \mu_f$.

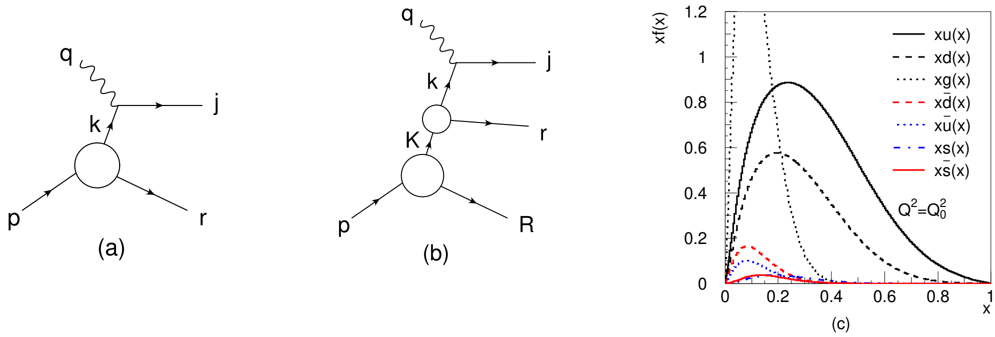


Fig. 12: Probing (a) a valence parton in the proton and (b) a sea parton in a hadronic fluctuation (letters are four-momenta) resulting in (c) parton distributions at the starting scale Q_0^2.

W and Higgs production. The effect of replacing the approximate results by the full splitting functions [6, 7] is illustrated in Figure 11. Especially at scales relevant to the above-mentioned processes, the previous approximations introduce an error of less than 0.2% for $x \gtrsim 10^{-3}$, and less than 1% even down to $x \simeq 10^{-5}$. Consequently the splitting-function approximations used for the evolution the parton distributions of Refs. [51, 69] are confirmed to a sufficient accuracy for high-scale processes at the LHC.

The unchanged unpolarized LO and NLO reference tables of Ref. [59] are not repeated here. Note that the one digit of the first (FFN) α_s value was mistyped in the header of Table 1 in that report[5], the correct value can be found in Table 3 below.

5 Non-perturbative x-shape of PDFs[6]

The x-shape of parton density functions at a low scale Q_0^2 is due to the dynamics of the bound state proton and is hence an unsolved problem of non-perturbative QCD. Usually this is described by parameterizations of data using more or less arbitrary functional forms. More understanding can be obtained by a recently developed physical model [70], which is phenomenologically successful in describing data.

The model gives the four-momentum k of a single probed valence parton (Fig. 12a) by assuming that, in the nucleon rest frame, the shape of the momentum distribution for a parton of type i and mass m_i can be taken as a Gaussian $f_i(k) = N(\sigma_i, m_i) \exp\left\{-\left[(k_0 - m_i)^2 + k_x^2 + k_y^2 + k_z^2\right]/2\sigma_i^2\right\}$, which may be motivated as a result of the many interactions binding the parton in the nucleon. The width of the distribution should be of order hundred MeV from the Heisenberg uncertainty relation applied to the nucleon size, *i.e.* $\sigma_i = 1/d_N$. The momentum fraction x of the parton is then defined as the light-cone fraction $x = k_+/p_+$ and is therefore invariant under longitudinal boosts (e.g. to the infinite momentum frame). Constraints are imposed on the final-state momenta to obtain a kinematically allowed final state, which also ensures that $0 < x < 1$ and $f_i(x) \to 0$ for $x \to 1$.

The sea partons are obtained using a hadronic basis for the non-perturbative dynamics of the bound state proton and considering hadronic fluctuations

$$|p\rangle = \alpha_0|p_0\rangle + \alpha_{p\pi^0}|p\pi^0\rangle + \alpha_{n\pi^+}|n\pi^+\rangle + \ldots + \alpha_{\Lambda K}|\Lambda K^+\rangle + \ldots \tag{5.21}$$

Probing a parton i in a hadron H of a baryon-meson fluctuation $|BM\rangle$ (Fig. 12b) gives a sea parton with light-cone fraction $x = x_H\, x_i$ of the target proton. The momentum of the probed hadron is given by a similar Gaussian, but with a separate width parameter σ_H. Also here, kinematic constraints ensure physically allowed final states.

[5]We thank H. Böttcher and J. Blümlein for pointing this out to us.

[6]Contributing author: G. Ingelman

Table 2: Reference results for the $N_\mathrm{f} = 4$ next-next-to-leading-order evolution for the initial conditions (4.16) – (4.18). The corresponding value of the strong coupling is $\alpha_\mathrm{s}(\mu_\mathrm{r}^2 = 10^4\ \mathrm{GeV}^2) = 0.110141$. The valence distributions s_v and c_v are equal for the input (4.17). The notation is explained below Eq. (4.17) and in the paragraph below Eq. (4.20).

NNLO, $N_\mathrm{f} = 4\,,\ \mu_\mathrm{f}^2 = 10^4\ \mathrm{GeV}^2$								
x	xu_v	xd_v	xL_-	$2xL_+$	xs_v	xs_+	xc_+	xg
$\mu_\mathrm{r}^2 = \mu_\mathrm{f}^2$								
10^{-7}	1.5287^{-4}	1.0244^{-4}	5.7018^{-6}	1.3190^{+2}	3.1437^{-5}	6.4877^{+1}	6.4161^{+1}	9.9763^{+2}
10^{-6}	6.9176^{-4}	4.4284^{-4}	2.5410^{-5}	6.8499^{+1}	9.4279^{-5}	3.3397^{+1}	3.2828^{+1}	4.9124^{+2}
10^{-5}	3.0981^{-3}	1.8974^{-3}	1.0719^{-4}	3.3471^{+1}	2.2790^{-4}	1.6059^{+1}	1.5607^{+1}	2.2297^{+2}
10^{-4}	1.3722^{-2}	8.1019^{-3}	4.2558^{-4}	1.5204^{+1}	3.6644^{-4}	7.0670^{+0}	6.7097^{+0}	9.0668^{+1}
10^{-3}	5.9160^{-2}	3.4050^{-2}	1.6008^{-3}	6.3230^{+0}	1.4479^{-4}	2.7474^{+0}	2.4704^{+0}	3.1349^{+1}
10^{-2}	2.3078^{-1}	1.2919^{-1}	5.5688^{-3}	2.2752^{+0}	-5.7311^{-4}	8.5502^{-1}	6.6623^{-1}	8.1381^{+0}
0.1	5.5177^{-1}	2.7165^{-1}	1.0023^{-2}	3.9019^{-1}	-3.0627^{-4}	1.1386^{-1}	5.9773^{-2}	9.0563^{-1}
0.3	3.5071^{-1}	1.3025^{-1}	3.0098^{-3}	3.5358^{-2}	-3.1891^{-5}	9.0480^{-3}	3.3061^{-3}	8.4186^{-2}
0.5	1.2117^{-1}	3.1528^{-2}	3.7742^{-4}	2.3867^{-3}	-2.7215^{-6}	5.7965^{-4}	1.7170^{-4}	8.1126^{-3}
0.7	2.0077^{-2}	3.0886^{-3}	1.3434^{-5}	5.4244^{-5}	-1.0106^{-7}	1.2936^{-5}	3.5304^{-6}	3.8948^{-4}
0.9	3.5111^{-4}	1.7783^{-5}	8.651^{-9}	2.695^{-8}	-1.476^{-10}	7.132^{-9}	2.990^{-9}	1.2136^{-6}
$\mu_\mathrm{r}^2 = 2\,\mu_\mathrm{f}^2$								
10^{-7}	1.3416^{-4}	8.7497^{-5}	4.9751^{-6}	1.3020^{+2}	2.1524^{-5}	6.4025^{+1}	6.3308^{+1}	1.0210^{+3}
10^{-6}	6.2804^{-4}	3.9406^{-4}	2.2443^{-5}	6.6914^{+1}	6.5149^{-5}	3.2602^{+1}	3.2032^{+1}	4.9626^{+2}
10^{-5}	2.9032^{-3}	1.7575^{-3}	9.6205^{-5}	3.2497^{+1}	1.5858^{-4}	1.5570^{+1}	1.5118^{+1}	2.2307^{+2}
10^{-4}	1.3206^{-2}	7.7673^{-3}	3.9093^{-4}	1.4751^{+1}	2.5665^{-4}	6.8388^{+0}	6.4807^{+0}	9.0162^{+1}
10^{-3}	5.8047^{-2}	3.3434^{-2}	1.5180^{-3}	6.1703^{+0}	1.0388^{-4}	2.6695^{+0}	2.3917^{+0}	3.1114^{+1}
10^{-2}	2.2930^{-1}	1.2857^{-1}	5.4626^{-3}	2.2492^{+0}	-3.9979^{-4}	8.4058^{-1}	6.5087^{-1}	8.0993^{+0}
0.1	5.5428^{-1}	2.7326^{-1}	1.0072^{-2}	3.9297^{-1}	-2.1594^{-4}	1.1439^{-1}	5.9713^{-2}	9.0851^{-1}
0.3	3.5501^{-1}	1.3205^{-1}	3.0557^{-3}	3.6008^{-2}	-2.2632^{-5}	9.2227^{-3}	3.3771^{-3}	8.5022^{-2}
0.5	1.2340^{-1}	3.2166^{-2}	3.8590^{-4}	2.4459^{-3}	-1.9420^{-6}	5.9487^{-4}	1.7699^{-4}	8.2293^{-3}
0.7	2.0597^{-2}	3.1751^{-3}	1.3849^{-5}	5.5722^{-5}	-7.2616^{-8}	1.3244^{-5}	3.5361^{-6}	3.9687^{-4}
0.9	3.6527^{-4}	1.8544^{-5}	9.050^{-9}	2.663^{-8}	-1.075^{-10}	6.713^{-9}	2.377^{-9}	1.2489^{-6}
$\mu_\mathrm{r}^2 = 1/2\,\mu_\mathrm{f}^2$								
10^{-7}	1.7912^{-4}	1.2521^{-4}	6.4933^{-6}_{*}	1.2714^{+2}	4.9649^{-5}	6.2498^{+1}	6.1784^{+1}	9.2473^{+2}
10^{-6}	7.7377^{-4}	5.1222^{-4}	2.8719^{-5}	6.7701^{+1}	1.4743^{-4}	3.2999^{+1}	3.2432^{+1}	4.6863^{+2}
10^{-5}	3.3184^{-3}	2.0760^{-3}	1.1977^{-4}	3.3644^{+1}	3.5445^{-4}	1.6147^{+1}	1.5696^{+1}	2.1747^{+2}
10^{-4}	1.4184^{-2}	8.4455^{-3}	4.6630^{-4}	1.5408^{+1}	5.6829^{-4}	7.1705^{+0}	6.8139^{+0}	8.9820^{+1}_{*}
10^{-3}	5.9793^{-2}	3.4418^{-2}	1.6996^{-3}	6.4042^{+0}	2.2278^{-4}	2.7892^{+0}	2.5128^{+0}	3.1336^{+1}
10^{-2}	2.3106^{-1}	1.2914^{-1}	5.7016^{-3}	2.2876^{+0}	-8.9125^{-4}	8.6205^{-1}	6.7377^{-1}	8.1589^{+0}
0.1	5.5039^{-1}	2.7075^{-1}	1.0031^{-2}	3.8850^{-1}	-4.7466^{-4}	1.1332^{-1}	5.9489^{-2}	9.0795^{-1}
0.3	3.4890^{-1}	1.2949^{-1}	2.9943^{-3}	3.5090^{-2}	-4.9304^{-5}	8.9667^{-3}	3.2670^{-3}	8.4309^{-2}
0.5	1.2026^{-1}	3.1269^{-2}	3.7428^{-4}	2.3729^{-3}	-4.1981^{-6}	5.7783^{-4}	1.7390^{-4}	8.1099^{-3}_{*}
0.7	1.9867^{-2}	3.0534^{-3}	1.3273^{-5}	5.4635^{-5}	-1.5541^{-7}	1.3275^{-5}	3.9930^{-6}	3.8824^{-4}
0.9	3.4524^{-4}	1.7466^{-5}	8.489^{-9}	3.030^{-8}	-2.255^{-10}	8.863^{-9}	4.803^{-9}	1.2026^{-6}

Table 3: As Table 2, but for the variable-N_f evolution using the flavour matching conditions of Ref. [65, 67, 68]. The corresponding values for the strong coupling $\alpha_\mathrm{s}(\mu_\mathrm{r}^2 = 10^4\ \mathrm{GeV}^2)$ are given by 0.115818, 0.115605 and 0.115410 for $\mu_\mathrm{r}^2/\mu_\mathrm{f}^2 = 0.5$, 1 and 2, respectively. For brevity the small, but non-vanishing valence distributions s_v, c_v and b_v are not displayed.

NNLO, $N_\mathrm{f} = 3 \ldots 5$, $\mu_\mathrm{f}^2 = 10^4$ GeV^2								
x	xu_v	xd_v	xL_-	$2xL_+$	xs_+	xc_+	xb_+	xg
$\mu_\mathrm{r}^2 = \mu_\mathrm{f}^2$								
10^{-7}	1.5978^{-4}	1.0699^{-5}	6.0090^{-6}	1.3916^{+2}	6.8509^{+1}	6.6929^{+1}	5.7438^{+1}	9.9694^{+3}
10^{-6}	7.1787^{-4}	4.5929^{-4}	2.6569^{-5}	7.1710^{+1}	3.5003^{+1}	3.3849^{+1}	2.8332^{+1}	4.8817^{+2}
10^{-5}	3.1907^{-3}	1.9532^{-3}	1.1116^{-4}	3.4732^{+1}	1.6690^{+1}	1.5875^{+1}	1.2896^{+1}	2.2012^{+2}
10^{-4}	1.4023^{-2}	8.2749^{-3}	4.3744^{-4}	1.5617^{+1}	7.2747^{+0}	6.7244^{+0}	5.2597^{+0}	8.8804^{+1}
10^{-3}	6.0019^{-2}	3.4519^{-2}	1.6296^{-3}	6.4173^{+0}	2.7954^{+0}	2.4494^{+0}	1.8139^{+0}	3.0404^{+1}
10^{-2}	2.3244^{-1}	1.3000^{-1}	5.6100^{-3}	2.2778^{+0}	8.5749^{-1}	6.6746^{-1}	4.5073^{-1}	7.7912^{+0}
0.1	5.4993^{-1}	2.7035^{-1}	9.9596^{-3}	3.8526^{-1}	1.1230^{-1}	6.4466^{-2}	3.7280^{-2}	8.5266^{-1}
0.3	3.4622^{-1}	1.2833^{-1}	2.9572^{-3}	3.4600^{-2}	8.8410^{-3}	4.0134^{-3}	2.1047^{-3}	7.8898^{-2}
0.5	1.1868^{-1}	3.0811^{-2}	3.6760^{-4}	2.3198^{-3}	5.6309^{-4}	2.3752^{-4}	1.2004^{-4}	7.6398^{-3}
0.7	1.9486^{-2}	2.9901^{-3}	1.2957^{-5}	5.2352^{-5}	1.2504^{-5}	5.6038^{-6}	2.8888^{-6}	3.7080^{-4}
0.9	3.3522^{-4}	1.6933^{-5}	8.209^{-9}	2.574^{-8}	6.856^{-9}	4.337^{-9}	2.679^{-9}	1.1721^{-6}
$\mu_\mathrm{r}^2 = 2\,\mu_\mathrm{f}^2$								
10^{-7}	1.3950^{-4}	9.0954^{-5}	5.2113^{-6}	1.3549^{+2}	6.6672^{+1}	6.5348^{+1}	5.6851^{+1}	1.0084^{+3}
10^{-6}	6.4865^{-4}	4.0691^{-4}	2.3344^{-5}	6.9214^{+1}	3.3753^{+1}	3.2772^{+1}	2.7818^{+1}	4.8816^{+2}
10^{-5}	2.9777^{-3}	1.8020^{-3}	9.9329^{-5}	3.3385^{+1}	1.6015^{+1}	1.5306^{+1}	1.2601^{+1}	2.1838^{+2}
10^{-4}	1.3452^{-2}	7.9078^{-3}	4.0036^{-4}	1.5035^{+1}	6.9818^{+0}	6.4880^{+0}	5.1327^{+0}	8.7550^{+1}
10^{-3}	5.8746^{-2}	3.3815^{-2}	1.5411^{-3}	6.2321^{+0}	2.7012^{+0}	2.3747^{+0}	1.7742^{+0}	3.0060^{+1}
10^{-2}	2.3063^{-1}	1.2923^{-1}	5.4954^{-3}	2.2490^{+0}	8.4141^{-1}	6.5083^{-1}	4.4354^{-1}	7.7495^{+0}
0.1	5.5279^{-1}	2.7222^{-1}	1.0021^{-2}	3.8897^{-1}	1.1312^{-1}	6.2917^{-2}	3.7048^{-2}	8.5897^{-1}
0.3	3.5141^{-1}	1.3051^{-1}	3.0134^{-3}	3.5398^{-2}	9.0559^{-3}	3.8727^{-3}	2.0993^{-3}	8.0226^{-2}
0.5	1.2140^{-1}	3.1590^{-2}	3.7799^{-4}	2.3919^{-3}	5.8148^{-4}	2.2376^{-4}	1.1918^{-4}	7.8098^{-3}
0.7	2.0120^{-2}	3.0955^{-3}	1.3462^{-5}	5.4194^{-5}	1.2896^{-5}	5.0329^{-6}	2.8153^{-6}	3.8099^{-4}
0.9	3.5230^{-4}	1.7849^{-5}	8.687^{-9}	2.568^{-8}	6.513^{-9}	3.390^{-9}	2.407^{-9}	1.2188^{-6}
$\mu_\mathrm{r}^2 = 1/2\,\mu_\mathrm{f}^2$								
10^{-7}	1.8906^{-4}	1.3200^{-4}	6.9268^{-6}	1.3739^{+2}	6.7627^{+1}	6.5548^{+1}	5.5295^{+1}	9.4403^{+2}
10^{-6}	8.1001^{-4}	5.3574^{-4}	3.0345^{-5}	7.2374^{+1}	3.5337^{+1}	3.3846^{+1}	2.7870^{+1}	4.7444^{+2}
10^{-5}	3.4428^{-3}	2.1524^{-3}	1.2531^{-4}	3.5529^{+1}	1.7091^{+1}	1.6065^{+1}	1.2883^{+1}	2.1802^{+2}
10^{-4}	1.4580^{-2}	8.6744^{-3}	4.8276^{-4}	1.6042^{+1}	7.4886^{+0}	6.8276^{+0}	5.3044^{+0}	8.9013^{+1}
10^{-3}	6.0912^{-2}	3.5030^{-2}	1.7393^{-3}	6.5544^{+0}	2.8656^{+0}	2.4802^{+0}	1.8362^{+0}	3.0617^{+1}
10^{-2}	2.3327^{-1}	1.3022^{-1}	5.7588^{-3}	2.2949^{+0}	8.6723^{-1}	6.7688^{-1}	4.5597^{-1}	7.8243^{+0}_{*}
0.1	5.4798^{-1}	2.6905^{-1}	9.9470^{-3}	3.8192^{-1}	1.1124^{-1}	6.7091^{-2}	3.7698^{-2}	8.4908^{-1}
0.3	3.4291^{-1}	1.2693^{-1}	2.9239^{-3}	3.4069^{-2}	8.6867^{-3}	4.3924^{-3}	2.1435^{-3}	7.8109^{-2}
0.5	1.1694^{-1}	3.0310^{-2}	3.6112^{-4}	2.2828^{-3}	5.5537^{-4}	2.7744^{-4}	1.2416^{-4}	7.5371^{-3}
0.7	1.9076^{-2}	2.9217^{-3}	1.2635^{-5}	5.2061^{-5}	1.2677^{-5}	7.2083^{-6}	3.0908^{-6}	3.6441^{-4}
0.9	3.2404^{-4}	1.6333^{-5}	7.900^{-9}	2.850^{-8}	8.407^{-9}	6.795^{-9}	3.205^{-9}	1.1411^{-6}

Table 4: Reference results for the $N_{\rm f} = 4$ (FFN) and the variable-$N_{\rm f}$ (VFN) polarized leading-order evolution of the initial distributions (4.20), shown together with these boundary conditions. The respective values for $\alpha_{\rm s}(\mu_{\rm r}^2 = \mu_{\rm f}^2 = 10^4\ {\rm GeV}^2)$ read 0.117574 (FFN) and 0.122306 (VFN). The notation is the same as for the unpolarized case.

x	xu_v	$-xd_v$	$-xL_-$	$-2xL_+$	xs_+	xc_+	xb_+	xg
Pol. input, $\mu_{\rm f}^2 = 2\ {\rm GeV}^2$								
10^{-7}	1.6366^{-5}	6.2946^{-6}	7.9433^{-5}	1.5887^{-3}	-3.9716^{-4}	0.0^{+0}	0.0^{+0}	4.7434^{-4}
10^{-6}	8.2024^{-5}	3.1548^{-5}	1.5849^{-4}	3.1698^{-3}	-7.9244^{-4}	0.0^{+0}	0.0^{+0}	1.5000^{-3}
10^{-5}	4.1110^{-4}	1.5811^{-4}	3.1621^{-4}	6.3241^{-3}	-1.5810^{-3}	0.0^{+0}	0.0^{+0}	4.7432^{-3}
10^{-4}	2.0604^{-3}	7.9245^{-4}	6.3052^{-4}	1.2610^{-2}	-3.1526^{-3}	0.0^{+0}	0.0^{+0}	1.4993^{-2}
10^{-3}	1.0326^{-2}	3.9716^{-3}	1.2501^{-3}	2.5003^{-2}	-6.2507^{-3}	0.0^{+0}	0.0^{+0}	4.7197^{-2}
10^{-2}	5.1723^{-2}	1.9886^{-2}	2.3412^{-3}	4.6825^{-2}	-1.1706^{-2}	0.0^{+0}	0.0^{+0}	1.4265^{-1}
0.1	2.4582^{-1}	9.1636^{-2}	2.3972^{-3}	4.7943^{-2}	-1.1986^{-2}	0.0^{+0}	0.0^{+0}	2.8009^{-1}
0.3	3.6473^{-1}	1.1370^{-1}	5.7388^{-4}	1.1478^{-2}	-2.8694^{-3}	0.0^{+0}	0.0^{+0}	1.3808^{-1}
0.5	2.5008^{-1}	5.7710^{-2}	6.3457^{-5}	1.2691^{-3}	-3.1729^{-4}	0.0^{+0}	0.0^{+0}	3.3146^{-2}
0.7	8.4769^{-2}	1.1990^{-2}	1.9651^{-6}	3.9301^{-5}	-9.8254^{-6}	0.0^{+0}	0.0^{+0}	3.0496^{-3}
0.9	4.4680^{-3}	2.1365^{-4}	9.689^{-10}	1.9378^{-8}	-4.8444^{-9}	0.0^{+0}	0.0^{+0}	1.4230^{-5}
LO, $N_{\rm f} = 4\,,\ \mu_{\rm f}^2 = 10^4\ {\rm GeV}^2$								
10^{-7}	4.8350_*^{-5}	1.8556^{-5}	1.0385^{-4}	3.5124^{-3}	-1.2370^{-3}	-7.1774^{-4}	0.0^{+0}	1.4116^{-2}
10^{-6}	2.3504^{-4}	9.0090^{-5}	2.0700^{-4}	7.7716^{-3}	-2.8508^{-3}	-1.8158^{-3}	0.0^{+0}	4.2163^{-2}
10^{-5}	1.1220^{-3}	4.2916^{-4}	4.1147^{-4}	1.6007^{-2}	-5.9463^{-3}	-3.8889^{-3}	0.0^{+0}	1.0922^{-1}
10^{-4}	5.1990^{-3}	1.9818^{-3}	8.0948^{-4}	2.8757^{-2}	-1.0331^{-2}	-6.2836^{-3}	0.0^{+0}	2.4069^{-1}
10^{-3}	2.2900^{-2}	8.6763^{-3}	1.5309^{-3}	4.0166^{-2}	-1.2428^{-2}	-4.7739^{-3}	0.0^{+0}	4.2181^{-1}
10^{-2}	9.1489^{-2}	3.4200^{-2}	2.4502^{-3}	3.3928^{-2}	-4.7126^{-3}	7.5385^{-3}	0.0^{+0}	4.9485^{-1}
0.1	2.6494^{-1}	9.1898^{-2}	1.5309^{-3}	8.5427^{-3}	3.3830^{-3}	1.1037^{-2}	0.0^{+0}	2.0503^{-1}
0.3	2.2668^{-1}	6.2946^{-2}	2.1104^{-4}	6.6698^{-4}	7.2173^{-4}	1.7769^{-3}	0.0^{+0}	3.3980^{-2}
0.5	9.7647^{-2}	1.9652^{-2}	1.4789^{-5}	-1.8850^{-5}	8.3371^{-5}	1.5732^{-4}	0.0^{+0}	4.3802^{-3}
0.7	1.9545^{-2}	2.3809^{-3}	2.7279^{-7}	-4.1807^{-6}	3.4543^{-6}	4.8183^{-6}	0.0^{+0}	2.6355^{-4}
0.9	4.1768^{-4}	1.7059^{-5}	5.494^{-11}	-7.6712^{-9}	4.1103^{-9}	4.3850^{-9}	0.0^{+0}	9.8421^{-7}
LO, $N_{\rm f} = 3 \ldots 5\,,\ \mu_{\rm f}^2 = 10^4\ {\rm GeV}^2$								
10^{-7}	4.9026^{-5}	1.8815^{-5}	1.0422^{-4}	3.5315^{-3}	-1.2447^{-3}	-7.2356^{-4}	-6.2276^{-4}	1.3726^{-2}
10^{-6}	2.3818^{-4}	9.1286^{-5}	2.0774^{-4}	7.8108^{-3}	-2.8667^{-3}	-1.8280^{-3}	-1.5301^{-3}	4.1011^{-2}
10^{-5}	1.1359^{-3}	4.3445^{-4}	4.1289^{-4}	1.6070^{-2}	-5.9705^{-3}	-3.9060^{-3}	-3.1196^{-3}	1.0615^{-1}
10^{-4}	5.2567^{-3}	2.0035^{-3}	8.1206^{-4}	2.8811^{-2}	-1.0345^{-2}	-6.2849^{-3}	-4.5871^{-3}	2.3343^{-1}
10^{-3}	2.3109^{-2}	8.7537^{-3}	1.5345^{-3}	4.0125^{-2}	-1.2390^{-2}	-4.7174^{-3}	-2.4822^{-3}	4.0743^{-1}
10^{-2}	9.2035^{-2}	3.4391^{-2}	2.4501^{-3}	3.3804^{-2}	-4.6512^{-3}	7.5994^{-3}	6.4665^{-3}	4.7445^{-1}
0.1	2.6478^{-1}	9.1762^{-2}	1.5206^{-3}	8.5181^{-3}	3.3438^{-3}	1.0947^{-2}	6.5223^{-3}	1.9402^{-1}
0.3	2.2495^{-1}	6.2376^{-2}	2.0811^{-4}	6.6195^{-4}	7.0957^{-4}	1.7501^{-3}	9.2045^{-4}	3.1960^{-2}
0.5	9.6318^{-2}	1.9353^{-2}	1.4496^{-5}	-1.8549^{-5}	8.1756^{-5}	1.5424^{-4}	7.8577^{-5}	4.1226^{-3}
0.7	1.9147^{-2}	2.3281^{-3}	2.6556^{-7}	-4.0936^{-6}	3.3746^{-6}	4.7024^{-6}	2.4901^{-6}	2.4888^{-4}
0.9	4.0430^{-4}	1.6480^{-5}	5.285^{-11}	-7.4351^{-9}	3.9818^{-9}	4.2460^{-9}	2.6319^{-9}	9.2939^{-7}

Table 5: Reference results for the polarized next-to-leading-order polarized evolution of the initial distributions (4.20) with $N_\mathrm{f} = 4$ quark flavours. The corresponding value of the strong coupling is $\alpha_\mathrm{s}(\mu_\mathrm{r}^2 = 10^4\ \mathrm{GeV}^2) = 0.110902$. As in the leading-order case, the valence distributions s_v and c_v vanish for the input (4.20).

Pol. NLO, $N_\mathrm{f} = 4$, $\mu_\mathrm{f}^2 = 10^4\ \mathrm{GeV}^2$							
x	xu_v	xd_v	xL_-	$2xL_+$	xs_+	xc_+	xg
$\mu_\mathrm{r}^2 = \mu_\mathrm{f}^2$							
10^{-7}	6.7336^{-5}	-2.5747^{-5}	-1.1434^{-4}	-5.2002^{-3}	-2.0528^{-3}	-1.5034^{-3}	2.6955^{-2}
10^{-6}	3.1280^{-4}	-1.1938^{-4}	-2.3497^{-4}	-1.0725^{-2}	-4.2774^{-3}	-3.1845^{-3}	6.5928^{-2}
10^{-5}	1.4180^{-3}	-5.3982^{-4}	-4.8579^{-4}	-1.9994^{-2}	-7.8594^{-3}	-5.6970^{-3}	1.4414^{-1}
10^{-4}	6.2085^{-3}	-2.3546^{-3}	-9.8473^{-4}	-3.1788^{-2}	-1.1749^{-2}	-7.5376^{-3}	2.7537^{-1}
10^{-3}	2.5741^{-2}	-9.7004^{-3}	-1.8276^{-3}	-3.8222^{-2}	-1.1427^{-2}	-3.6138^{-3}	4.3388^{-1}
10^{-2}	9.6288^{-2}	-3.5778^{-2}	-2.6427^{-3}	-2.6437^{-2}	-1.2328^{-3}	1.0869^{-2}	4.8281^{-1}
0.1	2.5843^{-1}	-8.9093^{-2}	-1.4593^{-3}	-7.5546^{-3}	3.4258^{-3}	1.0639^{-2}	2.0096^{-1}
0.3	2.1248^{-1}	-5.8641^{-2}	-1.9269^{-4}	-1.2210^{-3}	3.5155^{-4}	1.3138^{-3}	3.4126^{-2}
0.5	8.9180^{-2}	-1.7817^{-2}	-1.3125^{-5}	-9.1573^{-5}	1.9823^{-5}	8.5435^{-5}	4.5803^{-3}
0.7	1.7300^{-2}	-2.0885^{-3}	-2.3388^{-7}	-1.9691^{-6}	1.8480^{-7}	1.3541^{-6}	2.9526^{-4}
0.9	3.4726^{-4}	-1.4028^{-5}	-4.407^{-11}	-4.247^{-9}	-1.903^{-9}	-1.683^{-9}	1.2520^{-6}
$\mu_\mathrm{r}^2 = 2\,\mu_\mathrm{f}^2$							
10^{-7}	6.1781^{-5}	-2.3641^{-5}	-1.1137^{-4}	-4.6947^{-3}	-1.8092^{-3}	-1.2695^{-3}	2.2530^{-2}
10^{-6}	2.8974^{-4}	-1.1068^{-4}	-2.2755^{-4}	-9.8528^{-3}	-3.8580^{-3}	-2.7838^{-3}	5.7272^{-2}_{*}
10^{-5}	1.3281^{-3}	-5.0612^{-4}	-4.6740^{-4}	-1.8799^{-2}	-7.2908^{-3}	-5.1629^{-3}	1.2975^{-1}
10^{-4}	5.8891^{-3}	-2.2361^{-3}	-9.4412^{-4}	-3.0787^{-2}	-1.1292^{-2}	-7.1363^{-3}	2.5644^{-1}
10^{-3}	2.4777^{-2}	-9.3502^{-3}	-1.7632^{-3}	-3.8610^{-2}	-1.1658^{-2}	-3.9083^{-3}	4.1725^{-1}
10^{-2}	9.4371^{-2}	-3.5129^{-2}	-2.6087^{-3}	-2.8767^{-2}	-2.3430^{-3}_{*}	9.7922^{-3}_{*}	4.7804^{-1}
0.1	2.6008^{-1}	-8.9915^{-2}	-1.4923^{-3}	-8.3806^{-3}	3.1932^{-3}	1.0585^{-2}	2.0495^{-1}
0.3	2.1837^{-1}	-6.0497^{-2}	-2.0143^{-4}	-1.2157^{-3}	3.9810^{-4}	1.4042^{-3}	3.5366^{-2}
0.5	9.3169^{-2}	-1.8699^{-2}	-1.3954^{-5}	-7.9331^{-5}	3.0091^{-5}	9.9849^{-5}	4.7690^{-3}
0.7	1.8423^{-2}	-2.2357^{-3}	-2.5360^{-7}	-1.0062^{-6}	7.6483^{-7}	2.0328^{-6}	3.0796^{-4}
0.9	3.8293^{-4}	-1.5559^{-5}	-4.952^{-11}	-1.955^{-9}	-7.298^{-10}	-4.822^{-10}	1.3247^{-6}
$\mu_\mathrm{r}^2 = 1/2\,\mu_\mathrm{f}^2$							
10^{-7}	7.4443^{-5}	-2.8435^{-5}	-1.1815^{-4}	-5.7829^{-3}	-2.3341^{-3}	-1.7739^{-3}	3.2071^{-2}
10^{-6}	3.4143^{-4}	-1.3016^{-4}	-2.4482^{-4}	-1.1668^{-2}	-4.7305^{-3}	-3.6168^{-3}	7.5123^{-2}
10^{-5}	1.5256^{-3}	-5.8002^{-4}	-5.1085^{-4}	-2.1193^{-2}	-8.4295^{-3}	-6.2295^{-3}_{*}	1.5788^{-1}
10^{-4}	6.5726^{-3}	-2.4891^{-3}	-1.0409^{-3}	-3.2697^{-2}	-1.2166^{-2}	-7.8952^{-3}	2.9079^{-1}
10^{-3}	2.6766^{-2}	-1.0070^{-2}	-1.9171^{-3}	-3.7730^{-2}	-1.1160^{-2}	-3.2890^{-3}_{*}	4.4380^{-1}
10^{-2}	9.8073^{-2}	-3.6370^{-2}	-2.6942^{-3}	-2.4056^{-2}	-1.2354^{-4}_{*}	1.1929^{-2}	4.8272^{-1}
0.1	2.5628^{-1}	-8.8133^{-2}	-1.4304^{-3}	-6.9572^{-3}	3.5561^{-3}	1.0604^{-2}	1.9831^{-1}
0.3	2.0709^{-1}	-5.6988^{-2}	-1.8541^{-4}	-1.3308^{-3}	2.5993^{-4}	1.1855^{-3}	3.3524^{-2}
0.5	8.5835^{-2}	-1.7089^{-2}	-1.2463^{-5}	-1.1920^{-4}	2.6972^{-6}_{*}	6.4995^{-5}	4.5044^{-3}
0.7	1.6405^{-2}	-1.9723^{-3}	-2.1859^{-7}_{*}	-3.6817^{-6}	-7.4795^{-7}_{*}	3.4496^{-7}	2.9100^{-4}
0.9	3.2011^{-4}	-1.2870^{-5}	-4.000^{-11}	-8.173^{-9}	-3.886^{-9}	-3.686^{-9}	1.2230^{-6}

Table 6: As Table 5, but for the variable-N_f evolution using Eqs. (4.16), (4.17) and (4.20). The corresponding values for the strong coupling $\alpha_\mathrm{s}(\mu_\mathrm{r}^2 = 10^4\ \mathrm{GeV}^2)$ are given by 0.116461, 0.116032 and 0.115663 for $\mu_\mathrm{r}^2/\mu_\mathrm{f}^2 = 0.5$, 1 and 2, respectively.

Pol. NLO, $N_\mathrm{f} = 3\ \ldots 5\,,\ \mu_\mathrm{f}^2 = 10^4\ \mathrm{GeV}^2$								
x	xu_v	$-xd_v$	$-xL_-$	$-2xL_+$	xs_+	xc_+	xb_+	xg
$\mu_\mathrm{r}^2 = \mu_\mathrm{f}^2$								
10^{-7}	6.8787^{-5}_{*}	2.6297^{-5}	1.1496^{-4}	5.2176^{-3}	-2.0592^{-3}	-1.5076^{-3}	-1.2411^{-3}	2.5681^{-2}
10^{-6}	3.1881^{-4}	1.2165^{-4}	2.3638^{-4}	1.0770^{-2}	-4.2953^{-3}	-3.1979^{-3}	-2.4951^{-3}	6.3021^{-2}
10^{-5}	1.4413^{-3}	5.4856^{-4}	4.8893^{-4}	2.0077^{-2}	-7.8934^{-3}	-5.7228^{-3}	-4.1488^{-3}	1.3809^{-1}
10^{-4}	6.2902^{-3}	2.3849^{-3}	9.9100^{-4}	3.1883^{-2}	-1.1785^{-2}	-7.5596^{-3}	-4.8420^{-3}	2.6411^{-1}
10^{-3}	2.5980^{-2}	9.7872^{-3}	1.8364^{-3}	3.8224^{-2}	-1.1416^{-2}	-3.5879^{-3}	-1.1723^{-3}	4.1601^{-1}
10^{-2}	9.6750^{-2}	3.5935^{-2}	2.6452^{-3}	2.6306^{-2}	-1.1774^{-3}	1.0917^{-2}	8.1196^{-3}	4.6178^{-1}
0.1	2.5807^{-1}	8.8905^{-2}	1.4509^{-3}	7.4778^{-3}	3.4207^{-3}	1.0591^{-2}	6.1480^{-3}	1.9143^{-1}
0.3	2.1104^{-1}	5.8186^{-2}	1.9054^{-4}	1.2026^{-3}	3.4999^{-4}	1.3015^{-3}	7.2795^{-4}	3.2621^{-2}
0.5	8.8199^{-2}	1.7601^{-2}	1.2924^{-5}	8.9668^{-5}	1.9771^{-5}	8.4378^{-5}	5.2125^{-5}	4.4207^{-3}
0.7	1.7027^{-2}	2.0531^{-3}	2.2921^{-7}	1.9243^{-6}	1.8384^{-7}	1.3298^{-6}	1.2157^{-6}	2.8887^{-4}
0.9	3.3898^{-4}	1.3676^{-5}	4.284^{-11}	4.260^{-9}	-1.916^{-9}	-1.701^{-9}	-7.492^{-11}	1.2435^{-6}
$\mu_\mathrm{r}^2 = 2\,\mu_\mathrm{f}^2$								
10^{-7}	6.2819^{-5}_{*}	2.4035^{-5}	1.1180^{-4}	4.6896^{-3}	-1.8050^{-3}	-1.2637^{-3}	-1.0544^{-3}	2.1305^{-2}
10^{-6}	2.9408^{-4}	1.1232^{-4}	2.2855^{-4}	9.8538^{-3}	-3.8554^{-3}	-2.7780^{-3}	-2.2077^{-3}	5.4411^{-2}
10^{-5}	1.3450^{-3}	5.1245^{-4}	4.6965^{-4}	1.8815^{-2}	-7.2936^{-3}	-5.1597^{-3}	-3.8359^{-3}	1.2368^{-1}
10^{-4}	5.9485^{-3}	2.2582^{-3}	9.4866^{-4}	3.0816^{-2}	-1.1297^{-2}	-7.1323^{-3}	-4.7404^{-3}	2.4503^{-1}
10^{-3}	2.4951^{-2}	9.4134^{-3}	1.7698^{-3}	3.8618^{-2}	-1.1654^{-2}	-3.8925^{-3}	-1.5608^{-3}	3.9912^{-1}
10^{-2}	9.4706^{-2}	3.5243^{-2}	2.6108^{-3}	2.8761^{-2}	-2.3471^{-3}	9.7827^{-3}	7.5188^{-3}	4.5698^{-1}
0.1	2.5982^{-1}	8.9780^{-2}	1.4862^{-3}	8.3807^{-3}	3.1615^{-3}	1.0522^{-2}	6.1973^{-3}	1.9561^{-1}
0.3	2.1732^{-1}	6.0165^{-2}	1.9984^{-4}	1.2086^{-3}	3.9371^{-4}	1.3919^{-3}	7.6929^{-4}	3.3906^{-2}
0.5	9.2445^{-2}	1.8539^{-2}	1.3804^{-5}	7.8411^{-5}	2.9799^{-5}	9.8805^{-5}	5.7333^{-5}	4.6166^{-3}
0.7	1.8219^{-2}	2.2090^{-3}	2.5004^{-7}_{*}	9.8927^{-7}_{*}	7.5552^{-7}	2.0057^{-6}	1.4438^{-6}	3.0231^{-4}
0.9	3.7653^{-4}	1.5285^{-5}	4.855^{-11}	2.005^{-9}	-7.599^{-10}	-5.171^{-10}	3.809^{-10}	1.3232^{-6}
$\mu_\mathrm{r}^2 = 1/2\,\mu_\mathrm{f}^2$								
10^{-7}	7.6699^{-5}	2.9289^{-5}	1.1912^{-4}	5.8548^{-3}	-2.3667^{-3}	-1.8030^{-3}	-1.4521^{-3}	3.1009^{-2}
10^{-6}	3.5067^{-4}	1.3364^{-4}	2.4707^{-4}	1.1806^{-2}	-4.7934^{-3}	-3.6731^{-3}	-2.7846^{-3}	7.2690^{-2}
10^{-5}	1.5611^{-3}	5.9329^{-4}	5.1593^{-4}	2.1406^{-2}	-8.5248^{-3}	-6.3125^{-3}	-4.4072^{-3}	1.5274^{-1}
10^{-4}	6.6957^{-3}	2.5346^{-3}	1.0509^{-3}	3.2903^{-2}	-1.2252^{-2}	-7.9608^{-3}	-4.8402^{-3}	2.8097^{-1}
10^{-3}	2.7125^{-2}	1.0200^{-2}	1.9310^{-3}	3.7698^{-2}	-1.1127^{-2}	-3.2334^{-3}	-7.5827^{-4}	4.2756^{-1}
10^{-2}	9.8758^{-2}	3.6602^{-2}	2.6980^{-3}	2.3675^{-2}	5.1386^{-5}	1.2092^{-2}	8.6053^{-3}	4.6241^{-1}
0.1	2.5572^{-1}	8.7847^{-2}	1.4179^{-3}	6.7523^{-3}	3.5944^{-3}	1.0578^{-2}	6.0904^{-3}	1.8838^{-1}
0.3	2.0497^{-1}	5.6318^{-2}	1.8228^{-4}	1.2965^{-3}	2.6142^{-4}	1.1713^{-3}	6.8941^{-4}	3.1884^{-2}
0.5	8.4404^{-2}	1.6775^{-2}	1.2174^{-5}	1.1604^{-4}	2.8309^{-6}	6.3682^{-5}	4.7009^{-5}	4.3221^{-3}
0.7	1.6013^{-2}	1.9215^{-3}	2.1196^{-7}_{*}	3.6047^{-6}	-7.4260^{-7}	3.1714^{-7}	9.6419^{-7}	2.8268^{-4}
0.9	3.0848^{-4}	1.2377^{-5}	3.829^{-11}	8.129^{-9}	-3.873^{-9}	-3.681^{-9}	-6.816^{-10}	1.2009^{-6}

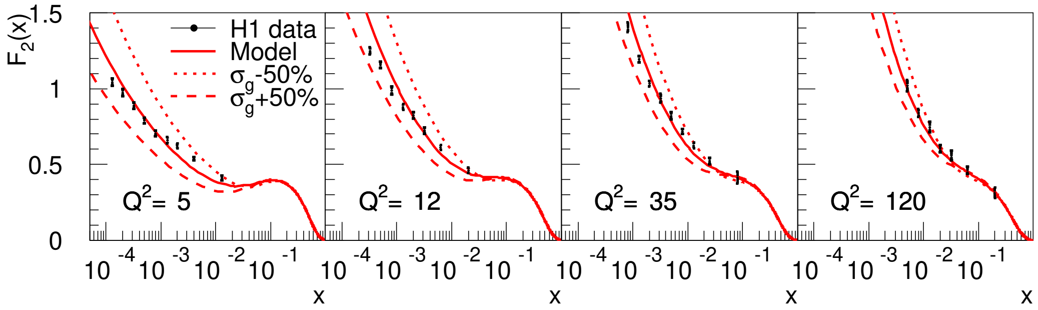


Fig. 13: $F_2(x, Q^2)$ from H1 compared to the model with $\pm 50\%$ variation of the width parameter σ_g of the gluon distribution.

Using a Monte Carlo method the resulting valence and sea parton x-distributions are obtained without approximations. These apply at a low scale Q_0^2 and the distributions at higher Q^2 are obtained using perturbative QCD evolution at next-to-leading order. To describe all parton distributions (Fig. 12c), the model has only four shape parameters and three normalization parameters, plus the starting scale:

$$\begin{array}{cccc} \sigma_u = 230 \text{ MeV} & \sigma_d = 170 \text{ MeV} & \sigma_g = 77 \text{ MeV} & \sigma_H = 100 \text{ MeV} \\ \alpha^2_{p\pi^0} = 0.45 & \alpha^2_{n\pi^+} = 0.14 & \alpha^2_{\Lambda K} = 0.05 & Q_0 = 0.75 \text{ GeV} \end{array} \tag{5.22}$$

These are determined from fits to data as detailed in [70] and illustrated in Fig. 13. The model reproduces the inclusive proton structure function and gives a natural explanation of observed quark asymmetries, such as the difference between the up and down valence distributions and between the anti-up and anti-down sea quark distributions. Moreover, its asymmetry in the momentum distribution of strange and anti-strange quarks in the nucleon is large enough to reduce the NuTeV anomaly to a level which does not give a significant indication of physics beyond the Standard Model.

Recent fits of PDF's at very low x and Q^2 have revealed problems with the gluon density, which in some cases even becomes negative. The reason for this is that the DGLAP evolution, driven primarily by the gluon at small x, otherwise gives too large parton densities and thereby a poor fit to F_2 in the genuine DIS region at larger Q^2. It has been argued [71] that the root of the problem is the application of the formalism for DIS also in the low-Q^2 region, where the momentum transfer is not large enough that the parton structure of the proton is clearly resolved. The smallest distance that can be resolved is basically given by the momentum transfer of the exchanged photon through $d = 0.2/\sqrt{Q^2}$, where d is in Fermi if Q^2 is in GeV^2. This indicates that partons are resolved only for $Q^2 \gtrsim 1\,\text{GeV}^2$. For $Q^2 \lesssim 1\,\text{GeV}^2$, there is no hard scale involved and a parton basis for the description is not justified. Instead, the interaction is here of a soft kind between the nearly on-shell photon and the proton. The cross section is then dominated by the process where the photon fluctuates into a virtual vector meson state which then interacts with the proton in a strong interaction. The quantum state of the photon can be expressed as $|\gamma\rangle = C_0|\gamma_0\rangle + \sum_V \frac{e}{f_V}|V\rangle + \int_{m_0} dm(\cdots)$. The sum is over $V = \rho^0, \omega, \phi \ldots$ as in the original vector meson dominance model (VDM), whereas the generalised vector meson dominance model (GVDM) also includes the integral over a continuous mass spectrum (not written out explicitly here).

Applied to ep at low Q^2 this leads to the expression [71]

$$\begin{aligned} F_2(x,Q^2) &= \frac{(1-x)Q^2}{4\pi^2\alpha} \left\{ \sum_{V=\rho,\omega,\phi} r_V \left(\frac{m_V^2}{Q^2+m_V^2}\right)^2 \left(1+\xi_V \frac{Q^2}{m_V^2}\right) \right. \\ &\quad \left. + r_C \left[(1-\xi_C)\frac{m_0^2}{Q^2+m_0^2} + \xi_C \frac{m_0^2}{Q^2} \ln\left(1+\frac{Q^2}{m_0^2}\right) \right] \right\} A_\gamma \frac{Q^{2\epsilon}}{x^\epsilon} \end{aligned} \tag{5.23}$$

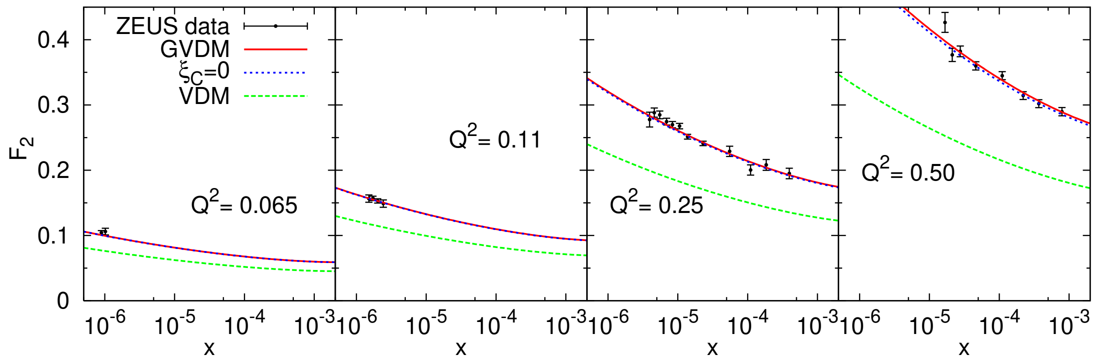


Fig. 14: F_2 data at low Q^2 from ZEUS compared to the full GVDM in eq. (5.23) (full curves), when excluding the longitudinal contribution of the continuum ($\xi_C = 0$) and excluding the continuous contribution altogether (setting $r_C = 0$) giving VDM.

where the hadronic cross-section $\sigma(ip \to X) = A_i s^{\epsilon} + B_i s^{-\eta} \approx A_i s^{\epsilon} \approx A_i (Q^2/x)^{\epsilon}$ has been used for the small-x region of interest. The parameters involved are all essentially known from GVDM phenomenology. With $\epsilon = 0.091$, $\xi = 0.34$, $m_0 = 1.5$ GeV and $A_\gamma = 71\,\mu$b, this GVDM model gives a good fit ($\chi^2/\mathrm{d.o.f.} = 87/66 = 1.3$) as illustrated in Fig. 14. Using this model at very low Q^2 in combination with the normal parton density approach at larger Q^2 it is possible to obtain a good description of data over the full Q^2 range [71]. This involves, however, a phenomenological matching of these two approaches, since a theoretically well justified combination is an unsolved problem.

Neglecting the GVDM component when fitting PDF's to data at small Q^2 may thus lead to an improper gluon distribution, which is not fully universal and therefore may give incorrect results when used for cross section calculations at LHC.

6 Towards precise determination of the nucleon PDFs [7]

The nucleon parton distribution functions (PDFs) available to the moment are extracted from the rather limited set of experimental distributions (the deep-inelastic scattering (DIS) structure functions, the Drell-Yan (DY) and jet production cross sections). Other high-energy processes potentially could provide additional constraints on PDFs, however insufficient theoretical understanding does not allow to use those data without risk of having uncontrolled theoretical inaccuracies. Even for the case of the existing global fits of the PDFs performed by the MRST and CTEQ groups missing next-to-next-to-leading (NNLO) order QCD corrections to the Drell-Yan and jet production cross sections are not small as compared to the accuracy of the corresponding data used and therefore might give non-negligible effect. In this section we outline progress in the QCD fits with consistent account of the NNLO corrections.

6.1 Impact of the NNLO evolution kernel fixation on PDFs

In order to allow account of the NNLO corrections in the fit of PDFs one needs analytical expressions for the 3-loop corrections to the QCD evolution kernel. Until recent times these expressions were known only in the approximate form of Ref. [61] derived from the partial information about the kernel, including the set of its Mellin moments and the low-x asymptotics [22, 23, 12] However with the refined calculations of Ref. [6, 7] the exact expression for the NNLO kernel has been available. These improvement is of particular importance for analysis of the low-x data including the HERA ones due to general rise of the high-order QCD correction at low x. We illustrate impact of the NNLO evolution kernel validation on PDFs using the case of fit to the global DIS data [72, 73, 74, 75, 76, 77]. The exact NNLO corrections to the DIS coefficient functions are know [78, 4] that allowed to perform approximate NNLO fit of PDFs

[7]Contributing author: S. I. Alekhin

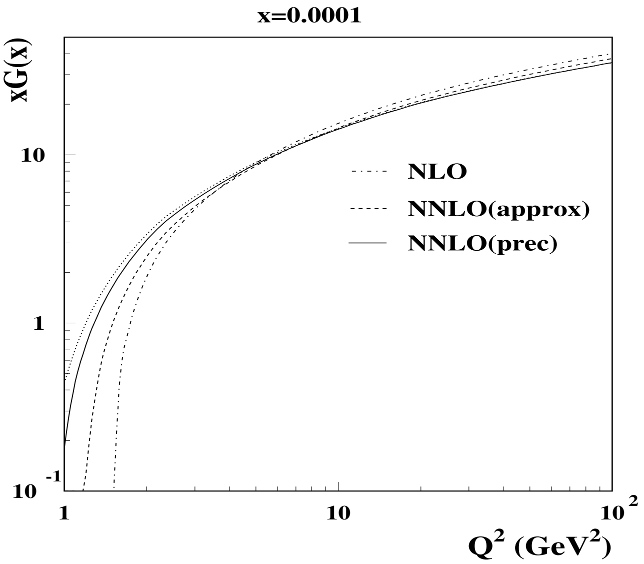


Fig. 15: The gluon distributions obtained in the different variants of PDFs fit to the DIS data (solid: the fit with exact NNLO evolution; dashes: the fit with approximate NNLO evolution; dots: the approximate NNLO gluons evolved with the exact NNLO kernel; dashed-dots: the NLO fit).

to these data [69] using the approximate NNLO corrections to the evolution kernel of Ref. [61]. Taking into account exact NNLO evolution kernel the analysis of Ref. [69] was updated recently to the exact NNLO case [79].

The gluon distributions at small x obtained in these two variants of the fit are compared in Fig.15. With the exact NNLO corrections the QCD evolution of gluon distribution at small x gets weaker and as a result at small x/Q the gluon distribution obtained using the precise NNLO kernel is quite different from the approximate one. In particular, the approximate NNLO gluon distribution is negative at $Q^2 \lesssim 1.3~\mathrm{GeV}^2$, while the precise one remains positive even below $Q^2 = 1~\mathrm{GeV}^2$. For the NLO case the positivity of gluons at small x/Q is even worse than for the approximate NNLO case due to the approximate NNLO corrections dampen the gluon evolution at small x too, therefore account of the NNLO corrections is crucial in this respect. (cf. discussion of Ref. [80]). Positivity of the PDFs is not mandatory beyond the QCD leading order, however it allows probabilistic interpretation of the parton model and facilitates modeling of the soft processes, such as underlying events in the hadron-hadron collisions at LHC. The change of gluon distribution at small x/Q as compared to the fit with approximate NNLO evolution is rather due the change in evolution kernel than due to shift in the fitted parameters of PDFs. This is clear from comparison of the exact NNLO gluon distribution to one obtained from the approximate NNLO fit and evolved to low Q using the exact NNLO kernel (see Fig.15). In the vicinity of crossover in the gluon distribution to the negative values its relative change due to variation of the evolution kernel is quite big and therefore further fixation of the kernel at small x discussed in Ref. [81] might be substantial for validation of the PDFs at low x/Q. For the higher-mass kinematics at LHC numerical impact of the NNLO kernel update is not dramatic. Change in the Higgs and W/Z bosons production cross sections due to more precise definition of the NNLO PDFs is comparable to the errors coming from the PDFs uncertainties, i.e. at the level of several percent.

6.2 NNLO fit of PDFs to the combined DIS and Drell-Yan data

The DIS process provide very clean source of information about PDFs both from experimental and theoretical side, however very poorly constrains the gluon and sea distributions at $x \gtrsim 0.3$. The well

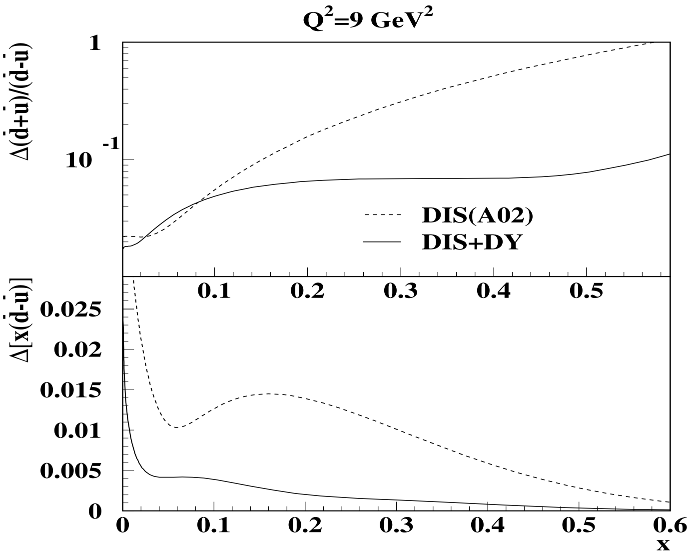


Fig. 16: Uncertainties in the non-strange sea distributions obtained from NNLO QCD fit to the DIS data combined with the fixed target Drell-Yan data (solid curves). The same uncertainties obtained in fit to the DIS data only [8] are given for comparison by dashes.

known way to improve precision of the sea distributions is to combine DIS data with the Drell-Yan ones. The cross section of process $NN \rightarrow l^+l^-$ reads

$$\sigma_{DY} \propto \sum_i \left[q_i(x_1)\bar{q}_i(x_2) + q_i(x_2)\bar{q}_i(x_1)\right] + \text{higher-order terms},$$

where $q(\bar{q})_i$ are the quarks(antiquarks) distribution and $x_{1,2}$ give the momentum fractions carried by each of the colliding partons. The quark distributions are determined by the DIS data with the precision of several percent in the wide region of x and therefore precision of the sea distribution extracted from the combined fit to the DIS and DY data is basically determined by the latter. The Fermilab fixed-target experiments provide measurements of the DY cross sections for the isoscalar target [82] and the ratio of cross sections for the deuteron and proton targets [83] with the accuracy better than 20% at $x \lesssim 0.6$. Fitting PDFs to these data combined with the global DIS data of Ref. [72, 73, 74, 75, 76, 77] we can achieve comparable precision in the sea distributions. Recent calculations of Ref. [84] allow to perform this fit with full account of the NNLO correction. Using these calculations the DY data of Refs. [82, 83] were included into the NNLO fit of Ref. [79] that leads to significant improvement in the precision of sea distributions (see Fig. 16). Due to the DY data on the deuteron/proton ratio the isospin asymmetry of sea is also improved. It is worth to note that the precision achieved for the total sea distribution is in good agreement to the rough estimates given above. The value of χ^2/NDP obtained in the fit is 1.1 and the spread of χ^2/NDP over separate experiments used in the fit is not dramatic, its biggest value is 1.4. We rescaled the errors in data for experiments with $\chi^2/\text{NDP} > 1$ in order to bring χ^2/NDP for this experiments to 1 and found that overall impact of this rescaling on the PDFs errors is marginal. This proofs sufficient statistical consistency of the data sets used in the fit and disfavors huge increase in the value of $\Delta\chi^2$ criterion suggested by the CTEQ collaboration for estimation of errors in the global fit of PDFs. A particular feature of the PDFs obtained is good stability with respect to the choice of factorization/renormalization scale in the DY cross section: Variation of this scale from $M_{\mu^+\mu^-}/2$ to $2M_{\mu^+\mu^-}$ leads to variation of PDFs comparable to their uncertainties due to errors in data.

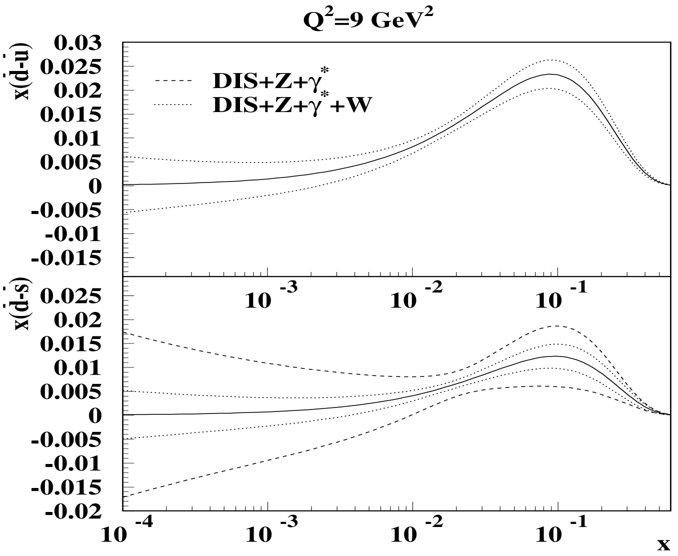


Fig. 17: The 1σ error band for $x(\bar{d}-\bar{u})$ (upper panel) and $x(\bar{d}-\bar{s})$ (lower panel) expected for the fit of PDFs to the LHC data combined with the global DIS ones. Dashed curves correspond to the case of $Z/\gamma*$-production, dots are for the combination $Z/\gamma*$- with the W^+/W^--production. Solid curves are for the central values obtained from the reference fit to the global DIS data

.

6.3 LHC data and flavor separation of the sea at small x

Combination of the existing DIS and fixed-target DY data provide good constraint on the total sea quarks distribution and allows separation of the $\bar{u}$- and $\bar{d}$-quark distributions up to the values of x sufficient for most practical applications at the LHC. At small x the total sea is also well constrained by the precise HERA data on the inclusive structure functions, however $\bar{u}/\bar{d}$ separation is poor in this region due to lack of the deuteron target data at HERA. The problem of the sea flavor separation is regularly masked due to additional constraints imposed on PDFs. In particular, most often the Regge-like behavior of the sea isospin asymmetry $x(\bar{d}-\bar{u}) \propto x^{a_{ud}}$ is assumed with a_{ud} selected around value of 0.5 motivated by the intercept of the meson trajectories. This assumption automatically provides constraint $\bar{d} = \bar{u}$ at $x \to 0$ and therefore leads to suppression of the uncertainties both in $\bar{u}$ and $\bar{d}$ at small x. If we do not assume the Regge-like behavior of $x(\bar{d}-\bar{u})$ its precision determined from the NNLO fit to the combined DIS and DY data of Section 1.2 is about 0.04 at $x = 10^{-4}$ furthermore this constraint is defined rather by assumption about the shape of PDFs at small x than by data used in the fit. The strange sea distribution is known much worse than the non-strange ones. It is essentially defined only by the CCFR experiment from the cross section of dimuon production in the neutrino nucleus collisions [85]. In this experiment the strange sea distribution was probed at $x = 0.01 \div 0.2$ and the shape obtained is similar to one of the non-strange sea with the strangeness suppression factor about 0.5. This is in clear disagreement with the Regge-like constraint on $x(\bar{d}-\bar{s})$ or $x(\bar{u}-\bar{s})$ and therefore we cannot use even this assumption to predict the strange sea at small x.

The LHC data on $\mu^+\mu^-$ production cross section can be used for further validation of the sea distributions at small x. Study of this process at the lepton pair masses down to 15 GeV will allow to probe PDFs at x down to 10^{-4}, while with both leptons detected full kinematics can be reliably reconstructed. In order to check impact of the foreseen LHC data on the sea flavor separation we generated sample of pseudo-data for the process $pp \to \mu^+\mu^-X$ at $\sqrt{s} = 14$ TeV with integral luminosity of 10 1/fb corresponding to the first stage of the LHC operation. In order to meet typical limitations of the LHC

Valence	a_u	0.718 ± 0.085
	b_u	3.81 ± 0.16
	ϵ_u	-1.56 ± 0.46
	γ_u	3.30 ± 0.49
	a_d	1.71 ± 0.20
	b_d	10.00 ± 0.97
	ϵ_d	-3.83 ± 0.23
	γ_d	4.64 ± 0.41
Sea	A_S	0.211 ± 0.016
	a_s	-0.048 ± 0.039
	b_s	2.20 ± 0.20
Glue	a_G	0.356 ± 0.095
	b_G	10.9 ± 1.4
	$\alpha_{\mathrm{s}}(M_{\mathrm{Z}})$	0.1132 ± 0.0015

Table 7: Values of the parameters obtained in the benchmark fit.

detectors only events with the lepton pair absolute rapidity less than 2.5 were accepted; other detector effects were not taken into account. For generation of these pseudo-data we used PDFs obtained in the dedicated version of fit [79] with the sea distributions parameterized as $xS_{u,d,s} = \eta_{u,d,s}x^a(1-x)^{b_{u,d,s}}$ with the constraints $\eta_u = \eta_d = \eta_s$ and $b_s = (b_u + b_d)/2$ imposed. These constraints are necessary for stability of the fit in view of limited impact of the DIS data on the flavor separation and, besides, the former one guarantees SU(3) symmetry in the sea distributions at small x. The generated pseudo-data were added to the basic DIS data sample and the errors in PDFs parameters were re-estimated with no constraints on the sea distributions imposed at this stage. Since dimuon data give extra information about the PDFs products they allow to disentangle the strange distribution, if an additional constraint on the non-strange sea distributions is set. The dashed curves in the lower panel of Fig.17 give the 1σ bands for $x(\bar{d}-\bar{s})$ as they are defined by the LHC simulated data combined with the global DIS ones given $(\bar{d}-\bar{u})$ is fixed. One can see that $\bar{d}/\bar{s}$ (and $\bar{u}/\bar{s}$) separation at the level of several percents would be feasible down to x=10^{-4} in this case. The supplementary constraint on $(\bar{d}-\bar{u})$ can be obtained from study of the W-boson charge asymmetry. To estimate impact of this process we simulated the single W^+- and W^--production data similarly to the case of the $\mu^+\mu^-$-production and took into account this sample too. In this case one can achieve separation of all three flavors with the precision better than 0.01 (see Fig.17). Note that strange sea separation is also improved due to certain sensitivity of the W-production cross section to the strange sea contribution. The estimates obtained refer to the ideal case of full kinematical reconstruction of the W-bosons events. For the case of using the charge asymmetry of muons produced from the W-decays the precision of the PDFs would be worse. Account of the backgrounds and the detector effects would also deteriorate it, however these losses can be at least partially compensated by rise of the LHC luminosity at the second stage of operation.

6.4 Benchmarking of the PDFs fit

For the available nucleon PDFs the accuracy at percent level is reached in some kinematical regions. For this reason benchmarking of the codes used in these PDFs fits is becoming important issue. A tool for calibration of the QCD evolution codes was provided by Les Houches workshop [59]. To allow benchmarking of the PDFs errors calculation we performed a test fit suggested in Les Houches workshop too. This fit reproduces basic features of the existing global fits of PDFs, but is simplified a lot to facilitate its reproduction. We use for the analysis data on the proton DIS structure functions F_2 obtained by the BCDMS, NM, H1, and ZEUS collaborations and ratio of the deuteron and proton structure functions F_2 obtained by the NMC. The data tables with full description of experimental errors taken into account are available online[8]. Cuts for the momentum transferred $Q^2 > 9$ GeV^2 and for invariant mass of the hadronic system $W^2 > 15$ GeV^2 are imposed in order to avoid influence of the power corrections and simplify calculations. The contribution of the Z-boson exchange at large Q is not taken into account for the same purpose. The PDFs are parameterized in the form

$$xp_i(x, 1\ GeV) = N_i x^{a_i}(1-x)^{b_i}(1+\epsilon_i\sqrt{x}+\gamma_i x),$$

to meet choice common for many popular global fits of PDFs. Some of the parameters ϵ_i and γ_i are set to zero since they were found to be consistent to zero within the errors. We assume isotopic symmetry for sea distribution and the strange sea is the same as the non-strange ones suppressed by factor of 0.5. Evolution of the PDFs is performed in the NLO QCD approximation within the $\overline{\text{MS}}$ scheme. The heavy quarks contribution is accounted in the massless scheme with the variable number of flavors (the thresholds for c- and b-quarks are 1.5 GeV and 4.5 GeV correspondingly). All experimental errors including correlated ones are taken into account for calculation of the errors in PDFs using the covariance matrix approach [86] and assuming linear propagation of errors. The results of the benchmark fit obtained with the code used in analysis of Refs. [69, 79] are given in Tables 7 and 8. The total number of the fitted PDF parameters left is 14. The normalization parameters N_i for the gluon and valence quark distributions are calculated from the momentum and fermion number conservation. The remaining normalization parameter A_S gives the total momentum carried by the sea distributions. Important note is that in view of many model assumptions made in the fit these results can be used mainly for the purposes of benchmarking rather for the phenomenological studies.

7 Benchmark Partons from DIS data and a Comparison with Global Fit Partons [9]

In this article I consider the uncertainties on partons arising from the errors on the experimental data that are used in a parton analysis. Various groups [87], [88], [69], [89], [76], [90], [91] have concentrated on the experimental errors and have obtained estimates of the uncertainties on parton distributions within a NLO QCD framework, using a variety of competing procedures. Here the two analyses, performed by myself and S. Alekhin (see Sec. 6) minimise the differences one obtains for the central values of the partons and the size of the uncertainties by fitting to exactly the same data sets with the same cuts, and using the same theoretical prescription. In order to be conservative we use only DIS data – BCDMS proton [73] and deuterium [74] fixed target data, NMC data on proton DIS and on the ratio $F_2^n(x,Q^2)/F_2^p(x,Q^2)$ [75], and H1 [76] and ZEUS [77] DIS data. We also apply cuts of $Q^2 = 9\text{GeV}^2$ and $W^2 = 15\text{GeV}^2$ in order to avoid the influence of higher twist. We each use NLO perturbative QCD in the $\overline{\text{MS}}$ renormalization and factorization scheme, with the zero-mass variable flavour number scheme and quark masses of $m_c = 1.5\text{GeV}$ and $m_b = 4.5\text{GeV}$. There is a very minor difference between $\alpha_S(\mu^2)$ used in the two fitting programs due to the different methods of implementing heavy quark thresholds (the differences being formally of higher order), as observed in the study by M. Whalley for this workshop [92]. If the couplings in the two approaches have the same value at $\mu^2 = M_Z^2$, then the MRST value is $\sim 1\%$ higher for $Q^2 \sim 20\text{GeV}^2$.

We each input our parton distributions at $Q_0^2 = 1\text{GeV}^2$ with a parameterization of the form

$$xf_i(x, Q_0^2) = A_i(1-x)^{b_i}(1+\epsilon_i x^{0.5}+\gamma_i x)x^{a_i}. \tag{7.24}$$

The input sea is constrained to be 40% up and anti-up quarks, 40% down and anti-down quarks, and 20% strange and antistrange. No difference between $\bar{u}$ and $\bar{d}$ is input. There is no negative term for the gluon, as introduced in [90], since this restricted form of data shows no strong requirement for it in order to obtain the best fit. Similarly we are able to set ϵ_g, γ_g, ϵ_S and γ_S all equal to zero. A_g is set by the momentum sum rule and A_{u_V} and A_{d_V} are set by valence quark number. Hence, there are nominally 13

[8]https://mail.ihep.ru/~alekhin/benchmark/TABLE

[9]Contributing author: R.S. Thorne.

	a_u	b_u	ϵ_u	γ_u	a_d	b_d	ϵ_d	γ_d	A_S	a_s	b_s	a_G	b_G	$\alpha_{\mathrm{s}}(M_{\mathrm{Z}})$
a_u	1.000	0.728	-0.754	-0.708	0.763	0.696	-0.444	0.215	-0.216	-0.473	-0.686	0.593	0.777	-0.006
b_u	0.728	1.000	-0.956	-0.088	0.377	0.620	-0.420	0.387	0.175	-0.182	-0.713	0.067	0.505	-0.337
ϵ_u	-0.754	-0.956	1.000	0.105	-0.388	-0.662	0.503	-0.485	-0.229	0.059	0.600	-0.047	-0.503	0.276
γ_u	-0.708	-0.088	0.105	1.000	-0.741	-0.390	0.219	0.107	0.597	0.591	0.310	-0.716	-0.675	-0.088
a_d	0.763	0.377	-0.388	-0.741	1.000	0.805	-0.622	0.248	-0.367	-0.509	-0.528	0.652	0.664	0.101
b_d	0.696	0.620	-0.662	-0.390	0.805	1.000	-0.904	0.728	0.017	-0.193	-0.512	0.272	0.576	-0.136
ϵ_d	-0.444	-0.420	0.503	0.219	-0.622	-0.904	1.000	-0.896	-0.132	-0.019	0.245	-0.038	-0.362	0.173
γ_d	0.215	0.387	-0.485	0.107	0.248	0.728	-0.896	1.000	0.346	0.240	-0.107	-0.241	0.120	-0.228
A_S	-0.216	0.175	-0.229	0.597	-0.367	0.017	-0.132	0.346	1.000	0.708	0.127	-0.375	-0.026	0.047
a_s	-0.473	-0.182	0.059	0.591	-0.509	-0.193	-0.019	0.240	0.708	1.000	0.589	-0.595	-0.241	-0.011
b_s	-0.686	-0.713	0.600	0.310	-0.528	-0.512	0.245	-0.107	0.127	0.589	1.000	-0.508	-0.402	-0.109
a_G	0.593	0.067	-0.047	-0.716	0.652	0.272	-0.038	-0.241	-0.375	-0.595	-0.508	1.000	0.565	0.587
b_G	0.777	0.505	-0.503	-0.675	0.664	0.576	-0.362	0.120	-0.026	-0.241	-0.402	0.565	1.000	-0.138
$\alpha_{\mathrm{s}}(M_{\mathrm{Z}})$	-0.006	-0.337	0.276	-0.088	0.101	-0.136	0.173	-0.228	0.047	-0.011	-0.109	0.587	-0.138	1.000

Table 8: Correlation coefficients for the parameters obtained in the benchmark fit.

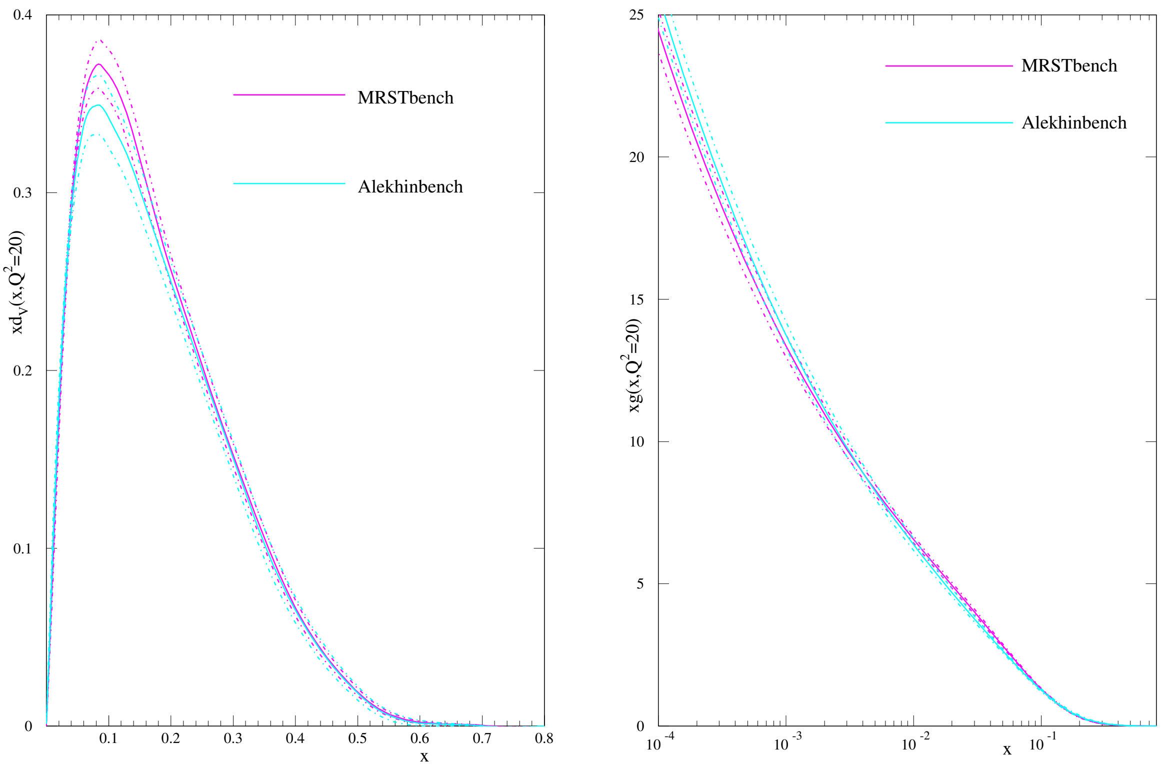


Fig. 18: Left plot: $xd_V(x, 20)$ from the MRST benchmark partons compared to that from the Alekhin benchmark partons. Right plot: $xg(x, 20)$ from the MRST benchmark partons compared to that from the Alekhin benchmark partons.

free parton parameters. However, the MRST fitting program exhibited instability in the error matrix due to a very high correlation between u_V parameters, so ϵ_u was set at its best fit value of $\epsilon_u = -1.56$, while 12 parameters were free to vary. The coupling was also allowed to vary in order to obtain the best fit. The treatment of the errors on the data was exactly as for the published partons with uncertainties for each group, i.e. as in [69] and [93]. This means that all detail on correlations between errors is included for the Alekhin fit (see Sec. 6), assuming that these errors are distributed in the Gaussian manner. The errors in the MRST fit are treated as explained in the appendix of [93], and the correlated errors are not allowed to move the central values of the data to as great an extent for the HERA data, and cannot do so at all for the fixed target data, where the data used are averaged over the different beam energies. The Alekhin approach is more statistically rigorous. The MRST approach is more pragmatic, reducing the ability of the data to move relative to the theory comparison by use of correlated errors (other than normalization), and is in some ways similar to the offset method [91]. The danger of this movement of data relative to theory has been suggested by the joint analysis of H1 and ZEUS data at this workshop (see [94]), where letting the joint data sets determine the movement due to correlated errors gives different results from when the data sets are compared to theoretical results.

7.1 Comparison Between the Benchmark Parton Distributions.

I compare the results of the two approaches to fitting the restricted data chosen for the benchmarking. The input parameters for the Alekhin fit are presented in Sec. 6. Those for the MRST type fit are similar, but there are some differences which are best illustrated by comparing the partons at a typical Q^2 for the data, e.g. $Q^2 = 20\text{GeV}^2$. A comparison is shown for the d_V quarks and the gluon in Fig. 18.

From the plots it is clear that there is generally good agreement between the parton distributions. The central values are usually very close, and nearly always within the uncertainties. The difference in the central values is mainly due to the different treatment of correlated errors, and partially due to the difference in the coupling definition. The uncertainties are similar in the two sets, but are generally about $1.2 - 1.5$ times larger for the Alekhin partons, due to the increased freedom in the use of the correlated experimental errors. The values of $\alpha_S(M_Z^2)$ are quite different, $\alpha_S(M_Z^2) = 0.1132 \pm 0.0015$ compared to 0.1110 ± 0.0012. However, as mentioned earlier, one expects a 1% difference due to the different threshold prescriptions – the MRST α_S would be larger at $Q^2 \sim 20\text{GeV}^2$, where the data are concentrated, so correspondingly to fit the data it receives a 1% shift downwards for $Q^2 = M_Z^2$. Once this systematic effect is taken into account, the values of $\alpha_S(M_Z^2)$ are very compatible. Hence, there is no surprising inconsistency between the two sets of parton distributions.

7.2 Comparison of the Benchmark Parton Distributions and Global Fit Partons.

It is also illuminating to show the comparison between the benchmark partons and the published partons from a global fit. This is done below for the MRST01 partons. For example, $u_V(x,Q^2)$ and $\bar{u}(x,Q^2)$ are shown in Fig. 19. It is striking that the uncertainties in the two sets are rather similar. This is despite the fact that the uncertainty on the benchmark partons is obtained from allowing $\Delta\chi^2 = 1$ in the fit while that for the MRST01 partons is obtained from $\Delta\chi^2 = 50$.[10] This illustrates the great improvement in precision which is obtained due to the increase in data from the relaxation of the cuts and the inclusion of types of data other than DIS. For the u_V partons, which are those most directly constrained by the DIS data in the benchmark fit, the comparison between the two sets of partons is reasonable, but hardly perfect – the central values differing by a few standard deviations. This is particularly important given that in this comparison the treatment of the data in the fit has been exactly the same in both cases. There is a minor difference in theoretical approach because of the simplistic treatment of heavy flavours in the benchmark fit. However, this would influence the gluon and sea quarks rather than valence quarks. Moreover, the region sensitive to this simplification would be $Q^2 \sim m_c^2$ (the lower charge weighting for bottom quarks greatly reducing the effect near $Q^2 = m_b^2$) which is removed by the Q^2 cut of 9GeV^2. Indeed, introducing the variable flavour number scheme usually used for the MRST partons modifies the benchmark partons only very minimally. Hence, if the statistical analysis is correct, the benchmark partons should agree with the global partons within their uncertainties (or at most 1.5 times their uncertainties, allowing for the effect of the correlated errors), which they do not. For the $\bar{u}$ partons the comparison is far worse, the benchmark partons being far larger at high x.

This disagreement in the high-x $\bar{u}$ partons can be understood better if one also looks at the high-x d_V distribution shown in Fig. 20. Here the benchmark distribution is very much smaller than for MRST01. However, the increase in the sea distribution, which is common to protons and neutrons, at high-x has allowed a good fit to the high-x BCDMS deuterium data even with the very small high-x d_V distribution. In fact it is a better fit than in [93]. However, the fit can be shown to break down with the additional inclusion of high-x SLAC data [72] on the deuterium structure function. More dramatically, the shape of the $\bar{u}$ is also completely incompatible with the Drell-Yan data usually included in the global fit, e.g. [82, 95]. Also in Fig. 20 we see that the d_V distributions are very different at smaller x. The benchmark set is markedly inconsistent with NMC data on $F_2^n(x,Q^2)/F_2^p(x,Q^2)$ which is at small x, but below the cut of $Q^2 = 9\text{GeV}^2$.

The gluon from the benchmark set is also compared to the MRST01 gluon in Fig. 21. Again there is an enormous difference at high x. Nominally the benchmark gluon has little to constrain it at high x. However, the momentum sum rule determines it to be very small in this region in order to get the best fit to HERA data, similar to the gluon from [76]. As such, the gluon has a small uncertainty and is many standard deviations from the MRST01 gluon. Indeed, the input gluon at high x is so small that its value at higher Q^2 is dominated by the evolution of u_V quarks to gluons, rather than by the input gluon. Hence, the uncertainty is dominated by the quark parton input uncertainty rather than its own, and since the up quark is well determined the uncertainty on the high-x gluon is small for the benchmark partons. The smallness of the high-x gluon results in the benchmark partons producing a very poor prediction indeed

[10]Though it is meant to be interpreted as a one sigma error in the former case and a 90% confidence limit in the latter.

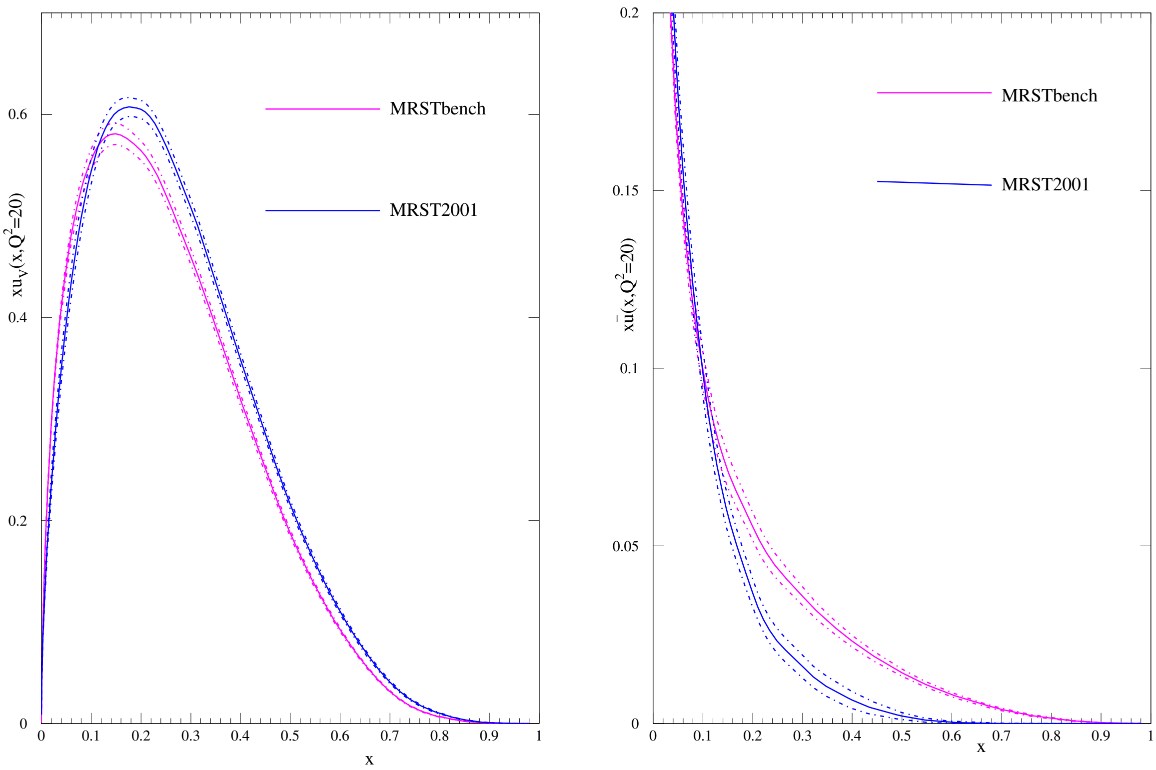


Fig. 19: Left plot: $xu_V(x,20)$ from the MRST benchmark partons compared to that from the MRST01 partons. Right plot: $x\bar{u}(x,20)$ from the MRST benchmark partons compared to that from the MRST01 partons with emphasis on large x.

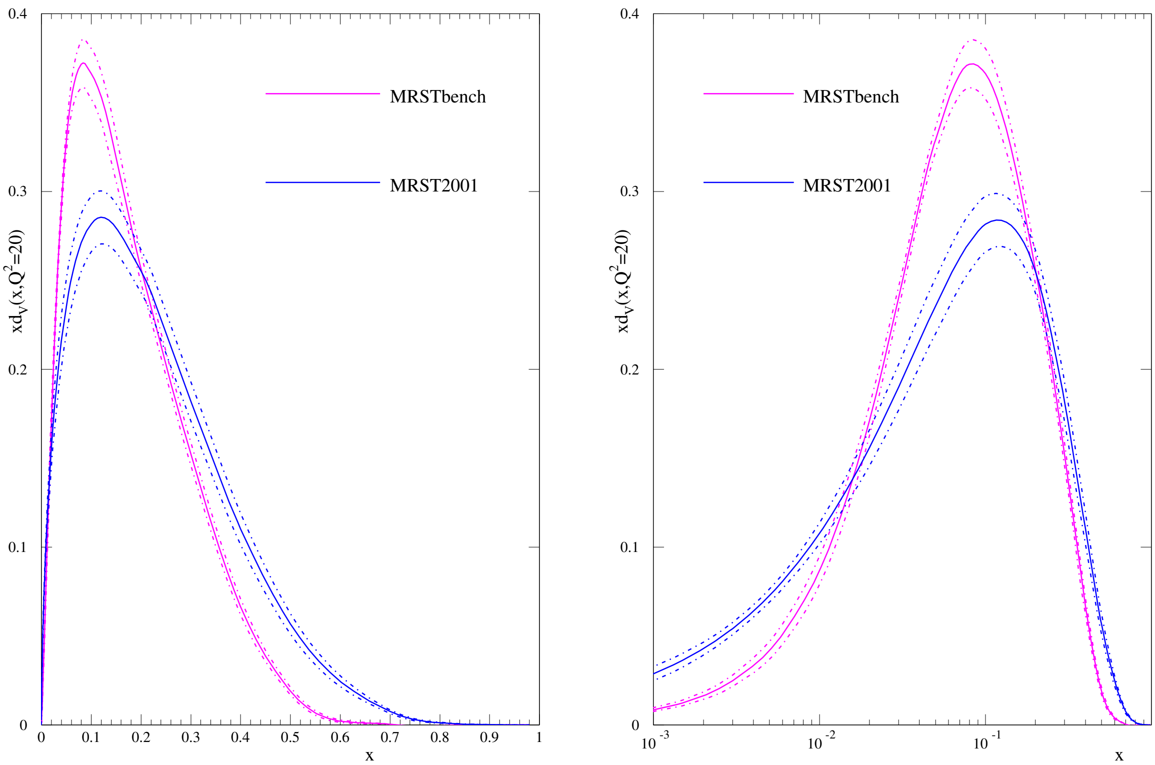


Fig. 20: Left plot: $xd_V(x,20)$ from the MRST benchmark partons compared to that from the MRST01 partons. Right plot: $xd_V(x,20)$ from the MRST benchmark partons compared to that from the MRST01 partons with emphasis on small x.

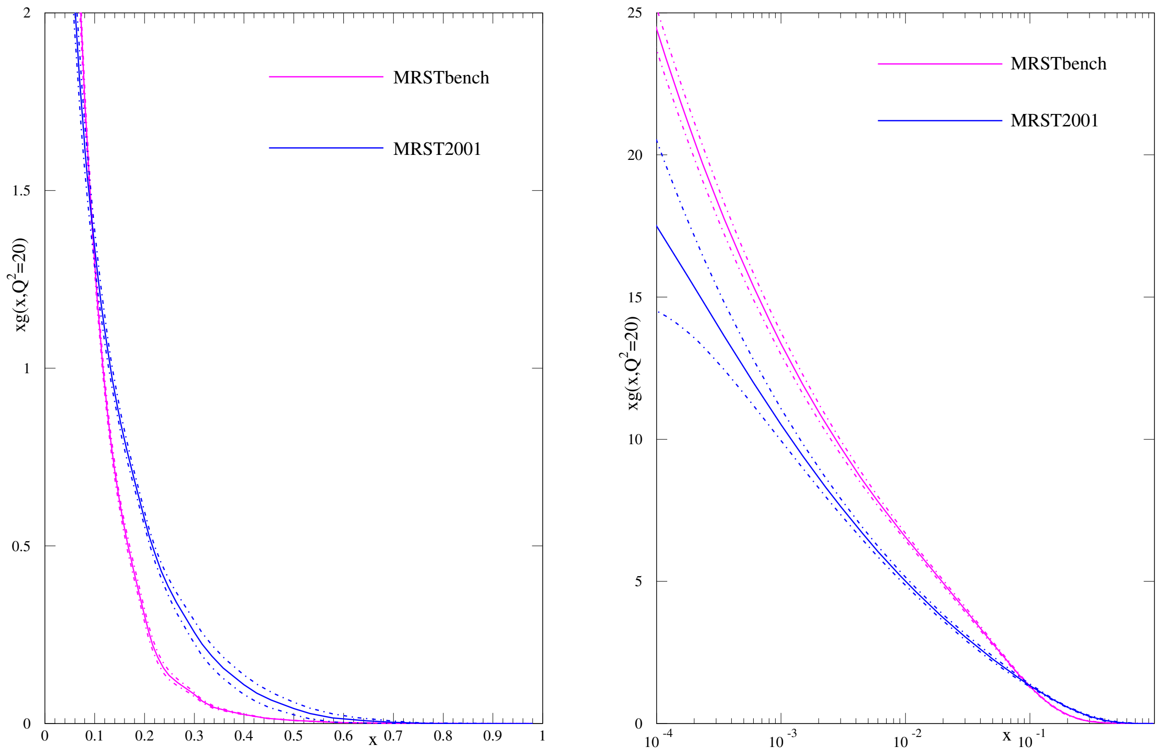


Fig. 21: Left plot: $xg(x, 20)$ from the MRST benchmark partons compared to that from the MRST2001 partons. Right plot: $xg(x, 20)$ from the MRST benchmark partons compared to that from the MRST2001 partons with emphasis on small x.

for the Tevatron jet data [96, 97], which are the usual data that constrain the high-x gluon in global fits.

It is also illustrative to look at small x. Here the benchmark gluon is only a couple of standard deviations from the MRST01 gluon, suggesting that its size is not completely incompatible with a good fit to the HERA small-x data at Q^2 below the benchmark cut. However, the uncertainty in the benchmark gluon is much smaller than in the MRST01 gluon, despite the much smaller amount of low-x data in the fit for the benchmark partons. This comes about as a result of the artificial choice made in the gluon input at Q_0^2. Since it does not have the term introduced in [93], allowing the freedom for the input gluon to be negative at very small x, the gluon is required by the fit to be valence-like. Hence, at input it is simply very small at small x. At higher Q^2 it becomes much larger, but in a manner driven entirely by evolution, i.e. it is determined by the input gluon at moderate x, which is well constrained. In this framework the small-x gluon does not have any intrinsic uncertainty – its uncertainty is a reflection of moderate x. This is a feature of e.g. the CTEQ6 gluon uncertainty [89], where the input gluon is valence-like. In this case the percentage gluon uncertainty does not get any larger once x reaches about 0.001. The alternative treatment in [93] gives the expected increase in the gluon uncertainty as $x \to 0$, since in this case the uncertainty is determined largely by that in the input gluon at small x. The valence-like input form for a gluon is an example of fine-tuning, the form being unstable to evolution in either direction. The artificial limit on the small-x uncertainty is a consequence of this.

7.3 Conclusions.

I have demonstrated that different approaches to fitting parton distributions that use exactly the same data and theoretical framework produce partons that are very similar and have comparable uncertainties. There are certainly some differences due to the alternative approaches to dealing with experimental errors, but these are relatively small. However, the partons extracted using a very limited data set are completely incompatible, even allowing for the uncertainties, with those obtained from a global fit with an identical treatment of errors and a minor difference in theoretical procedure. This implies that the inclusion of more data from a variety of different experiments moves the central values of the partons in a manner indicating either that the different experimental data are inconsistent with each other, or that the theoretical framework is inadequate for correctly describing the full range of data. To a certain extent both explanations are probably true. Some data sets are not entirely consistent with each other (even if they are seemingly equally reliable). Also, there are a wide variety of reasons why NLO perturbative QCD might require modification for some data sets, or in some kinematic regions [98]. Whatever the reason for the inconsistency between the MRST benchmark partons and the MRST01 partons, the comparison exhibits the dangers in extracting partons from a very limited set of data and taking them seriously. It also clearly illustrates the problems in determining the true uncertainty on parton distributions.

8 Stability of PDF fits [11]

One of the issues raised at the workshop is the reliability of determinations of parton distribution functions (PDFs), which might be compromised for example by the neglect of NNLO effects or non-DGLAP evolution in the standard analysis, or hidden assumptions made in parameterizing the PDFs at nonperturbative scales. We summarize the results of the CTEQ PDF group on this issue. For the full story see [80].

8.1 Stability of PDF determinations

The stability of NLO global analysis was seriously challenged by an analysis [98] which found a 20% variation in the cross section predicted for W production at the LHC – a critical "standard candle" process for hadron colliders – when certain cuts on input data are varied. If this instability were confirmed, it would significantly impact the phenomenology of a wide range of physical processes for the Tevatron Run II and the LHC. The CTEQ PDF group therefore performed an independent study of this issue within their global analysis framework. In addition, to explore the dependence of the results on assumptions about the parameterization of PDFs at the starting scale $Q_0 = 1.3\,\mathrm{GeV}$, we also studied the effect of allowing a negative gluon distribution at small x – a possibility that is favored by the MRST NLO analysis, and that is closely tied to the W cross section controversy.

The stability of the global analysis was investigated by varying the inherent choices that must be made to perform the analysis. These choices include the selection of experimental data points based on kinematic cuts, the functional forms used to parameterize the initial nonperturbative parton distribution functions, and the treatment of α_s.

The stability of the results is most conveniently measured by differences in the global χ^2 for the relevant fits. To quantitatively define a change of χ^2 that characterizes a significant change in the quality of the PDF fit is a difficult issue in global QCD analysis. In the context of the current analysis, we have argued that an increase by $\Delta\chi^2 \sim 100$ (for ~ 2000 data points) represents roughly a 90% confidence level uncertainty on PDFs due to the uncertainties of the current input experimental data [99, 100, 101, 89]. In other words, PDFs with $\chi^2 - \chi^2_{\mathrm{BestFit}} > 100$ are regarded as not tolerated by current data.

The CTEQ6 and previous CTEQ global fits imposed "standard" cuts $Q > 2\,\mathrm{GeV}$ and $W > 3.5\,\mathrm{GeV}$ on the input data set, in order to suppress higher-order terms in the perturbative expansion and the effects of resummation and power-law ("higher twist") corrections. We examined the effect of stronger cuts on Q to see if the fits are stable. We also examined the effect of imposing cuts on x, which should serve to suppress any errors due to deviations from DGLAP evolution, such as those predicted by BFKL. The idea is that any inconsistency in the global fit due to data points near the boundary of the accepted region will be revealed by an improvement in the fit to the data that remain after those near-boundary points have been removed. In other words, the decrease in χ^2 for the subset of data that is retained, when the PDF shape parameters are refitted to that subset alone, measures the degree to which

[11] Contributing authors: J. Huston, J. Pumplin.

Cuts	$Q_{\min}$	$x_{\min}$	N_{pts}	χ^2_{1926}	χ^2_{1770}	χ^2_{1588}	$\sigma_W^{\text{LHC}} \times B_{\ell\nu}$ [nb]
standard	2 GeV	0	1926	2023	1850	1583	20.02
intermediate	2.5 GeV	0.001	1770	–	1849	1579	20.10
strong	3.162 GeV	0.005	1588	–	–	1573	20.34

Table 9: Comparisons of three fits with different choices of the cuts on input data at the Q and x values indicated. In these fits, a conventional positive-definite gluon parameterization was used.

the fit to that subset was distorted in the original fit by compromises imposed by the data at low x and/or low Q.

The main results of this study are presented in Table 9. Three fits are shown, from three choices of the cuts on input data as specified in the table. They are labeled 'standard', 'intermediate' and 'strong'. N_{pts} is the number of data points that pass the cuts in each case, and $\chi^2_{N_{\text{pts}}}$ is the χ^2 value for that subset of data. The fact that the changes in χ^2 in each column are insignificant compared to the uncertainty tolerance is strong evidence that our NLO global fit results are very stable with respect to choices of kinematic cuts.

Cuts	$Q_{\min}$	$x_{\min}$	N_{pts}	χ^2_{1926}	χ^2_{1770}	χ^2_{1588}	$\sigma_W^{\text{LHC}} \times B_{\ell\nu}$ [nb]
standard	2 GeV	0	1926	2011	1845	1579	19.94
intermediate	2.5 GeV	0.001	1770	–	1838	1574	19.80
strong	3.162 GeV	0.005	1588	–	–	1570	19.15

Table 10: Same as Table 9 except that the gluon parameterization is extended to allow negative values.

We extended the analysis to a series of fits in which the gluon distribution $g(x)$ is allowed to be negative at small x, at the scale $Q_0 = 1.3\,\text{GeV}$ where we begin the DGLAP evolution. The purpose of this additional study is to determine whether the feature of a negative gluon PDF is a key element in the stability puzzle, as suggested by the findings of [98]. The results are presented in Table 10. Even in this extended case, we find no evidence of instability. For example, χ^2 for the subset of 1588 points that pass the *strong* cuts increases only from 1570 to 1579 when the fit is extended to include the full standard data set.

Comparing the elements of Table 9 and Table 10 shows that our fits with $g(x) < 0$ have slightly smaller values ofχ^2: e.g., 2011 versus 2023 for the standard cuts. However, the difference $\Delta\chi^2 = 12$ between these values is again not significant according to our tolerance criterion.

8.2 W cross sections at the LHC

The last columns of Tables 9 and 10 show the predicted cross section for $W^+ + W^-$ production at the LHC. This prediction is also very stable: it changes by only 1.6% for the positive-definite gluon parameterization, which is substantially less than the overall PDF uncertainty of σ_W estimated previously with the standard cuts. For the negative gluon parameterization, the change is 4%–larger, but still less than the overall PDF uncertainty. These results are explicitly displayed, and compared to the MRST results in Fig. 22. We see that this physical prediction is indeed insensitive to the kinematic cuts used for the fits, and to the assumption on the positive definiteness of the gluon distribution.

We also studied the stability of the prediction for σ_W using the Lagrange Multiplier (LM) method of Refs. [99, 100, 101]. Specifically, we performed a series of fits to the global data set that are constrained to specific values of σ_W close to the best-fit prediction. The resulting variation of χ^2 versus σ_W measures the uncertainty of the prediction. We repeated the constrained fits for each case of fitting

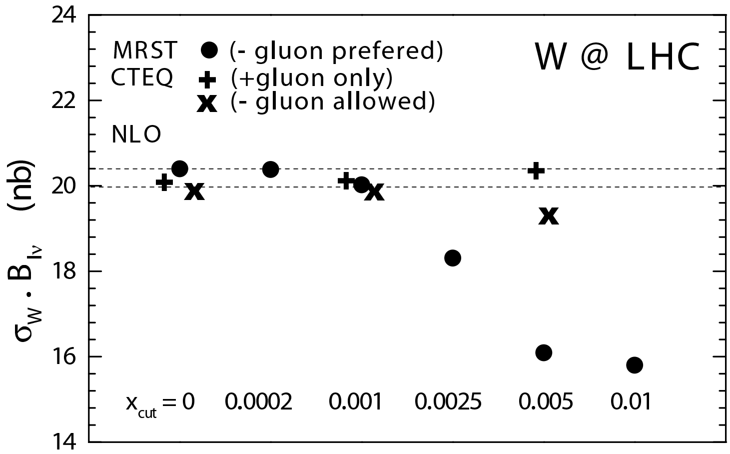


Fig. 22: Predicted total cross section of $W^+ + W^-$ production at the LHC for the fits obtained in our stability study, compared to the NLO results of Ref. [98]. The Q-cut values associated with the CTEQ points are given in the two tables. The overall PDF uncertainty of the prediction is $\sim 5\%$.

choices (parameterization and kinematic cuts). In this way we gain an understanding of the stability of the uncertainty, in addition to the stability of the central prediction.

Figure 23 shows the results of the LM study for the three sets of kinematic cuts described in Table 9, all of which have a positive-definite gluon distribution. The χ^2 shown along the vertical axis is

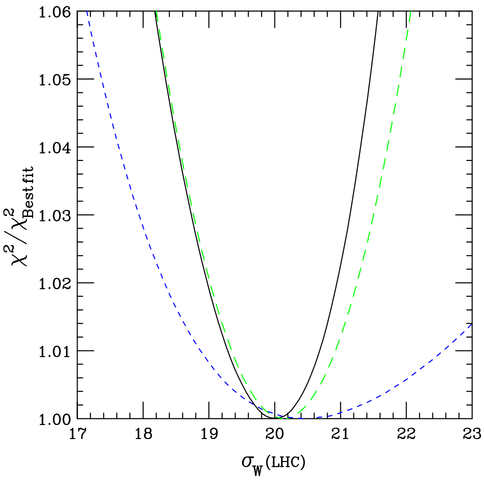


Fig. 23: Lagrange multiplier results for the W cross section (in nb) at the LHC using a positive-definite gluon. The three curves, in order of decreasing steepness, correspond to the three sets of kinematic cuts labeled standard/intermediate/strong in Table 9.

normalized to its value for the best fit in each series. In all three series, χ^2 depends almost quadratically on σ_W. We observe several features:

- The location of the minimum of each curve represents the best-fit prediction for $\sigma_W^{\rm LHC}$ for the corresponding choice of cuts. The fact that the three minima are close together displays the stability of the predicted cross section already seen in Table 9.
- Although more restrictive cuts make the global fit less sensitive to possible contributions from resummation, power-law and other nonperturbative effects, the loss of constraints caused by the removal of precision HERA data points at small x and low Q results directly in increased uncertainties on the PDF parameters and their physical predictions. This is shown in Fig. 23 by

the increase of the width of the curves with stronger cuts. The uncertainty of the predicted σ_W increases by more than a factor of 2 in going from the standard cuts to the strong cuts.

Figure 24 shows the results of the LM study for the three sets of kinematic cuts described in Table 10, all of which have a gluon distribution which is allowed to go negative.

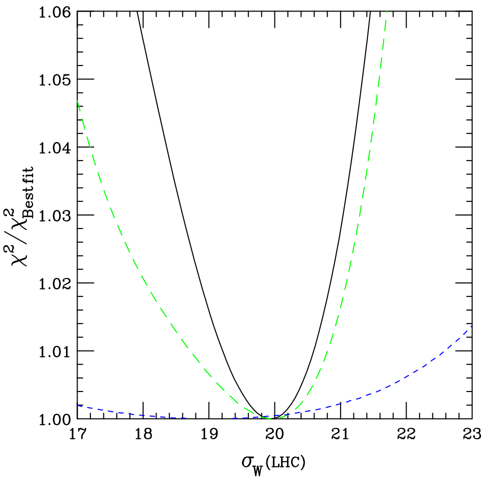


Fig. 24: Lagrange multiplier results for the W cross section (in nb) at the LHC using a functional form where the gluon is not required to be positive-definite. The three curves, in order of decreasing steepness, correspond to the three sets of kinematic cuts labeled standard/intermediate/strong in Table 10.

We observe:

- Removing the positive definiteness condition necessarily lowers the value of χ^2, because more possibilities are opened up in the χ^2 minimization procedure. But the decrease is insignificant compared to other sources of uncertainty. Thus, a negative gluon PDF is allowed, but not required.
- The minima of the two curves occur at approximately the same σ_W. Allowing a negative gluon makes no significant change in the central prediction – merely a decrease of about $1\,\%$, which is small compared to the overall PDF uncertainty.
- For the standard set of cuts, allowing a negative gluon PDF would expand the uncertainty range only slightly. For the intermediate and strong cuts, allowing a negative gluon PDF would significantly expand the uncertainty range.

We examined a number of aspects of our analysis that might account for the difference in conclusions between our stability study and that of [98]. A likely candidate seems to be that in order to obtain stability, it is necessary to allow a rather free parametrization of the input gluon distribution. This suspicion is seconded by recent work by MRST [102], in which a different gluon parametrization appears to lead to a best-fit gluon distribution that is close to that of CTEQ6. In summary, we found that the NLO PDFs and their physical predictions at the Tevatron and LHC are quite stable with respect to variations of the kinematic cuts and the PDF parametrization after all.

8.3 NLO and NNLO

In recent years, some preliminary next-to-next-leading-order (NNLO) analyses for PDFs have been carried out either for DIS alone [103], or in a global analysis context [51] – even if all the necessary hard cross sections, such as inclusive jet production, are not yet available at this order. Determining the parton distributions at NNLO is obviously desirable on theoretical grounds, and it is reasonable to plan for having a full set of tools for a true NNLO global analysis in place by the time LHC data taking begins. At the moment, however, NNLO fitting is not a matter of pressing necessity, since the difference between NLO and NNLO appears to be very small compared to the other uncertainties in the PDF analysis. This

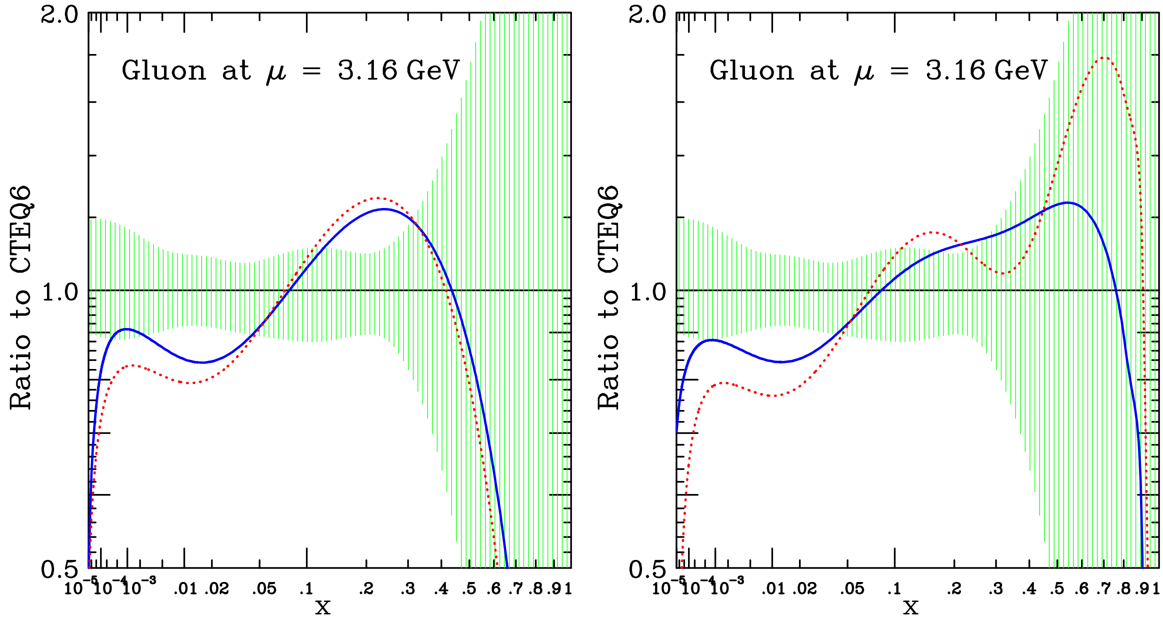


Fig. 25: Left: mrst2002 NLO (solid) and NNLO (dotted); Right: mrst2004 NLO (solid) and NNLO (dotted); Shaded region is uncertainty according to the 40 eigenvector sets of CTEQ6.1.

is demonstrated in Fig. 25, which shows the NLO and NNLO gluon distributions extracted by the MRST group. The difference between the two curves is much smaller than the other uncertainties measured by the 40 eigenvector uncertainty sets of CTEQ6.1, which is shown by the shaded region. The difference is also much smaller than the difference between CTEQ and MRST best fits. Similar conclusions [104] can be found using the NLO and NNLO fits by Alekhin.

9 The neural network approach to parton distributions [12]

The requirements of precision physics at hadron colliders, as has been emphasized through this workshop, have recently led to a rapid improvement in the techniques for the determination of parton distribution functions (pdfs) of the nucleon. Specifically it is now mandatory to determine accurately the uncertainty on these quantities, and the different collaborations performing global pdf analysis [51, 105, 69] have performed estimations of these uncertainties using a variety of techniques. The main difficulty is that one is trying to determine the uncertainty on a function, that is, a probability measure in a space of functions, and to extract it from a finite set of experimental data, a problem which is mathematically ill-posed. It is also known that the standard approach to global parton fits have several shortcomings: the bias introduced by choosing fixed functional forms to parametrize the parton distributions (also known as *model dependence*), the problems to assess faithfully the pdf uncertainties, the combination of inconsistent experiments, and the lack of general, process-independent error propagation techniques. Although the problem of quantifying the uncertainties in pdfs has seen a huge progress since its paramount importance was raised some years ago, until now no unambiguous conclusions have been obtained.

In this contribution we present a novel strategy to address the problem of constructing unbiased parametrizations of parton distributions with a faithful estimation of their uncertainties, based on a combination of two techniques: Monte Carlo methods and neural networks. This strategy, introduced in [106, 107], has been first implemented to address the marginally simpler problem of parametrizing deep-inelastic structure functions $F(x, Q^2)$, which we briefly summarize now. In a first step we construct a Monte Carlo sampling of the experimental data (generating artificial data replicas), and then we

[12]Contributing authors: L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, J. Rojo

train neural networks to each data replica, to construct a probability measure in the space of structure functions $\mathcal{P}\left[F(x,Q^2)\right]$. The probability measure constructed in this way contains all information from experimental data, including correlations, with the only assumption of smoothness. Expectation values and moments over this probability measure are then evaluated as averages over the trained network sample,

$$\left\langle \mathcal{F}\left[F(x,Q^2)\right]\right\rangle = \int \mathcal{D}F\mathcal{P}\left[F(x,Q^2)\right]\mathcal{F}\left[F(x,Q^2)\right] = \frac{1}{N_{\rm rep}}\sum_{k=1}^{N_{\rm rep}}\mathcal{F}\left(F^{({\rm net})(k)}(x,Q^2)\right) \ . \quad (9.25)$$

where $\mathcal{F}[F]$ is an arbitrary function of $F(x,Q^2)$.

The first step is the Monte Carlo sampling of experimental data, generating $N_{\rm rep}$ replicas of the original $N_{\rm dat}$ experimental data,

$$F_i^{({\rm art})(k)} = \left(1+r_N^{(k)}\sigma_N\right)\left[F_i^{({\rm exp})} + r_i^{s,(k)}\sigma_i^{stat} + \sum_{l=1}^{N_{sys}} r^{l,(k)}\sigma_i^{sys,l}\right] , \qquad i=1,\ldots,N_{\rm dat} \ , \quad (9.26)$$

where r are gaussian random numbers with the same correlation as the respective uncertainties, and $\sigma^{stat},\sigma^{sys},\sigma_N$ are the statistical, systematic and normalization errors. The number of replicas $N_{\rm rep}$ has to be large enough so that the replica sample reproduces central values, errors and correlations of the experimental data.

The second step consists on training a neural network[13] on each of the data replicas. Neural networks are specially suitable to parametrize parton distributions since they are unbiased, robust approximants and interpolate between data points with the only assumption of smoothness. The neural network training consist on the minimization for each replica of the χ^2 defined with the inverse of the experimental covariance matrix,

$$\chi^{2(k)} = \frac{1}{N_{\rm dat}}\sum_{i,j=1}^{N_{\rm dat}}\left(F_i^{({\rm art})(k)} - F_i^{({\rm net})(k)}\right){\rm cov}_{ij}^{-1}\left(F_j^{({\rm art})(k)} - F_j^{({\rm net})(k)}\right) \ . \quad (9.27)$$

Our minimization strategy is based on Genetic Algorithms (introduced in [108]), which are specially suited for finding global minima in highly nonlinear minimization problems.

The set of trained nets, once is validated through suitable statistical estimators, becomes the sought-for probability measure $\mathcal{P}\left[F(x,Q^2)\right]$ in the space of structure functions. Now observables with errors and correlations can be computed from averages over this probability measure, using eq. (9.25). For example, the average and error of a structure function $F(x,Q^2)$ at arbitrary (x,Q^2) can be computed as

$$\left\langle F(x,Q^2)\right\rangle = \frac{1}{N_{\rm rep}}\sum_{k=1}^{N_{\rm rep}}F^{({\rm net})(k)}(x,Q^2), \quad \sigma(x,Q^2) = \sqrt{\langle F(x,Q^2)^2\rangle - \langle F(x,Q^2)\rangle^2} \ . \quad (9.28)$$

A more detailed account of the application of the neural network approach to structure functions can be found in [107], which describes the most recent NNPDF parametrization of the proton structure function[14].

Hence this strategy can be used also to parametrize parton distributions, provided one now takes into account perturbative QCD evolution. Therefore we need to define a suitable evolution formalism.

[13] For a more throughly description of neural network, see [106] and references therein

[14] The source code, driver program and graphical web interface for our structure function fits is available at `http://sophia.ecm.ub.es/f2neural`.

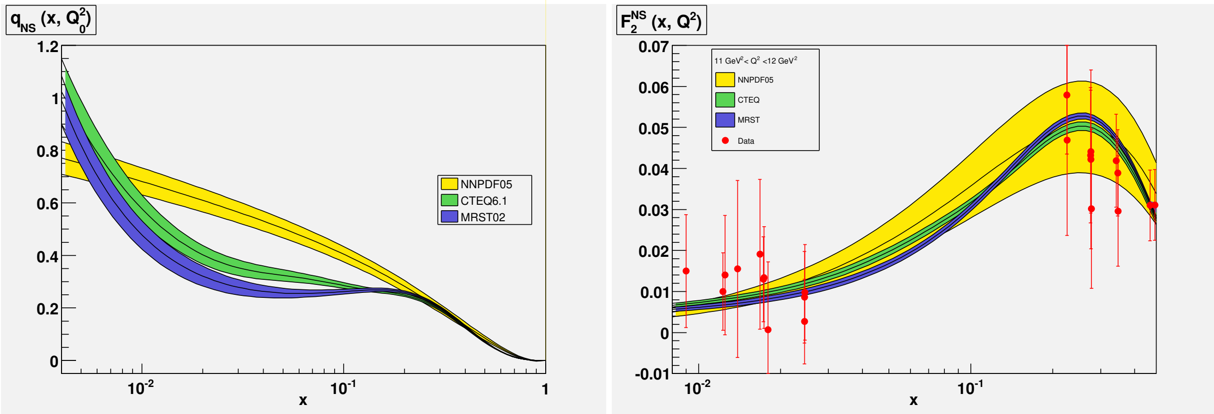


Fig. 26: Preliminary results for the NNPDF q_{NS} fit at $Q_0^2 = 2\,\mathrm{GeV}^2$, and the prediction for $F_2^{NS}(x, Q^2)$ compared with the CTEQ and MRST results.

Since complex neural networks are not allowed, we must use the convolution theorem to evolve parton distributions in $x-$space using the inverse $\Gamma(x)$ of the Mellin space evolution factor $\Gamma(N)$, defined as

$$q(N, Q^2) = q(N, Q_0^2)\Gamma\left(N, \alpha_s\left(Q^2\right), \alpha_s\left(Q_0^2\right)\right) \;, \tag{9.29}$$

The only subtlety is that the x-space evolution factor $\Gamma(x)$ is a distribution, which must therefore be regulated at $x = 1$, yielding the final evolution equation,

$$q(x, Q^2) = q(x, Q_0^2) \int_x^1 dy\, \Gamma(y) + \int_x^1 \frac{dy}{y}\Gamma(y)\left(q\left(\frac{x}{y}, Q_0^2\right) - yq(x, Q_0^2)\right) \;, \tag{9.30}$$

where in the above equation $q(x, Q_0^2)$ is parametrized using a neural network. At higher orders in perturbation theory coefficient functions $C(N)$ are introduced through a modified evolution factor, $\tilde{\Gamma}(N) \equiv \Gamma(N)C(N)$. We have benchmarked our evolution code with the Les Houches benchmark tables [59] at NNLO up to an accuracy of 10^{-5}. The evolution factor $\Gamma(x)$ and its integral are computed and interpolated before the neural network training in order to have a faster fitting procedure.

As a first application of our method, we have extracted the nonsinglet parton distribution $q_{NS}(x, Q_0^2) = \frac{1}{6}\left(u + \bar{u} - d - \bar{d}\right)(x, Q_0^2)$ from the nonsinglet structure function $F_2^{NS}(x, Q^2)$ as measured by the NMC [75] and BCDMS [73, 74] collaborations. The preliminary results of a NLO fit with fully correlated uncertainties [109] can be seen in fig. 26 compared to other pdfs sets. Our preliminary results appear to point in the direction that the uncertainties at small x do not allow, provided the current experimental data, to determine if $q_{NS}(x, Q^2)$ grows at small x, as supported by different theoretical arguments as well as by other global parton fits. However, more work is still needed to confirm these results. Only additional nonsinglet structure function data at small x could settle in a definitive way this issue[15].

Summarizing, we have described a general technique to parametrize experimental data in an bias-free way with a faithful estimation of their uncertainties, which has been successfully applied to structure functions and that now is being implemented in the context of parton distribution. The next step will be to construct a full set of parton distributions from all available hard-scattering data using the strategy described in this contribution.

References

[1] Van Neerven, W. L. and Zijlstra, E. B., Phys. Lett. **B272**, 127 (1991).

[15]Like the experimental low x deuteron structure function which would be measured in an hypothetical electron-deuteron run at HERA II, as it was pointed out during the workshop by M. Klein ([110]) and C. Gwenlan

[2] Zijlstra, E. B. and van Neerven, W. L., Phys. Lett. **B273**, 476 (1991).

[3] Zijlstra, E. B. and van Neerven, W. L., Phys. Lett. **B297**, 377 (1992).

[4] Zijlstra, E. B. and van Neerven, W. L., Nucl. Phys. **B383**, 525 (1992).

[5] Moch, S. and Vermaseren, J. A. M., Nucl. Phys. **B573**, 853 (2000).

[6] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. **B688**, 101 (2004).

[7] Vogt, A. and Moch, S. and Vermaseren, J. A. M., Nucl. Phys. **B691**, 129 (2004).

[8] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Phys. Lett. **B606**, 123 (2005).

[9] Vermaseren, J. A. M. and Vogt, A. and Moch, S., Nucl. Phys. **B724**, 3 (2005).

[10] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. **B646**, 181 (2002).

[11] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. Proc. Suppl. **135**, 137 (2004).

[12] Catani, S. and Hautmann, F., Nucl. Phys. **B427**, 475 (1994).

[13] Fadin, V. S. and Lipatov, L. N., Phys. Lett. **B429**, 127 (1998).

[14] Kuraev, E. A. and Lipatov, L. N. and Fadin, Victor S., Sov. Phys. JETP **45**, 199 (1977).

[15] Balitsky, I. I. and Lipatov, L. N., Sov. J. Nucl. Phys. **28**, 822 (1978).

[16] Jaroszewicz, T., Phys. Lett. **B116**, 291 (1982).

[17] Kirschner, R. and Lipatov, L. N., Nucl. Phys. **B213**, 122 (1983).

[18] Blümlein, J. and Vogt, A., Phys. Lett. **B370**, 149 (1996).

[19] G. Altarelli et al., *Resummation*. These proceedings.

[20] Van Neerven, W. L. and Vogt, A., Nucl. Phys. **B603**, 42 (2001).

[21] Larin, S. A. and van Ritbergen, T. and Vermaseren, J. A. M., Nucl. Phys. **B427**, 41 (1994).

[22] Larin, S. A. and Nogueira, P. and van Ritbergen, T. and Vermaseren, J. A. M., Nucl. Phys. **B492**, 338 (1997).

[23] Retey, A. and Vermaseren, J. A. M., Nucl. Phys. **B604**, 281 (2001).

[24] Blümlein, J. and Kurth, S., Phys. Rev. **D60**, 014018 (1999).

[25] Blümlein, J., Comput. Phys. Commun. **159**, 19 (2004).

[26] Remiddi, E. and Vermaseren, J. A. M., Int. J. Mod. Phys. **A15**, 725 (2000).

[27] Borwein, J. M. and Bradley, D. M. and Broadhurst, D. J. and Lisonek, P., Trans. Am. Math. Soc. **353**, 907 (2001).

[28] Blümlein, J., *in preparation*.

[29] Blümlein, J. and Moch, S., *in preparation*.

[30] Blümlein, J., Nucl. Phys. Proc. Suppl. **135**, 225 (2004).

[31] Blümlein, J., Comput. Phys. Commun. **133**, 76 (2000).

[32] Blümlein, J. and Moch, S., Phys. Lett. **B614**, 53 (2005).

[33] Nielsen, N., Nova Acta Leopoldina (Halle) **90**, 123 (1909).

[34] Matsuura, T. and van der Marck, S. C. and van Neerven, W. L., Nucl. Phys. **B319**, 570 (1989).

[35] Hamberg, R. and van Neerven, W. L. and Matsuura, T., Nucl. Phys. **B359**, 343 (1991).

[36] Ravindran, V. and Smith, J. and van Neerven, W. L., Nucl. Phys. **B682**, 421 (2004).

[37] Catani, S. and de Florian, D. and Grazzini, M., JHEP **05**, 025 (2001).

[38] Harlander, R. V. and Kilgore, W. B., Phys. Rev. **D64**, 013015 (2001).

[39] Harlander, R. V. and Kilgore, W. B., Phys. Rev. Lett. **88**, 201801 (2002).

[40] Harlander, R. V. and Kilgore, W. B., JHEP **10**, 017 (2002).

[41] Anastasiou, C. and Melnikov, K., Nucl. Phys. **B646**, 220 (2002).

[42] Ravindran, V. and Smith, J. and van Neerven, W. L., Nucl. Phys. **B665**, 325 (2003).

[43] Rijken, P. J. and van Neerven, W. L., Phys. Lett. **B386**, 422 (1996).

[44] Rijken, P. J. and van Neerven, W. L., Nucl. Phys. **B487**, 233 (1997).

[45] Rijken, P. J. and van Neerven, W. L., Phys. Lett. **B392**, 207 (1997).

[46] Blümlein, J. and Ravindran, V., Nucl. Phys. Proc. Suppl. **135**, 24 (2004).

[47] Blümlein, J. and Ravindran, V., Nucl. Phys. **B716**, 128 (2005).

[48] Blümlein, J. and Böttcher, H. and Guffanti, A., Nucl. Phys. Proc. Suppl. **135**, 152 (2004).

[49] Blumlein, J. and Bottcher, H. and Guffanti, A., *in preparation*.

[50] Alekhin, S., Phys. Rev. **D63**, 094022 (2001).

[51] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Phys. Lett. **B531**, 216 (2002).

[52] Botje, M. and Klein, M. and Pascaud, C. (1996).

[53] Blümlein, J. and Riemersma, S. and van Neerven, W. L. and Vogt, A., Nucl. Phys. Proc. Suppl. **51C**, 97 (1996).

[54] Blümlein, J. and Guffanti, A., *Scheme-invariant nnlo evolution for unpolarized dis structure functions*. Preprint hep-ph/0411110, 2004.

[55] Furmanski, W. and Petronzio, R., Zeit. Phys. **C11**, 293 (1982).

[56] Catani, S., Z. Phys. **C75**, 665 (1997).

[57] Blümlein, J. and Ravindran, V. and van Neerven, W. L., Nucl. Phys. **B586**, 349 (2000).

[58] Alekhin, S. and Blümlein, J., Phys. Lett. **B594**, 299 (2004).

[59] Giele, W. et al, *The qcd/sm working group: Summary report*. Preprint hep-ph/0204316, 2002.

[60] Blümlein, J. and Böttcher, H., Nucl. Phys. **B636**, 225 (2002).

[61] Van Neerven, W. L. and Vogt, A., Phys. Lett. **B490**, 111 (2000).

[62] Blümlein, J. et al., *A detailed comparison of nlo qcd evolution codes*. Preprint hep-ph/9609400, 1996.

[63] Vogt, A., Comput. Phys. Commun. **170**, 65 (2005).

[64] Lai, H. L. et al., Eur. Phys. J. **C12**, 375 (2000).

[65] Buza, M. and Matiounine, Y. and Smith, J. and van Neerven, W. L., Eur. Phys. J. **C1**, 301 (1998).

[66] Chuvakin, A. and Smith, J., Comput. Phys. Commun. **143**, 257 (2002).

[67] Larin, S. A. and van Ritbergen, T. and Vermaseren, J. A. M., Nucl. Phys. **B438**, 278 (1995).

[68] Chetyrkin, K. G. and Kniehl, B. A. and Steinhauser, M., Phys. Rev. Lett. **79**, 2184 (1997).

[69] Alekhin, S., Phys. Rev. **D68**, 014002 (2003).

[70] Alwall, J. and Ingelman, G., Phys. Rev. **D71**, 094015 (2005);
Alwall, J. and Ingelman, G., Phys. Rev. **D70**, 111505 (2004).

[71] Alwall, J. and Ingelman, G., Phys. Lett. **B596**, 77 (2004).

[72] Whitlow, L. W. and Riordan, E. M. and Dasu, S. and Rock, Stephen and Bodek, A., Phys. Lett. **B282**, 475 (1992).

[73] Benvenuti, A. C. et al., Phys. Lett. **B223**, 485 (1989).

[74] Benvenuti, A. C. et al., Phys. Lett. **B237**, 592 (1990).

[75] Arneodo, M. et al., Nucl. Phys. **B483**, 3 (1997).

[76] Adloff, C. et al., Eur. Phys. J. **C21**, 33 (2001).

[77] Chekanov, S. et al., Eur. Phys. J. **C21**, 443 (2001).

[78] Kazakov, D. I. and Kotikov, A. V., Phys. Lett. **B291**, 171 (1992).

[79] Alekhin, S., *Parton distribution functions from the precise nnlo qcd fit*. Preprint hep-ph/0508248, 2005.

[80] Huston, J. and Pumplin, J. and Stump, D. and Tung, W. K., JHEP **06**, 080 (2005).

[81] Altarelli, G. and Ball, R. D. and Forte, S., Nucl. Phys. **B674**, 459 (2003).

[82] Moreno, G. et al., Phys. Rev. **D43**, 2815 (1991).

[83] Towell, R. S. et al., Phys. Rev. **D64**, 052002 (2001).

[84] Anastasiou, C. and Dixon, L. J. and Melnikov, K. and Petriello, F., Phys. Rev. Lett. **91**, 182002 (2003).

[85] Bazarko, A. O. et al., Z. Phys. **C65**, 189 (1995).

[86] Alekhin, S., *Statistical properties of the estimator using covariance matrix*. Preprint hep-ex/0005042, 2000.

[87] Botje, M., Eur. Phys. J. **C14**, 285 (2000).

[88] Giele, W.T. and Keller, S., Phys. Rev. **D58**, 094023 (1998);
Giele, Walter T. and Keller, Stephane A. and Kosower, David A., *Parton distribution function uncertainties*. Preprint hep-ph/0104052, 2001.

[89] Pumplin, J. et al., JHEP **07**, 012 (2002).

[90] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C28**, 455 (2003).

[91] ZEUS Coll., Chekanov, S. et al., Phys. Rev. **D67**, 012007 (2003).

[92] Whalley, M., *talk at this workshop*.

[93] Martin, A. D. et al., Eur. Phys. J **C23**, 73 (2002).

[94] A. Cooper-Sarkar, C. Gwenlan, *Comparison and combination of zeus and h1 pdf analyses*. These proceedings.

[95] Webb, J. C. and others, *Absolute drell-yan dimuon cross sections in 800-gev/c p p and p d collisions*. Preprint hep-ex/0302019, 2003.

[96] Abbott, T. et al., Phys. Rev. Lett. **86**, 1707 (2001).

[97] Affolder, T. et al., Phys. Rev. **D64**, 032001 (2001).

[98] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C35**, 2004 (2004).

[99] Pumplin, J. and Stump, D. R. and Tung, W. K., Phys. Rev. **D65**, 014011 (2002).

[100] Stump, D. et al., Phys. Rev. **D65**, 014012 (2002).

[101] Pumplin, J. et al., Phys. Rev. **D65**, 014013 (2002).

[102] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Phys. Lett. **B604**, 61 (2004).

[103] Alekhin, S., *Nnlo parton distributions from deep-inelastic scattering data*. Preprint hep-ph/0311184, 2003.

[104] Pumplin, J. et al., *Parton distributions*. Preprint hep-ph/0507093, 2005.

[105] Stump, D. et al., JHEP **10**, 046 (2003).

[106] Forte, S. and Garrido, L. and Latorre, J. I. and Piccione, A., JHEP **05**, 062 (2002).

[107] Del Debbio, L. and Forte, S. and Latorre, J. I. and Piccione, A. and Rojo, J., JHEP **03**, 080 (2005).

[108] Rojo, J. and Latorre, J. I., JHEP **01**, 055 (2004).

[109] Rojo, J. and Del Debbio, L. and Forte, S. and Latorre, J. I. and Piccione, A., *The neural network approach to parton fitting*. Preprint hep-ph/0505044, 2005.

[110] M. Klein, B. Reisert, *Determination of the light quark momentum distributions at low x at hera*. These proceedings.

Resummation

G. Altarelli, J. Andersen, R. D. Ball, M. Ciafaloni, D. Colferai, G. Corcella, S. Forte, L. Magnea, A. Sabio Vera, G. P. Salam, A. Staśto

1 Introduction [1]

An accurate perturbative determination of the hard partonic cross-sections (coefficient functions) and of the anomalous dimensions which govern parton evolution is necessary for the precise extraction of parton densities. Recent progress in the determination of higher order contributions to these quantities has been reviewed in [1]. As is well known, such high-order perturbative calculations display classes of terms containing large logarithms, which ultimately signal the breakdown of perturbation theory. Because these terms are scale–dependent and in general non universal, lack of their inclusion can lead to significant distortion of the parton densities in some kinematical regions, thereby leading to loss of accuracy if parton distributions extracted from deep-inelastic scattering (DIS) or the Drell-Yan (DY) processes are used at the LHC.

Logarithimic enhancement of higher order perturbative contribution may take place when more than one large scale ratio is present. In DIS and DY this happen in the two opposite limits when the center-of-mass energy of the partonic collision is much higher than the characteristic scale of the process, or close to the threshold for the production of the final state. These correspond respectively to the small x and large x kinematical regions, where $0 \leq x \leq 1$ is defined in terms of the invariant mass M^2 of the non-leptonic final state as $M^2 = \frac{(1-x)Q^2}{x}$. The corresponding perturbative contributions are respectively enhanced by powers of $\ln \frac{1}{x}$ and $\ln(1-x)$, or, equivalently, in the space of Mellin moments, by powers of $\frac{1}{N}$ and $\ln N$, where $N \to 0$ moments dominate as $x \to 0$ while $N \to \infty$ moments dominate as $x \to 1$.

The theoretical status of small x and large x resummation is somewhat different. Large x logs are well understood and the corresponding perturbative corrections have been determined to all orders with very high accuracy. Indeed, the coefficients that determine their resummation can be extracted from fixed-order perturbative computations. Their resummation for DY and DIS was originally derived in [2, 3] and extended on very general grounds in [4]. The coefficients of the resulting exponentiation have now been determined so that resummation can now be performed exactly at N^2LL [5, 6], and to a very good approximation at N^3LL [7, 8, 9], including even some non-logarithmic terms [10]. On the other hand, small x logs are due to the fact that at high energies, due to the opening of phase space, both collinear [11, 12, 13] and high-energy [14, 15, 16, 17] logarithms contribute, and thus the coefficients required for their resummation can only be extracted from a simultaneous resolution of the DGLAP equation, which resums collinear logarithms, and the BFKL equation, which resum the high-energy logarithms. Although the determination of the kernels of these two equations has dramatically progressed in the last several years, thanks to the computation of the N^2LO DGLAP kernel [6, 18] and of the NLO BFKL kernel [14, 15, 16, 17, 19, 20], the formalism which is needed to combine these two equations, as required for sucessful phenomenology, has only recently progressed to the point of being usable for realistic applications [21, 22, 23, 24, 25, 26, 27, 28, 29, 30].

In practice, however, neither small x nor large x resummation is systematically incorporated in current parton fits, so data points for which such effects may be important must be discarded. This is especially unsatisfactory in the case of large x resummation, where resummed results (albeit with a varying degree of logarithmic accuracy) are available for essentially all processes of interest for a global parton fit, in particular, besides DIS and DY, prompt photon production [31, 32], jet production [33, 34] and heavy quark electroproduction [35, 36]. Even if one were to conclude that resummation is

[1] Subsection coordinator: S. Forte

not needed, either because (at small x) it is affected by theoretical uncertainties or because (at large x) its effects are small, this conclusion could only be arrived at after a careful study of the impact of resummation on the determination of parton distributions, which is not available so far.

The purpose of this section is to provide a first assessment of the potential impact of the inclusion of small x and large x resummation on the determination of parton distributions. In the case of large x, this will be done by determining resummation effects on parton distributions extracted from structure functions within a simplified parton fit. In the case of small x, this will be done through a study of the impact of small x resummation on splitting functions, as well as the theoretical uncertainty involved in the resummation process, in particular by comparing the results obtained within the approach of ref. [21, 22, 23] and that of ref. [24, 25, 26, 27, 28, 29, 30]. We will also discuss numerical approaches to the solution of the small-x (BFKL) evolution equation.

2 Soft gluons

With the current level of theoretical control of soft gluon resummations, available calculations for DIS or DY should be fully reliable over most of the available phase space. Specifically, one expects current (resummed) predictions for DIS structure functions to apply so long as the leading power correction can be neglected, *i.e.* so long as $W^2 \sim (1-x)Q^2 >> \Lambda^2$, with $x = x_{Bj}$. Similarly, for the inclusive DY cross section, one would expect the same to be true so long as $(1-z)^2Q^2 >> \Lambda^2$, where now $z = Q^2/\hat{s}$, with $\hat{s} = x_1x_2S$ the partonic center of mass energy squared. Indeed, as already mentioned, a consistent inclusion of resummation effects in parton fits is feasible with present knowledge: on the one hand, recent fits show that consistent parton sets can be obtained by making use of data from a single process (DIS) (see [37, 38] and Ref. [39]), on the other hand, even if one adopts the philosophy of global fits, resummed calculations are available for all processes of interest.

In practice, however, currently available global parton fits are based on NLO, or N^{2}LO fixed-order perturbative calculations, so data points which would lie within the expected reach of resummed calculations cannot be fit consistently and must be discarded. The effect is that large-x quark distributions become less constrained, which has consequences on the gluon distribution, as well as on medium-x quark distributions, through sum rules and evolution. The pool of untapped information is growing, as more data at large values of x have become available from, say, the NuTeV collaboration at Fermilab [40, 41]. A related issue is the fact that a growing number of QCD predictions for various processes of interest at the LHC are now computed including resummation effects in the hard partonic cross sections, which must be convoluted with parton densities in order to make predictions at hadron level. Such predictions are not fully consistent, since higher order effects are taken into account at parton level, but disregarded in defining the parton content of the colliding hadrons.

It is therefore worthwile to provide an assessment of the potential impact of resummation on parton distributions. Here, we will do this by computing resummation effects on quark distributions in the context of a simplified parton fit.

2.1 General Formalism in DIS

Deep Inelastic Scattering structure functions $F_i(x, Q^2)$ are given by the convolution of perturbative coefficient functions, typically given in the $\overline{\mathrm{MS}}$ factorization scheme, and parton densities. The coefficient functions C_i^q for quark-initiated DIS present terms that become large when the Bjorken variable x for the partonic process is close to $x = 1$, which forces gluon radiation from the incoming quark to be soft or collinear. At $\mathcal{O}(\alpha_s)$, for example, the coefficient functions can be written in the form

$$C_i^q\left(x, \frac{Q^2}{\mu_F^2}, \alpha_s(\mu^2)\right) = \delta(1-x) + \frac{\alpha_s(\mu^2)}{2\pi} H_i^q\left(x, \frac{Q^2}{\mu_F^2}\right) + \mathcal{O}\left(\alpha_s^2\right) . \qquad (1)$$

Treating all quarks as massless, the part of H_i^q which contains terms that are logarithmically enhanced as $x \to 1$ reads

$$H^q_{i,\mathrm{soft}}\left(x, \frac{Q^2}{\mu_F^2}\right) = 2C_F \left\{ \left[\frac{\ln(1-x)}{1-x}\right]_+ + \frac{1}{(1-x)_+}\left(\frac{\ln Q^2}{\mu_F^2} - \frac{3}{4}\right)\right\} . \tag{2}$$

In moment space, where soft resummation is naturally performed, the contributions proportional to $\alpha_s[\ln(1-x)/(1-x)]_+$ and to $\alpha_s[1/(1-x)]_+$ correspond to double ($\alpha_s \ln^2 N$) and single ($\alpha_s \ln N$) logarithms of the Mellin variable N. The Mellin transform of Eq. (2) in fact reads, at large N,

$$\hat{H}^q_{i,\mathrm{soft}}\left(N, \frac{Q^2}{\mu_F^2}\right) = 2C_F \left\{ \frac{1}{2}\ln^2 N + \left[\gamma_E + \frac{3}{4} - \frac{\ln Q^2}{\mu_F^2}\right] \ln N \right\} . \tag{3}$$

All terms growing logarithmically with N, as well as all N-independent terms corresponding to contributions proportional to $\delta(1-x)$ in x-space, have been shown to exponentiate. In particular, the pattern of exponentiation of logarithmic singularities is nontrivial: one finds that the coefficient functions can be written as

$$\hat{C}_i^q\left(N, \frac{Q^2}{\mu_F^2}, \alpha_s(\mu^2)\right) = \mathcal{R}\left(N, \frac{Q^2}{\mu_F^2}, \alpha_s(\mu^2)\right) \Delta\left(N, \frac{Q^2}{\mu_F^2}, \alpha_s(\mu^2)\right) , \tag{4}$$

where $\mathcal{R}(N, Q^2/\mu_F^2, \alpha_s(\mu^2))$ is a finite remainder, nonsingular as $N \to \infty$, while [4]

$$\ln \Delta\left(N, \frac{Q^2}{\mu_F^2}, \alpha_s(\mu^2)\right) = \int_0^1 dx \frac{x^{N-1}-1}{1-x}\left\{\int_{\mu_F^2}^{(1-x)Q^2} \frac{dk^2}{k^2} A\left[\alpha_s(k^2)\right] + B\left[\alpha_s\left(Q^2(1-x)\right)\right]\right\} . \tag{5}$$

In Eq. (5) the leading logarithms (LL), of the form $\alpha_s^n \ln^{n+1} N$, are generated at each order by the function A. Next-to-leading logarithms (NLL), on the other hand, of the form $\alpha_s^n \ln^n N$, require the knowledge of the function B. In general, resumming N^{k}LL to all orders requires the knowledge of the function A to $k+1$ loops, and of the function B to k loops. In the following, we will adopt the common standard of NLL resummation, therefore we need the expansions

$$A(\alpha_s) = \sum_{n=1}^{\infty} \left(\frac{\alpha_s}{\pi}\right)^n A^{(n)} \ ; \ B(\alpha_s) = \sum_{n=1}^{\infty} \left(\frac{\alpha_s}{\pi}\right)^n B^{(n)} \tag{6}$$

to second order for A and to first order for B. The relevant coefficients are

$$\begin{aligned} A^{(1)} &= C_F , \\ A^{(2)} &= \frac{1}{2} C_F \left[C_A\left(\frac{67}{18} - \frac{\pi^2}{6}\right) - \frac{5}{9} n_f\right] , \\ B^{(1)} &= -\frac{3}{4} C_F . \end{aligned} \tag{7}$$

Notice that in Eq. (5) the term $\sim A(\alpha_s(k^2))/k^2$ resums the contributions of gluons that are both soft and collinear, and in fact the anomalous dimension A can be extracted order by order from the residue of the singularity of the nonsinglet splitting function as $x \to 1$. The function B, on the other hand, is related to collinear emission from the final state current jet.

In [35, 36] soft resummation was extended to the case of heavy quark production in DIS. In the case of heavy quarks, the function $B(\alpha_s)$ needs to be replaced by a different function, called $S(\alpha_s)$ in [36], which is characteristic of processes with massive quarks, and includes effects of large-angle soft radiation. In the following, we shall consider values of Q^2 much larger than the quark masses and employ the resummation results in the massless approximation, as given in Eq. (5).

2.2 Simplified parton fit

We would like to use large-x resummation in the DIS coefficient functions to extract resummed parton densities from DIS structure function data. Large-x data typically come from fixed-target experiments: in the following, we shall consider recent charged-current (CC) data from neutrino-iron scattering, collected by the NuTeV collaboration [40, 41], and neutral-current (NC) data from the NMC [42] and BCDMS [43, 44] collaborations.

Our strategy will be to make use of data at different, fixed values of Q^2. We will extract from these data moments of the corresponding structure functions, with errors; since such moments factor into a product of moments of parton densities times moments of coefficient functions, computing parton moments with errors is straightforward. We then compare NLO to resummed partons in Mellin space, and subsequently provide a translation back to x-space by means of simple parametrization. Clearly, given the limited data set we are working with, our results will be affected by comparatively large errors, and we will have to make simplifying assumptions in order to isolate specific quark densities. Resummation effects are, however, clearly visible, and we believe that our fit provides a rough quantitative estimate of their size. A more precise quantitative analysis would have to be performed in the context of a global fit.

The first step is to construct a parametrization of the chosen data. An efficient and faithful parametrization of the NMC and BCDMS neutral-current structure functions was provided in [45, 46], where a large sample of Monte Carlo copies of the original data was generated, taking properly into account errors and correlations, and a neural network was trained on each copy of the data. One can then use the ensemble of networks as a faithful and unbiased representation of the probability distribution in the space of structure functions. We shall make use of the nonsinglet structure function $F_2^{\rm ns}(x,Q^2)$ extracted from these data, as it is unaffected by gluon contributions, and provides a combination of up and down quark densities which is independent of the ones we extract from charged current data (specifically, $F_2^{\rm ns}(x,Q^2)$ gives $u-d$).

As far as the NuTeV data are concerned, we shall consider the data on the CC structure functions F_2 and F_3. The structure function F_3 can be written as a convolution of the coefficient function C_3^q with quark and antiquark distributions, with no gluon contribution, as

$$xF_3 = \frac{1}{2}\left(xF_3^{\nu} + xF_3^{\bar\nu}\right) = x\left[\sum_{q,q'} |V_{qq'}|^2\,(q-\bar q)\otimes C_3^q\right]. \tag{8}$$

We consider data for F_3 at $Q^2 = 12.59$ and 31.62 GeV^2, and, in order to compute moments, we fit them using the functional form

$$xF_3(x) = Cx^{-\rho}(1-x)^{\sigma}(1+kx)\,. \tag{9}$$

The best-fit values of C, ρ and δ, along with the χ^2 per degree of freedom, are given in [47]. Here we show the relevant NuTeV data on xF_3, along with our best-fit curves, in Fig. 1.

The analysis of NuTeV data on F_2 is slightly complicated by the fact that gluon-initiated DIS gives a contribution, which, however, is not enhanced but suppressed at large x. We proceed therefore by taking the gluon density from a global fit, such as the NLO set CTEQ6M [48], and subtract from F_2 the gluon contribution point by point. We then write F_2 as

$$F_2 \equiv \frac{1}{2}\left(F_2^{\nu} + F_2^{\bar\nu}\right) = x\sum_{q,q'} |V_{qq'}|^2\left[(q+\bar q)\otimes C_2^q + g\otimes C_2^g\right] \equiv F_2^q + F_2^g\,, \tag{10}$$

and fit only the quark-initiated part F_2^q, using the same parametrization as in Eq. (9). Fig. 2 shows the data on F_2^q and the best fit curves, as determined in Ref. [47]. After the subtraction of the gluon contribution from F_2, the structure functions we are considering (F_2^q, xF_3 and $F_2^{\rm ns}$) are all given in factorized form as

$$F_i(x,Q^2) = x\int_x^1 \frac{d\xi}{\xi}\,q_i\left(\xi,\mu_F^2\right)C_i^q\left(\frac{x}{\xi},\frac{Q^2}{\mu_F^2},\alpha_s(\mu^2)\right), \tag{11}$$

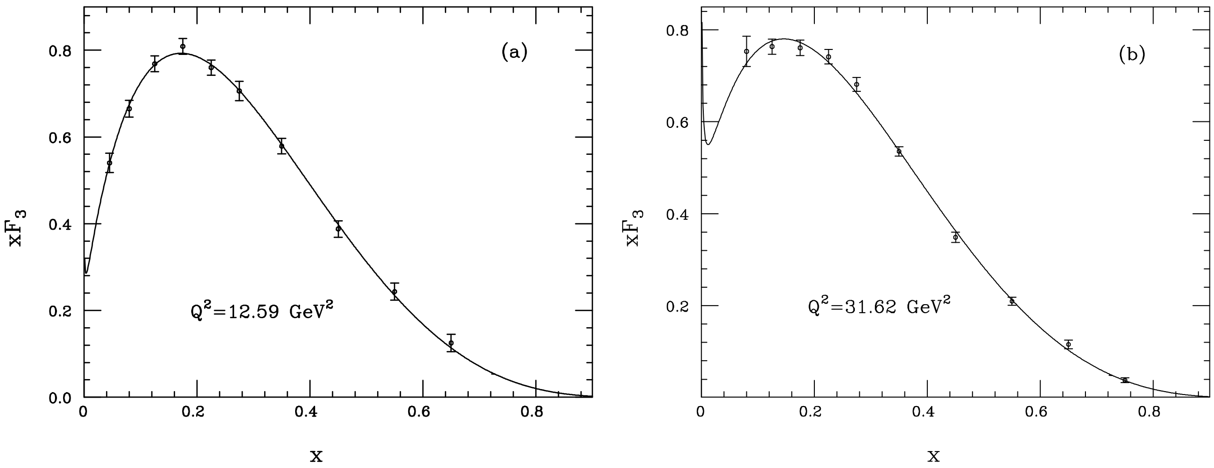


Fig. 1: NuTeV data on the structure function xF_3, at $Q^2 = 12.59$ GeV^2 (a) and at $Q^2 = 31.62$ GeV^2 (b), along with the best fit curve parametrized by Eq. (9).

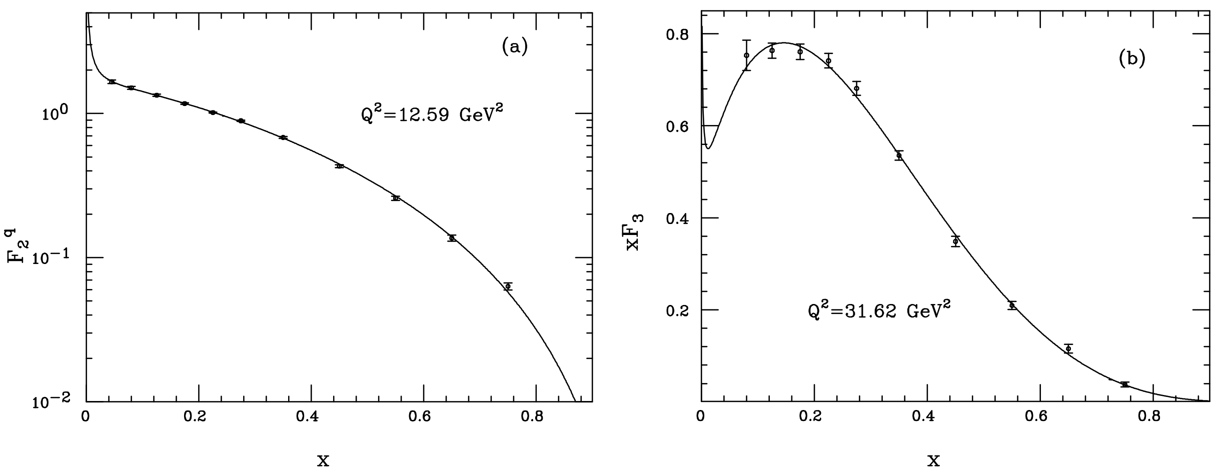


Fig. 2: NuTeV data on the quark-initiated contribution F_2^q to the structure function F_2, for $Q^2 = 12.59$ GeV^2 (a), and $Q^2 = 31.62$ GeV^2 (b). The solid lines are the best-fit predictions.

where C_i^q is the relevant coefficient function and q_i is a combination of quark and antiquark distributions only. Hereafter, we shall take $\mu = \mu_F = Q$ for the factorization and renormalization scales. At this point, to identify individual quark distributions from this limited set of data, we need to make some simplifying assumptions. Following [47], we assume isospin symmetry of the sea, $\bar{u} = \bar{d}$, $s = \bar{s}$ and we further impose a simple proportionality relation expressing the antistrange density in terms of the other antiquarks, $\bar{s} = \kappa\bar{u}$. As in [47], we shall present results for $\kappa = \frac{1}{2}$. With these assumptions, we can explicit solve for the remaining three independent quark densities (up, down, and, say, strange), using the three data sets we are considering.

Taking the Mellin moments of Eq. (11), the convolution becomes an ordinary product and we can extract NLO or NLL-resummed parton densities, according to whether we use NLO or NLL coefficient functions. More precisely,

$$\hat{q}_i^{\mathrm{NLO}}(N, Q^2) = \frac{\hat{F}_i(N-1, Q^2)}{\hat{C}_i^{\mathrm{NLO}}(N, 1, \alpha_s(Q^2))} \; ; \quad \hat{q}_i^{\mathrm{res}}(N, Q^2) = \frac{\hat{F}_i(N-1, Q^2)}{\hat{C}_i^{\mathrm{res}}(N, 1, \alpha_s(Q^2))} \,. \tag{12}$$

After extracting the combinations q_i, one can derive the individual quark densities, at NLO and including NLL large-x resummation. We concentrate our analysis on the up quark distribution, since experimental errors on the structure functions are too large to see an effect of the resummation on the other quark densities, such as d or s, with the limited data set we are using.

2.3 Impact of the resummation

We present results for moments of the up quark distribution in Figs. 3 and 4. Resummation effects

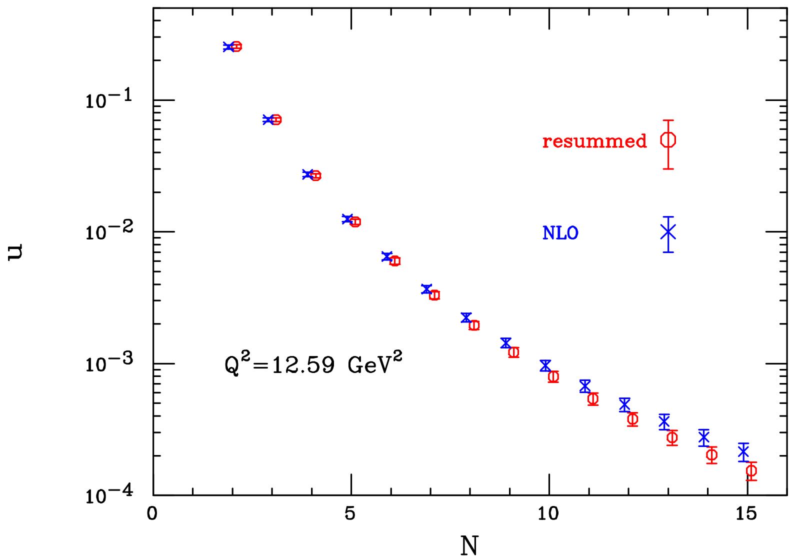


Fig. 3: NLO and resummed moments of the up quark distribution at $Q^2 = 12.59$ GeV^2

become statistically significant around $N \sim 6-7$ at both values of Q^2. Notice that high moments of the resummed up density are *suppressed* with respect to the NLO density, as a consequence of the fact that resummation in the $\overline{\mathrm{MS}}$ scheme enhances high moments of the coefficient functions.

In order to illustrate the effect in the more conventional setting of x-space distributions, we fit our results for the moments to a simple parametrization of the form $u(x) = Dx^{-\gamma}(1-x)^{\delta}$. Our best fit values for the parameters, with statistical errors, are given in Table (1), and the resulting distributions are displayed in Fig. 5, with one standard deviation uncertainty bands. Once again, the effect of soft resummation is clearly visible at large x: it suppresses the quark densities extracted from the given structure function data with respect to the NLO prediction.

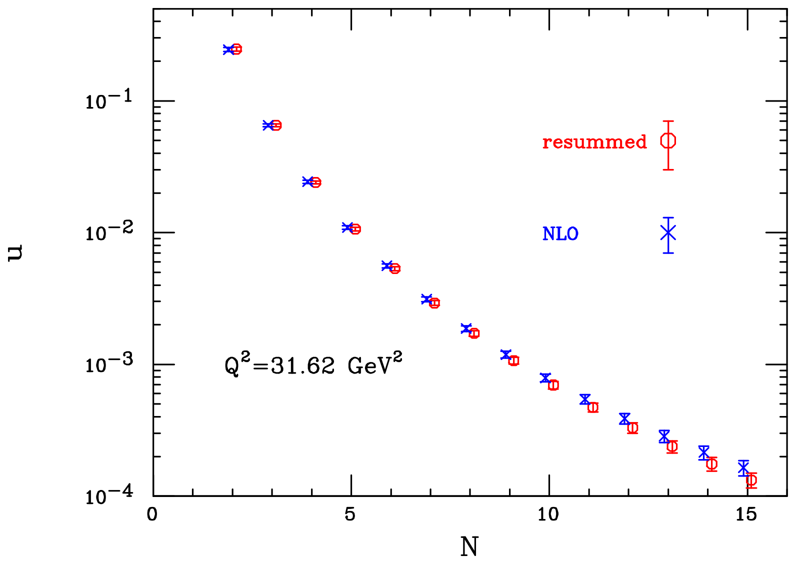


Fig. 4: As in Fig. 3, but at $Q^2 = 31.62$ GeV^2.

Q^2	PDF	D	γ	δ
12.59	NLO	3.025 ± 0.534	0.418 ± 0.101	3.162 ± 0.116
	RES	4.647 ± 0.881	0.247 ± 0.109	3.614 ± 0.128
31.62	NLO	2.865 ± 0.420	0.463 ± 0.086	3.301 ± 0.098
	RES	3.794 ± 0.583	0.351 ± 0.090	3.598 ± 0.104

Table 1: Best fit values and errors for the up-quark x-space parametrization, at the chosen values of Q^2.

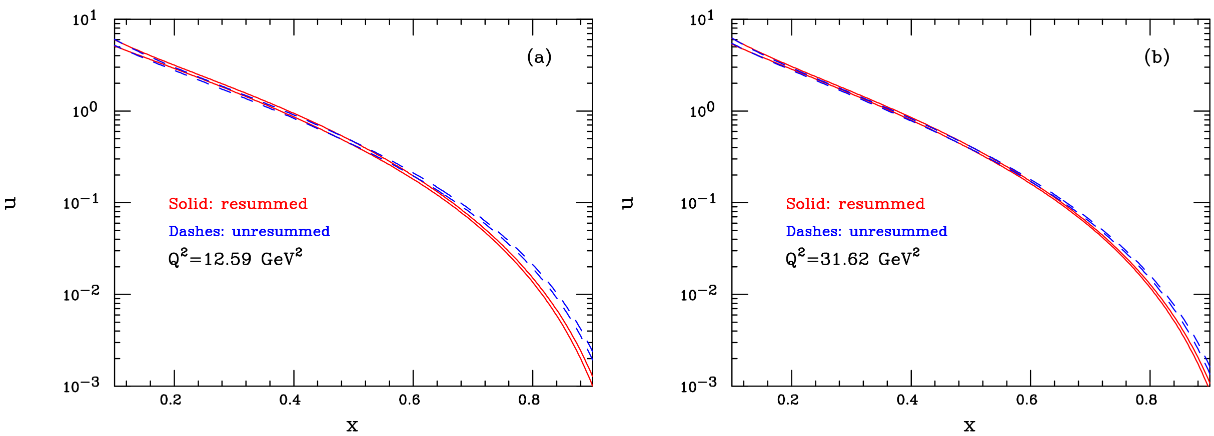


Fig. 5: NLO and resummed up quark distribution at $Q^2 = 12.59$ GeV^2 (a) and at $Q^2 = 31.62$ GeV^2, using the parametrization given in the text. The band corresponds to one standard deviation in parameter space.

In order to present the effect more clearly, we show in Fig. 6 the normalized deviation of the NLL-resummed prediction from the NLO one, i.e. $\Delta u(x) = (u_{\rm NLO}(x) - u_{\rm res}(x))\,/u_{\rm NLO}(x)$, at the two chosen values of Q^2 and for the central values of the best-fit parameters. We note a change in the sign of Δu in the neighborhhod of the point $x = 1/2$: although our errors are too large for the effect to be statistically significant, it is natural that the suppression of the quark distribution at large x be compensated by an enhancement at smaller x. In fact, the first moment of the coefficient function is

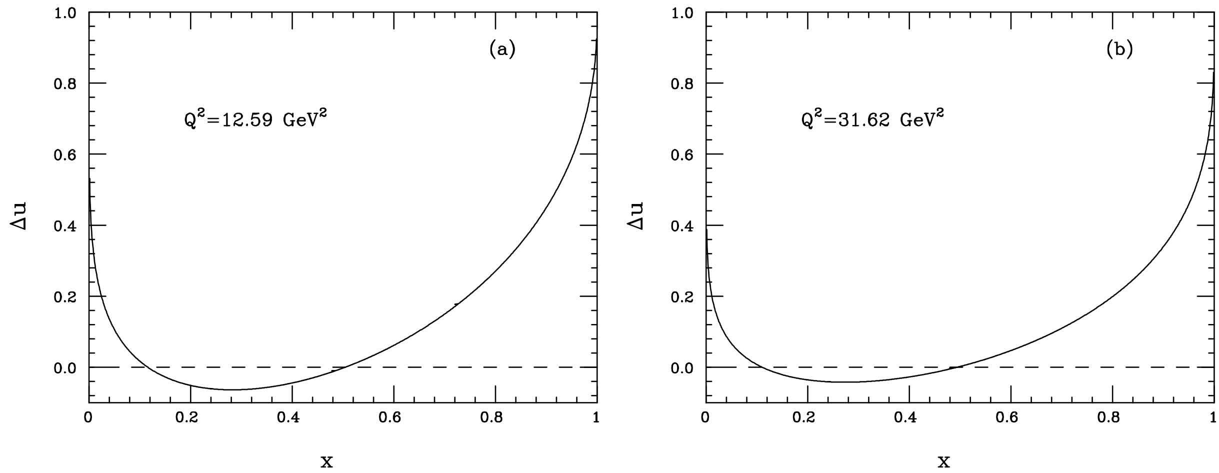


Fig. 6: Central value for the relative change in the up quark distribution, $\Delta u(x) \equiv (u_{\rm NLO}(x) - u_{\rm res}(x))\,/u_{\rm NLO}(x)$, at $Q^2 = 12.59$ (a) and 31.62 GeV^2 (b).

unaffected by the resummation: thus C_i^q, being larger at large x, must become smaller at small x. The further sign change at $x \sim 0.1$, on the other hand, should not be taken too seriously, since our sample includes essentially no data at smaller x, and of course we are using an x-space parametrization of limited flexibility.

Finally, we wish to verify that the up-quark distributions extracted by our fits at $Q^2 = 12.59$ and 31.62 GeV^2 are consistent with perturbative evolution. To achieve this goal, we evolve our N-space results at $Q^2 = 31.62$ GeV^2 down to 12.59 GeV^2, using NLO Altarelli–Parisi anomalous dimensions, and compare the evolved moments with the direct fit at 12.59 GeV^2. Figures 7 and 8 show that the results of our fits at 12.59 GeV^2 are compatible with the NLO evolution within the confidence level of one standard deviation. Note however that the evolution of resummed moments appears to give less consistent results, albeit within error bands: this can probably be ascribed to a contamination between pertubative resummation and power corrections, which we have not disentangled in our analysis.

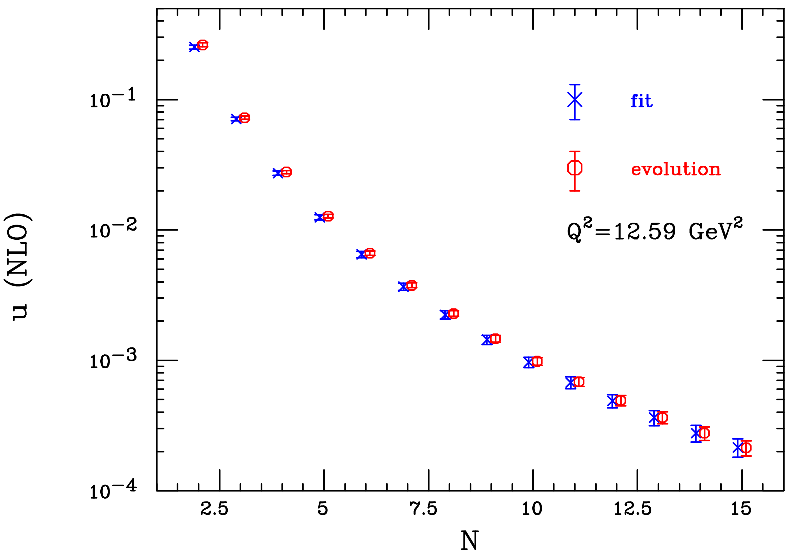


Fig. 7: Comparison of fitted moments of the NLO up quark distribution, at $Q^2 = 12.59$ GeV^2, with moments obtained via NLO evolution from $Q^2 = 31.62$ GeV^2.

Qualitatively, the observed effect on the up quark distribution is easily described, at least within the

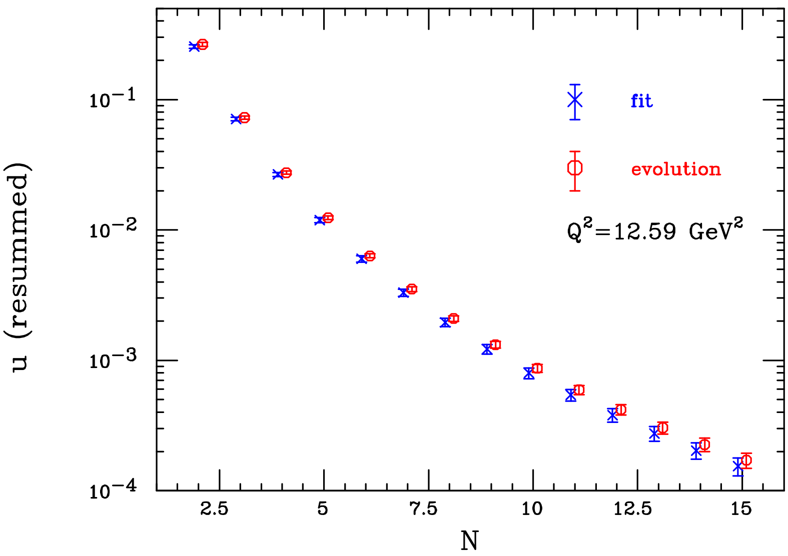


Fig. 8: As in Fig. 7, but comparing NLL-resummed moments of the up quark density.

limits of a simple parametrization like the one we are employing: resummation increases the exponent δ, responsible for the power-law decay of the distribution at large x, by about 10% to 15% at moderate Q^2. The exponent γ, governing the small-x behavior, and the normalization D, are then tuned so that the first finite moment (the momentum sum rule) may remain essentially unaffected.

In conclusion, our results indicate that quark distributions are suppressed at large x by soft gluon effects. Quantitatively, we observe an effect ranging between 10% and 20% when $0.6 < x < 0.8$ at moderate Q^2, where we expect power corrections not to play a significant role. Clearly, a more detailed quantitative understanding of the effect can be achieved only in the context of a broader and fully consistent fit. We would like however to notice two things: first, the effect of resummations propagates to smaller values of x, through the fact that the momentum sum rule is essentially unaffected by the resummation; similarly, evolution to larger values of Q^2 will shift the Sudakov suppression to smaller x. A second point is that, in a fully consistent treatment of hadronic cross section, there might be a partial compensation between the typical Sudakov enhancement of the partonic process and the Sudakov suppression of the quark distribution: the compensation would, however, be channel-dependent, since gluon-initiated partonic processes would be unaffected. We believe it would be interesting, and phenomenologically relevant, to investigate these issues in the context of a more comprehensive parton fit.

3 Small x

Small x structure functions are dominated by the flavour singlet contribution, whose coefficient functions and anomalous dimensions receive logarithmic enhancements, which make perturbation theory converge more slowly. In the small x, i.e. high energy limit, the cross section is quasi-constant and characterised by the effective expansion parameter $\langle \alpha_s(\boldsymbol{k}^2)\rangle \log \frac{1}{x} \log \frac{\boldsymbol{k}^2_{\rm max}}{\boldsymbol{k}^2_{\rm min}}$, where $x = Q^2/s$, $\boldsymbol{k}^2 \lesssim Q^2$ is the transverse momentum of the exchanged gluon, s is the photon-proton centre of mass energy squared and Q^2 is the hard scale. Such expansion parameter can be large, due to both the double-logs and to the fact that $\langle \boldsymbol{k}^2 \rangle$ may drift towards the non-perturbative region. Even assuming that truly non-perturbative effects are factored out — as is the case for structure functions — the problem remains of resumming the perturbative series with both kinds of logarithms [11, 12, 13, 14, 15, 16, 17]

In the BFKL approach one tries to resum the high-energy logarithms first, by an evolution equation in $\log 1/x$, whose $\boldsymbol{k}$-dependent evolution kernel is calculated perturbatively in α_s. However, the leading kernel [14, 15, 16, 17] overestimates the hard cross-section, and subleading ones [49, 19, 20] turn out to be large and of alternating sign, pointing towards an instability of the leading-log x (Lx) hierarchy. The problem is that, for any given value of the hard scales $Q, Q_0 \ll \sqrt{s}$ — think, for definiteness, of $\gamma^*(Q)$-$\gamma^*(Q_0)$ collisions —, the contributing kernels contain collinear enhancements in all $\boldsymbol{k}$-orderings of the exchanged gluons of type $\sqrt{s} \gg \cdots \boldsymbol{k}_1 \gg \boldsymbol{k}_2 \cdots$, or $\sqrt{s} \gg \cdots \boldsymbol{k}_2 \gg \boldsymbol{k}_1 \cdots$ and so on, to all orders in α_s. Such enhancements are only partly taken into account by any given truncation of the Lx hierarchy, and they make it unstable. In the DGLAP evolution equation one resums collinear logarithms first, but fixed order splitting functions do contain [6, 18] high-energy logarithms also, and a further resummation is needed.

Two approaches to the simultaneous resummation of these two classes of logs have recently reached the stage where their phenomenological application can be envisaged. The renormalisation group improved (CCSS) approach [21, 22, 23, 50] is built up within the BFKL framework, by improving the whole hierarchy of subleading kernels in the collinear region, so as to take into account all the $\boldsymbol{k}$-orderings mentioned before, consistently with the RG. In the duality (ABF) approach [51, 24, 25, 26, 27, 28, 29, 30] one concentrates on the problem of obtaining an improved anomalous dimension (splitting function) for DIS which reduces to the ordinary perturbative result at large N (large x), thereby automatically satisfying renormalization group constraints, while including resummed BFKL corrections at small N (small x), determined through the renormalization-group improved (i.e. running coupling) version of the BFKL kernel.

We will briefly review the theoretical underpinnings of these two approaches in turn, and then compare phenomenological results obtained in both approaches. Note that we shall use the notation of the CCSS or ABF papers in the corresponding sections, in order to enable a simpler connection with the original literature, at the price of some notational discontinuity. In particular, $\ln \frac{1}{x}$ is called Y by CCSS and ξ by ABF; the Mellin variable conjugate to $\ln \frac{1}{x}$ is called ω by CCSS and N by ABF; and the Mellin variable conjugated to $\ln \frac{Q^2}{\boldsymbol{k}^2}$ is called γ by CCSS and M by ABF.

3.1 The renormalisation group improved approach

The basic problem which is tackled in the CCSS approach [21, 22, 23, 50] is the calculation of the (azimuthally averaged) gluon Green function $\mathcal{G}(Y; k, k_0)$ as a function of the magnitudes of the external gluon transverse momenta $k \equiv |\boldsymbol{k}|,\ k_0 \equiv |\boldsymbol{k}_0|$ and of the rapidity $Y \equiv \log \frac{s}{kk_0}$. This is not yet a hard cross section, because one needs to incorporate the impact factors of the probes [52, 53, 54, 55, 56, 57, 58, 59]. Nevertheless, the Green function exhibits most of the physical features of the hard process, if we think of $k^2,\ k_0^2$ as external (hard) scales. The limits $k^2 \gg k_0^2$ ($k_0^2 \gg k^2$) correspond conventionally to the ordered (anti-ordered) collinear limit. By definition, in the ω-space conjugate to Y (so that $\hat{\omega} = \partial_Y$) one sets

$$\mathcal{G}_\omega(\boldsymbol{k}, \boldsymbol{k}_0) \equiv [\omega - \mathcal{K}_\omega]^{-1}(\boldsymbol{k}, \boldsymbol{k}_0)\,, \tag{13}$$

$$\omega \mathcal{G}_\omega(\boldsymbol{k}, \boldsymbol{k}_0) = \delta^2(\boldsymbol{k} - \boldsymbol{k}_0) + \int d^2\boldsymbol{k}'\ \mathcal{K}_\omega(\boldsymbol{k}, \boldsymbol{k}')\mathcal{G}_\omega(\boldsymbol{k}', \boldsymbol{k}_0)\ , \tag{14}$$

where $\mathcal{K}_\omega(\boldsymbol{k}, \boldsymbol{k}')$ is a kernel to be defined, whose $\omega = 0$ limit is related to the BFKL Y-evolution kernel discussed before.

In order to understand the RG constraints, it is useful to switch from $\boldsymbol{k}$-space to γ-space, where the variable γ is conjugated to $t \equiv \log \boldsymbol{k}^2/\boldsymbol{k}_0^2$ at fixed Y, and to make the following kinematical remark: the ordered (anti-ordered) region builds up scaling violations in the Bjorken variable $x = \boldsymbol{k}^2/s$ ($x_0 = \boldsymbol{k}_0^2/s$) and, if x (x_0) is fixed instead of $kk_0/s = e^{-Y}$, the variable conjugated to t is shifted [60] by an ω-dependent amount, and becomes $\gamma + \frac{\omega}{2} \sim \partial_{\log \boldsymbol{k}^2}$ $\ (1 - \gamma + \frac{\omega}{2} \sim \partial_{\log \boldsymbol{k}_0^2})$. Therefore, the characteristic function $\chi_\omega(\gamma)$ of $\mathcal{K}_\omega$ (with a factor α_s factored out) must be singular when either one of the variables is small, as shown (in the frozen α_s limit) by

$$\frac{1}{\omega}\chi_\omega(\gamma) \to \left[\frac{1}{\gamma+\frac{\omega}{2}} + \frac{1}{1-\gamma+\frac{\omega}{2}} + \cdots\right]\left[\gamma_{gg}^{(1)}(\alpha_s,\omega) + \cdots\right] , \tag{15}$$

where $\gamma_{gg}^{(1)}$ is the one-loop gluon anomalous dimension, and further orders may be added. Eq. (15) ensures the correct DGLAP evolution in either one of the collinear limits (because, e.g., $\gamma + \frac{\omega}{2} \sim \partial_{\log \boldsymbol{k}^2}$) and is ω-dependent, because of the shifts. Since higher powers of ω are related to higher subleading powers of α_s [61], this ω-dependence of the constraint (15) means that the whole hierarchy of subleading kernels is affected.

To sum up, the kernel $\mathcal{K}_\omega$ is constructed so as to satisfy the RG constraint (15) and to reduce to the exact Lx + NLx BFKL kernels in the $\omega \to 0$ limit; it is otherwise interpolated on the basis of various criteria (e.g., momentum conservation), which involve a "scheme" choice.

The resulting integral equation has been solved in [21, 22, 23] by numerical matrix evolution methods in $\boldsymbol{k}$- and x-space. Furthermore, introducing the integrated gluon density g, the resummed splitting function $P_{\text{eff}}(x, Q^2)$ is defined by the evolution equation

$$\frac{\partial g(x,Q^2)}{\partial \log Q^2} = \int \frac{\mathrm{d}z}{z}\, P_{\text{eff}}\big(z, \alpha_s(Q^2)\big)\, g\Big(\frac{x}{z}, Q^2\Big) , \tag{16}$$

and has been extracted [21, 22, 23] by a numerical deconvolution method [62]. Note that in the RGI approach the running of the coupling is treated by adopting in (14) the off-shell dependence of α_s suggested by the BFKL and DGLAP limits, and then solving the ensuing integral equation numerically.

It should be noted that the RGI approach has the somewhat wider goal of calculating the off-shell gluon density (13), not only its splitting function. Therefore, a comparison with the ABF approach, to be discussed below, is possible in the "on-shell" limit, in which the homogeneous (eigenvalue) equation of RGI holds. In the frozen coupling limit we have simply

$$\chi_\omega(\alpha_s, \gamma - \tfrac{\omega}{2}) = \omega , \qquad (\chi_\omega \text{ is at scale } kk_0) . \tag{17}$$

Solving Eq. (17) for either ω or γ, we are able to identify the effective characteristic function and its dual anomalous dimension

$$\omega = \chi_{\text{eff}}(\alpha_s, \gamma) ; \quad \gamma = \gamma_{\text{eff}}(\alpha_s, \omega) , \tag{18}$$

in the same spirit as the ABF approach [51, 24, 25, 26, 27, 28, 29, 30].

3.2 The duality approach

As already mentioned, in the ABF approach one constructs an improved anomalous dimension (splitting function) for DIS which reduces to the ordinary perturbative result at large N (large x) given by:

$$\gamma(N, \alpha_s) = \alpha_s \gamma_0(N) \; + \; \alpha_s^2 \gamma_1(N) \; + \; \alpha_s^3 \gamma_2(N) \; \ldots . \tag{19}$$

while including resummed BFKL corrections at small N (small x) which are determined by the aforementioned BFKL kernel $\chi(M, \alpha_s)$:

$$\chi(M, \alpha_s) = \alpha_s \chi_0(M) \; + \; \alpha_s^2 \chi_1(M) \; + \; \ldots , \tag{20}$$

which is the Mellin transform of the $\omega \to 0$, angular averaged kernel $\mathcal{K}$ eq. 14 with respect to $t = \ln \frac{k^2}{k_0^2}$. The main theoretical tool which enables this construction is the duality relation between the kernels χ and γ

$$\chi(\gamma(N, \alpha_s), \alpha_s) = N, \tag{21}$$

[compare eq. (18)] which is a consequence of the fact that the solutions of the BFKL and DGLAP equations coincide at leading twist [63, 51, 24]. Further improvements are obtained exploiting the symmetry under gluon interchange of the BFKL gluon-gluon kernel and through the inclusion of running coupling effecs.

By using duality, one can construct a more balanced expansion for both γ and χ, the "double leading" (DL) expansion, where the information from χ is used to include in γ all powers of α_s/N and, conversely γ is used to improve χ by all powers of α_s/M. A great advantage of the DL expansion is that it resums the collinear poles of χ at $M = 0$, enabling the imposition of the physical requirement of momentum conservation $\gamma(1, \alpha_s) = 0$, whence, by duality:

$$\chi(0, \alpha_s) = 1. \tag{22}$$

This procedure eliminates in a model independent way the alternating sign poles $+1/M, -1/M^2, \ldots..$ that appear in χ_0, χ_1,.... These poles make the perturbative expansion of χ unreliable even in the central region of M: e.g., $\alpha_s\chi_0$ has a minimum at $M = 1/2$, while, at realistic values of α_s, $\alpha_s\chi_0 + \alpha_s^2\chi_1$ has a maximum.

At this stage, while the poles at $M = 0$ are eliminated, those at $M = 1$ remain, so that the DL expansion is still not finite near $M = 1$. The resummation of the $M = 1$ poles can be accomplished by exploiting the collinear-anticollinear symmetry, as suggested in the CCSS approach discussed above. In Mellin space, this symmetry implies that at the fixed-coupling level the kernel χ for evolution in $\ln \frac{s}{kk_0}$ must satisfy $\chi(M) = \chi(1 - M)$. This symmetry is however broken by the DIS choice of variables $\ln \frac{1}{x} = \ln \frac{s}{Q^2}$ and by the running of the coupling. In the fixed coupling limit the kernel $\chi_{\rm DIS}$, dual to the DIS anomalous dimension, is related to the symmetric one χ_σ through the implicit equation [49]

$$\chi_{\rm DIS}(M + 1/2\chi_\sigma(M)) = \chi_\sigma(M), \tag{23}$$

to be compared to eq. (17) of the CCSS approach.

Hence, the $M = 1$ poles can be resummed by performing the double-leading resummation of $M = 0$ poles of $\chi_{\rm DIS}$, determining the associated χ_σ through eq. (23), then symmetrizing it, and finally going back to DIS variables by using eq. (23) again in reverse. Using the momentum conservation eq. (22) and eq. (23), it is easy to show that $\chi_\sigma(M)$ is an entire function of M, with $\chi_\sigma(-1/2) = \chi_\sigma(3/2) = 1$ and has a minimum at $M = 1/2$. Through this procedure one obtains order by order from the DL expansion a symmetrized DL kernel $\chi_{\rm DIS}$, and its corresponding dual anomalous dimension γ. The kernel $\chi_{\rm DIS}$ has to all orders a minimum and satisfies a momentum conservation constraint $\chi_{\rm DIS}(0) = \chi_{\rm DIS}(2) = 1$.

The final ingredient of the ABF approach is a treatment of the running coupling corrections to the resummed terms. Indeed, their inclusion in the resummed anomalous dimension greatly softens the asymptotic behavior near $x = 0$. Hence, the dramatic rise of structure functions at small x, which characterized resummations based on leading–order BFKL evolution, and is ruled out phenomenologically, is replaced by a much milder rise. This requires a running coupling generalization of the duality equation (21), which is possible noting that in M space the running coupling $\alpha_s(t)$ becomes a differential operator, since $t \to d/dM$. Hence, the BFKL evolution equation for double moments $G(N, M)$, which is an algebraic equation at fixed coupling, becomes a differential equation in M for running coupling. In the ABF approach, one solves this differential equation analytically when the kernel is replaced by its quadratic approximation near the minimum. The solution is expressed in terms of an Airy function if the kernel is linear in α_s, for example in the case of $\alpha_s\chi_0$, or of a Bateman function in the more general case of a non linear dependence on α_s as is the case for the DL kernels. The final result for the improved anomalous dimension is given in terms of the DL expansion plus the "Airy" or "Bateman" anomalous dimension, with the terms already included in the DL expansion subtracted away.

For example, at leading DL order, i.e. only using $\gamma_0(N)$ and $\chi_0(M)$, the improved anomalous dimension is

$$\gamma_I^{NL}(\alpha_s, N) = \left[\alpha_s\gamma_0(N) + \alpha_s^2\gamma_1(N) + \gamma_s(\frac{\alpha_s}{N}) - \frac{n_c\alpha_s}{\pi N}\right] + \gamma_A(c_0, \alpha_s, N) - \frac{1}{2} + \sqrt{\frac{2}{\kappa_0\alpha_s}[N - \alpha_s c_0]}. \tag{24}$$

The terms within square brackets give the LO DL approximation, i.e. they contain the fixed–coupling information from γ_0 and (through γ_s) from χ_0. The "Airy" anomalous dimension $\gamma_A(c_0, \alpha_s, N)$ contains the running coupling resummation, i.e. it is the exact solution of the running coupling BFKL equation which corresponds to a quadratic approximation to χ_0 near $M = 1/2$. The last two terms subtract the contributions to $\gamma_A(c_0, \alpha_s, N)$ which are already included in γ_s and γ_0. In the limit $\alpha_s \to 0$ with N fixed, $\gamma_I(\alpha_s, N)$ reduces to $\alpha_s\gamma_0(N) + O(\alpha_s^2)$. For $\alpha_s \to 0$ with α_s/N fixed, $\gamma_I(\alpha_s, N)$ reduces to $\gamma_s(\frac{\alpha_s}{N}) + O(\alpha_s^2/N)$, i.e. the leading term of the small x expansion. Thus the Airy term is subleading in both limits. However, if $N \to 0$ at fixed α_s, the Airy term replaces the leading singularity of the DL anomalous dimension, which is a square root branch cut, with a simple pole, located on the real axis at rather smaller N, thereby softening the small x behaviour. The quadratic approximation is sufficient to give the correct asymptotic behaviour up to terms which are of subleading order in comparison to those included in the DL expression in eq. (24).

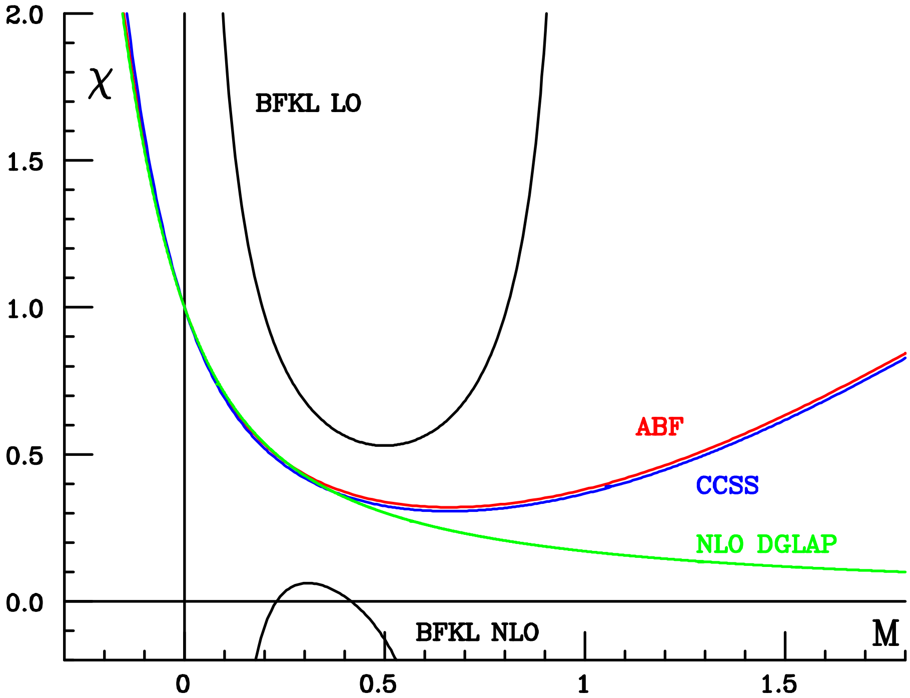


Fig. 9: The kernel χ (BFKL characteristic function) for fixed coupling ($\beta_0 = 0$) $\alpha_s = 0.2$ and $n_f = 0$. The BFKL curves are the LO and NLO truncations of eq. (20), the DGLAP curve is the dual eq. (21) of the NLO anomalous dimension eq. (19), while the CCSS and ABF curves are respectively the solution ω of eq. (17) and the solution $\chi_{\rm DIS}$ of eq. (23).

The running coupling resummation procedure can be applied to a symmetrized kernel, which possesses a minimum to all orders, and then extended to next-to-leading order [29, 30]. This entails various technical complications, specifically related to the nonlinear dependence of the symmetrized kernel on α_s, to the need to include interference between running coupling effects and the small x resummation, and to the consistent treatment of next-to-leading log Q^2 terms, in particular those related to the running of the coupling. It should be noted that even though the ABF appraoch is limited to the description of leading-twist evolution at zero-momentum transfer, it leads to a pair of systematic dual perturbative expansions for the χ and γ kernels. Hence, comparison with the CCSS approach is possible for instance by comparing the NLO ABF kernel to the RG improved Lx+NLx CCSS kernel.

3.3 Comparison of results

Even though the basic underlying physical principles of the CCSS and ABF approaches are close, there are technical differences in the construction of the resummed RG-improved (CCSS) or symmetrized DL (ABF) kernel, in the derivation from it of an anomalous dimension and associated splitting function, and in the inclusion of running coupling effects. Therefore, we will compare results for the resummed fixed-coupling χ kernel (BFKL characteristic function), then the corresponding fixed-coupling splitting functions, and finally the running coupling splitting functions which provide the final result in both approaches. In order to assess the phenomenological impact on parton evolution we will finally compare the convolution of the splitting function with a "typical" gluon distribution.

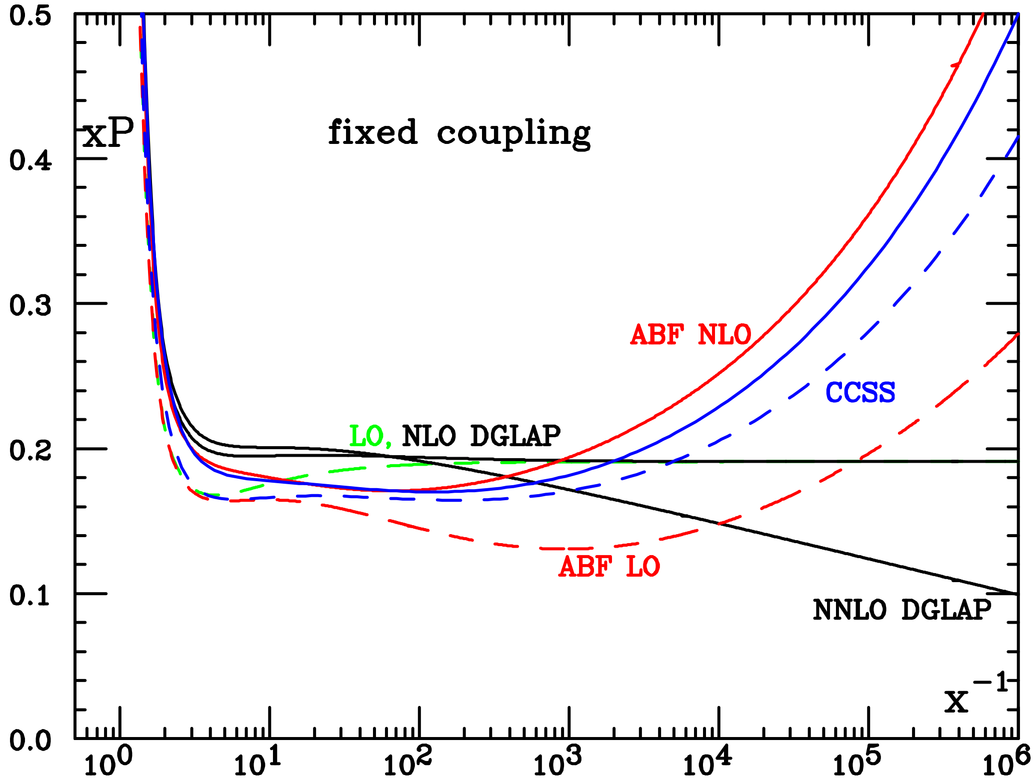


Fig. 10: The fixed coupling ($\beta_0 = 0$) $xP_{gg}(x)$ splitting function, evaluated with $\alpha_s = 0.2$ and $n_f = 0$. The dashed curves are LO for DGLAP, NLx+LO for CCSS and symmetrized LO DL for ABF, while the solid curves are NLO and NNLO for DGLAP, NLx+NLO for CCSS and symmetrized NLO DL for ABF.

In fig. 9 we compare the solution, ω, to the on-shell constraint, eq. (17) for the RGI CCSS result, and the solution $\chi_{\rm DIS}$ of eq. (23) for the symmetrized NLO DL ABF result. The pure Lx and NLx (BFKL) and next-to-leading $\ln Q^2$ (DGLAP) are also shown. All curves are determined with frozen coupling ($\beta_0 = 0$), and with $n_f = 0$, in order to avoid complications related to the diagonalization of the DGLAP anomalous dimension matrix and to the choice of scheme for the quark parton distribution. The resummed CCSS and ABF results are very close, in that they coincide by construction at the momentum conservation points $M = \frac{1}{2}$ and $M = 2$, and differ only in the treatment of NLO DGLAP terms. In comparison to DGLAP, the resummed kernels have a minimum, related to the fact that both collinear and anticollinear logs are resummed. In comparison to BFKL, which has a minimum at LO but not NLO, the resummed kernels always have a perturbatively stable minimum, characterized by a lower intercept than leading–order BFKL: specifically, when $\alpha_s = 0.2$, $\lambda \sim 0.3$ instead of $\lambda \sim 0.5$. This corresponds to a softer small x rise of the associated splitting function.

The fixed–coupling resummed splitting functions up to NLO are shown in figure 10, along with the unresummed DGLAP splitting functions up to NNLO.[2] In the CCSS approach the splitting function is determined by explicitly solving eq. (14) with the kernel corresponding to figure 9, and then applying the numerical deconvolution procedure of [62]. For $n_f = 0$ the NLO DGLAP splitting function has the property that it vanishes at small x — this makes it relatively straightforward to combine not just LO DGLAP but also NLO DGLAP with the NLLx resummation. Both the CCSS NLx+LO and NLx+NLO curves are shown in figure 10. On the other hand, in the ABF approach the splitting function is the inverse Mellin transform of the anomalous dimension obtained using duality eq. (21) from the symmetrized DL χ kernel. Hence, the LO and NLO resummed result respectively reproduce all information contained in the LO and NLO χ and γ kernel with the additional constraint of collinear-anticollinear symmetry. Both the ABF LO and NLO results are shown in figure 10.

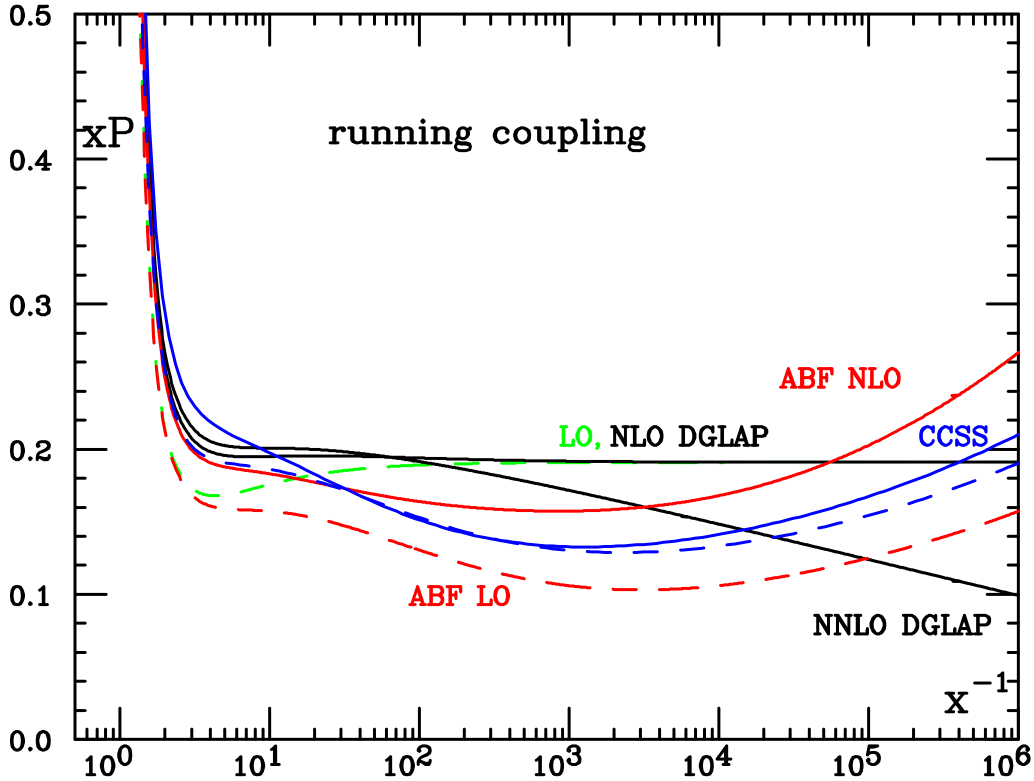


Fig. 11: The running coupling $xP_{gg}(x)$ splitting function, evaluated with $\alpha_s = 0.2$ and $n_f = 0$. The various curves correspond to the same cases as in figure 10.

In comparison to unresummed results, the resummed splitting functions display the characteristic rise at small x of fixed-coupling leading-order BFKL resummation, though the small x rise is rather milder ($\sim x^{-0.3}$ instead of $\sim x^{-0.5}$ for $\alpha_s = 0.2$). At large x there is good agreement between the resummed results and the corresponding LO (dashed) or NLO (solid) DGLAP curves. At small x the difference between the ABF LO and CCSS NLx+LO (dashed) curves is mostly due to the inclusion in CCSS of BFKL NLx terms, as well as to differences in the symmetrization procedure. When comparing CCSS NLx+NLO with ABF NLO this difference is reduced, and , being only due the way the symmetrization is implemented, it might be taken as an estimate of the intrinsic ambiguity of the fixed–coupling resummation procedure. At intermediate x the NLO resummed splitting functions is of a similar order of magnitude as the NLO DGLAP result even down to quite small x, but with a somewhat different shape, characterized by a shallow dip at $x \sim 10^{-2}$, until the small x rise sets in for $x \sim 10^{-3}$. It has

[2] Starting from NLO one needs also to specify a factorisation scheme. Small-x results are most straightforwardly obtained in the Q_0 scheme, while fixed-order splitting functions are quoted in the $\overline{\mathrm{MS}}$ scheme (for discussions of the relations between different schemes see [64, 65, 25, 50]).

been suggested [66] that in the small α_s limit this dip can be explained as a consequence of the interplay between the $-{\alpha_s}^3 \ln x$ NNLO term of xP_{gg} (also present in the resummation) and the first positive resummation effects which start with an ${\alpha_s}^4 \ln^3 1/x$ term. The unstable small x drop of the NNLO DGLAP result appears to be a consequence of the unresummed $\frac{\alpha_s^3}{N^2}$ double pole in the NNLO anomalous dimension.

The running-coupling resummed splitting functions are displayed in figure 11. Note that the unresummed curves are the same as in the fixed coupling case since their dependence on α_s is just through a prefactor of α_s^k, whereas in the resummed case there is an interplay between the running of the coupling and the structure of the small-x logs. All the resummed curves display a considerable softening of the small x behaviour in comparison to their fixed-coupling counterparts, due to the softening of the leading small x singularity in the running-coupling case [21, 26]. As a consequence, the various resummed results are closer to each other than in the fixed-coupling case, and also closer to the unresummed LO and NLO DGLAP results. The resummed perturbative expansion appears to be stable, subject to moderate theoretical ambiguity, and qualitatively close to NLO DGLAP.

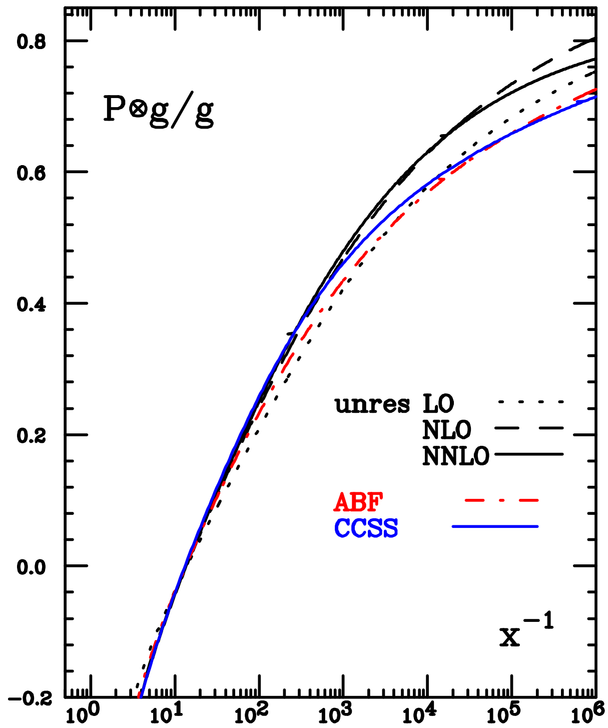


Fig. 12: Convolution of resummed and fixed-order P_{gg} splitting functions with a toy gluon distribution, eq. (25), normalised to the gluon distribution itself, with $\alpha_s = 0.2$ and $n_f = 0$. The resummed CCSS and ABF curves are obtained using respectively the CCSS NLx+NLO and the ABF NLO splitting function shown in fig. 11.

Finally, to appreciate the impact of resummation it is useful to investigate not only the properties of the splitting function, but also its convolution with a physically reasonable gluon distribution. We take the following toy gluon

$$xg(x) = x^{-0.18}(1-x)^5 \,, \tag{25}$$

and show in fig. 12 the result of its convolution with various splitting functions of fig. 11. The differences between resummed and unresummed results, and between the CCSS and ABF resummations are partly washed out by the convolution, even though the difference between the unresummed LO and NLO DGLAP results is clearly visible. In particular, differences between the fixed-order and resummed convolution start to become significant only for $x \lesssim 10^{-2} - 10^{-3}$, even though resummation effects started to be visible in the splitting functions at somewhat larger x.

It should be kept in mind that it is only the gg entry of the singlet splitting function matrix that has so far been investigated at this level of detail and that the other entries may yet reserve surprises.

3.4 Explicit solution of the BFKL equation by Regge exponentiation

The CCSS approach of section 3.1 exploits a numerical solution of the BFKL equation in which the gluon Green's function is represented on a grid in x and k. This method provides an efficient determination of the azimuthally averaged Green's function and splitting functions — for percent accuracy, up to $Y = 30$, it runs in a few seconds — for a wide range of physics choices, e.g. pure NLx, various NLx+NLO schemes. Here we will discuss an alternative framework suitable to solve numerically the NLL BFKL integral equation [67], based on Monte Carlo generation of events, which can also be applied to the study of different resummation schemes and DIS, but so far has been investigated for simpler NLL BFKL kernels and Regge–like configurations. This method has the advantage that it automatically provides information about azimuthal decorrelations as well as the pattern of final-state emissions.

This appproach relies on the fact that, as shown in Ref. [67], it is possible to trade the simple and double poles in ϵ, present in $D = 4 + 2\epsilon$ dimensional regularisation, by a logarithmic dependence on an effective gluon mass λ. This λ dependence numerically cancels out when the full NLL BFKL evolution is taken into account for a given center–of–mass energy, a consequence of the infrared finiteness of the full kernel. The introduction of this mass scale, differently to the original work of Ref. [49] was performed without angular averaging the NLL kernel.

With such reguralisation of the infrared divergencies it is then convenient to iterate the NLL BFKL equation for the t–channel partial wave, generating, in this way, multiple poles in the complex ω–plane. The positions of these singularities are set at different values of the gluon Regge trajectory depending on the transverse momenta of the Reggeized gluons entering the emission vertices. At this point it is possible to Mellin transform back to energy space and obtain an iterated form for the solution of the NLL BFKL equation:

$$\begin{aligned} f(\mathbf{k}_a, \mathbf{k}_b, \mathrm{Y}) &= e^{\omega_0^\lambda(\mathbf{k}_a)\mathrm{Y}}\, \delta^{(2)}(\mathbf{k}_a - \mathbf{k}_b) \\ &+ \sum_{n=1}^{\infty} \prod_{i=1}^{n} \int d^2\mathbf{k}_i \int_0^{y_{i-1}} dy_i \left[\frac{\theta\left(\mathbf{k}_i^2 - \lambda^2\right)}{\pi \mathbf{k}_i^2} \xi\left(\mathbf{k}_i\right) + \widetilde{\mathcal{K}}_r \left(\mathbf{k}_a + \sum_{l=0}^{i-1} \mathbf{k}_l, \mathbf{k}_a + \sum_{l=1}^{i} \mathbf{k}_l \right) \right] \\ &\times e^{\omega_0^\lambda\left(\mathbf{k}_a + \sum_{l=1}^{i-1}\mathbf{k}_l\right)(y_{i-1} - y_i)}\, e^{\omega_0^\lambda\left(\mathbf{k}_a + \sum_{l=1}^{i}\mathbf{k}_l\right) y_n}\, \delta^{(2)}\left(\sum_{l=1}^{n} \mathbf{k}_l + \mathbf{k}_a - \mathbf{k}_b \right), \end{aligned} \tag{26}$$

where the strong ordering in longitudinal components of the parton emission is encoded in the nested integrals in rapidity with an upper limit set by the logarithm of the total energy in the process, $y_0 = \mathrm{Y}$. The first term in the expansion corresponds to two Reggeized gluons propagating in the t–channel with no additional emissions. The exponentials carry the dependence on the Regge gluon trajectory, *i.e.*

$$\omega_0^\lambda\left(\mathbf{q}\right) \;=\; -\bar{\alpha}_s \ln \frac{\mathbf{q}^2}{\lambda^2} + \frac{\bar{\alpha}_s^2}{4} \left[\frac{\beta_0}{2N_c} \ln \frac{\mathbf{q}^2}{\lambda^2} \ln \frac{\mathbf{q}^2 \lambda^2}{\mu^4} + \left(\frac{\pi^2}{3} - \frac{4}{3} - \frac{5}{3}\frac{\beta_0}{N_c} \right) \ln \frac{\mathbf{q}^2}{\lambda^2} + 6\zeta(3) \right], \tag{27}$$

corresponding to no–emission probabilities between two consecutive effective vertices. Meanwhile, the real emission is built out of two parts, the first one:

$$\xi\left(\mathrm{X}\right) \;\equiv\; \bar{\alpha}_s + \frac{\bar{\alpha}_s^2}{4} \left(\frac{4}{3} - \frac{\pi^2}{3} + \frac{5}{3}\frac{\beta_0}{N_c} - \frac{\beta_0}{N_c} \ln \frac{\mathrm{X}}{\mu^2} \right), \tag{28}$$

which cancels the singularities present in the trajectory order by order in perturbation theory, and the second one: $\widetilde{\mathcal{K}}_r$, which, although more complicated in structure, does not generate ϵ singularities when integrated over the full phase space of the emissions, for details see Ref. [67].

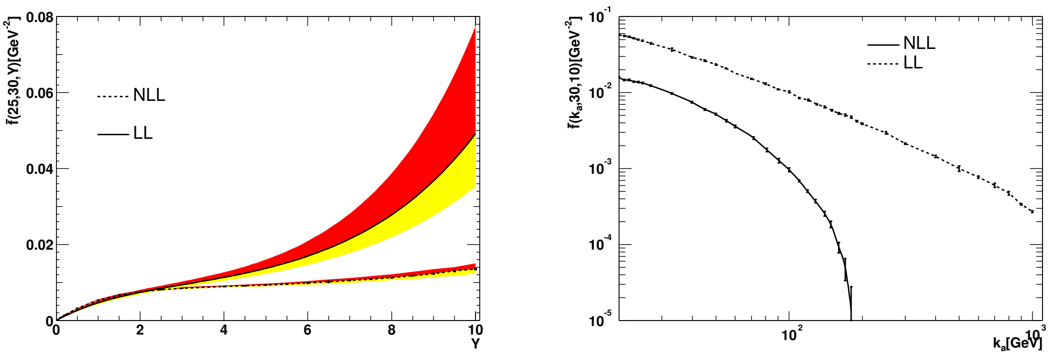


Fig. 13: Analysis of the gluon Green's function as obtained from the NLL BFKL equation. The plot to the left shows the evolution in rapidity of the gluon Green's function at LL and NLL for fixed $k_a = 25$ GeV and $k_b = 30$ GeV. The plot on the right hand side shows the dependence on k_a for fixed $k_b = 30$ GeV and $Y = 10$.

The numerical implementation and analysis of the solution as in Eq. (26) was performed in Ref. [68]. As in previous studies the intercept at NLL was proved to be lower than at leading–logarithmic (LL) accuracy. In this approach the kernel is not expanded on a set of functions derived from the LL eigenfunctions, and there are no instabilities in energy associated with a choice of functions breaking the $\gamma \leftrightarrow 1-\gamma$ symmetry, with γ being the variable Mellin–conjugate of the transverse momenta. This is explicitly shown at the left hand side of Fig. 13 where the coloured bands correspond to uncertainties from the choice of renormalisation scale. Since the exponential growth at NLL is slower than at LL, there is little overlap between the two predictions, and furthermore these move apart for increasing rapidities. The NLL corrections to the intercept amount to roughly 50% and are stable with increasing rapidities.

In transverse momentum space the NLL corrections are stable when the two transverse scales entering the forward gluon Green's function are of similar magnitude. However, when the ratio between these scales departs largely from unity, the perturbative convergence is poor, driving, as it is well–known, the gluon Green's function into an oscillatory behaviour with regions of negative values along the period of oscillation. This behaviour is demonstrated in the second plot of Fig 13.

The way the perturbative expansion of the BFKL kernel is improved by simultaneous resummation of energy and collinear logs has been discussed in sections 3.1,3.2. In particular, the original approach based on the introduction in the NLL BFKL kernel of an all order resummation of terms compatible with renormalisation group evolution described in ref. [60] (and incorporated in the CCSS approach of section 3.1) can be implemented in the iterative method here explained [69] (the method of ref. [60] was combined with the imposition of a veto in rapidities in refs. [70, 71, 72]). The main idea is that the solution to the ω–shift proposed in ref. [60]

$$\begin{aligned}\omega &= \bar{\alpha}_s \left(1+\left(\mathrm{a}+\frac{\pi^2}{6}\right)\bar{\alpha}_s\right)\left(2\psi(1)-\psi\left(\gamma+\frac{\omega}{2}-\mathrm{b}\,\bar{\alpha}_s\right)-\psi\left(1-\gamma+\frac{\omega}{2}-\mathrm{b}\,\bar{\alpha}_s\right)\right)\\ &+ \bar{\alpha}_s^2\left(\chi_1(\gamma)+\left(\frac{1}{2}\chi_0(\gamma)-\mathrm{b}\right)\left(\psi'(\gamma)+\psi'(1-\gamma)\right)-\left(\mathrm{a}+\frac{\pi^2}{6}\right)\chi_0(\gamma)\right), \qquad (29)\end{aligned}$$

can be very accurately approximated by the sum of the approximated solutions to the shift at each of the poles in γ of the LL eigenvalue of the BFKL kernel. This provides an effective "solution" of Eq. (29) of the form [69]

$$\begin{aligned}\omega &= \bar{\alpha}_s\chi_0(\gamma)+\bar{\alpha}_s^2\chi_1(\gamma)+\left\{\sum_{m=0}^{\infty}\left[\left(\sum_{n=0}^{\infty}\frac{(-1)^n(2n)!}{2^n n!(n+1)!}\frac{\left(\bar{\alpha}_s+\mathrm{a}\,\bar{\alpha}_s^2\right)^{n+1}}{(\gamma+m-\mathrm{b}\,\bar{\alpha}_s)^{2n+1}}\right.\right.\right.\\ &- \left.\left.\left.\frac{\bar{\alpha}_s}{\gamma+m}-\bar{\alpha}_s^2\left(\frac{\mathrm{a}}{\gamma+m}+\frac{\mathrm{b}}{(\gamma+m)^2}-\frac{1}{2(\gamma+m)^3}\right)\right)\right]+\{\gamma\to1-\gamma\}\right\}, \qquad (30)\end{aligned}$$

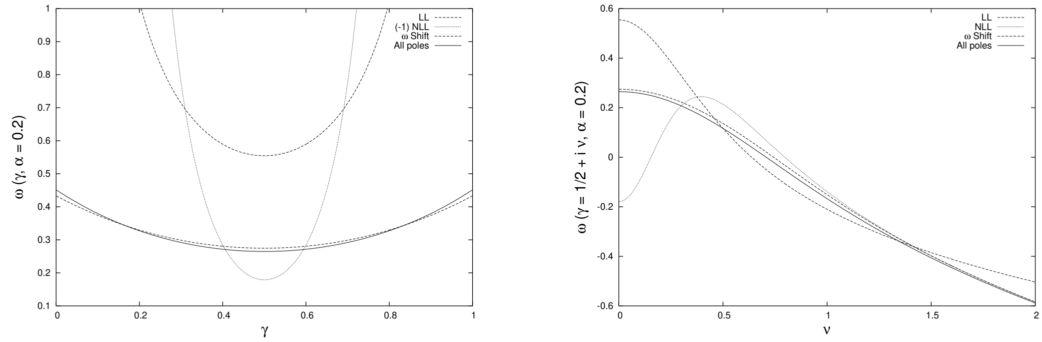


Fig. 14: The γ–representation of the LL and NLL scale invariant kernels together with the collinearly–improved kernel by an ω–shift and the "all–poles" resummation.

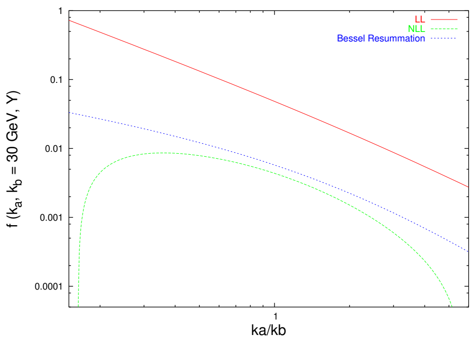


Fig. 15: The behaviour of the NLL gluon Green's function using the Bessel resummation.

where χ_0 and χ_1 are, respectively, the LL and NLL scale invariant components of the kernel in γ representation with the collinear limit

$$\chi_1(\gamma) \simeq \frac{\mathrm{a}}{\gamma} + \frac{\mathrm{b}}{\gamma^2} - \frac{1}{2\gamma^3}, \;\; \mathrm{a} = \frac{5}{12}\frac{\beta_0}{N_c} - \frac{13}{36}\frac{n_f}{N_c^3} - \frac{55}{36}, \;\; \mathrm{b} = -\frac{1}{8}\frac{\beta_0}{N_c} - \frac{n_f}{6N_c^3} - \frac{11}{12}. \tag{31}$$

The numerical solution to Eq. (29) and the value of expression (30) are compared in Fig. 14. The stability of the perturbative expansion is recovered in all regions of transverse momenta with a prediction for the intercept of 0.3 at NLL for $\bar{\alpha}_s = 0.2$, a result valid up to the introduction of scale invariance breaking terms. The implementation of expression (30) in transverse momentum space is simple given that the transverse components decouple from the longitudinal in this form of the collinear resummation [69]. The prescription is to remove the term $-\frac{\bar{\alpha}_s^2}{4}\ln^2\frac{q^2}{k^2}$ from the real emission kernel, $\mathcal{K}_r\left(\vec{q},\vec{k}\right)$, and replace it with

$$\left(\frac{q^2}{k^2}\right)^{-\mathrm{b}\bar{\alpha}_s\frac{|k-q|}{k-q}} \sqrt{\frac{2\left(\bar{\alpha}_s + \mathrm{a}\,\bar{\alpha}_s^2\right)}{\ln^2\frac{q^2}{k^2}}} J_1\left(\sqrt{2\left(\bar{\alpha}_s + \mathrm{a}\,\bar{\alpha}_s^2\right)\ln^2\frac{q^2}{k^2}}\right) - \bar{\alpha}_s - \mathrm{a}\,\bar{\alpha}_s^2 + \mathrm{b}\,\bar{\alpha}_s^2\frac{|k-q|}{k-q}\ln\frac{q^2}{k^2}, \tag{32}$$

with J_1 the Bessel function of the first kind. This prescription does not affect angular dependences and generates a well–behaved gluon Green's function as can be seen in Fig. 15 where the oscillations in the collinear and anticollinear regions of phase space are consistently removed. At present, work is in progress to study the effect of the running of the coupling in this analysis when the Bessel resummation is introduced in the iterative procedure of Ref. [67].

A great advantage of the iterative method here described is that the solution to the NLL BFKL equation is generated integrating the phase space using a Monte Carlo sampling of the different parton configurations. This allows for an investigation of the diffusion properties of the BFKL kernel as shown

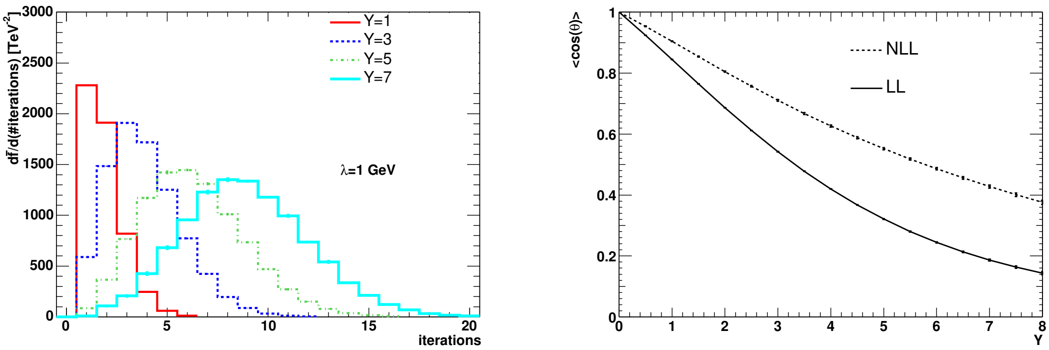


Fig. 16: Distribution in the number of iterations and angular dependence of the NLL gluon Green's function.

in ref. [73], and provides a good handle on the average multiplicities and angular dependences of the evolution. Multiplicities can be extracted from the Poisson–like distribution in the number of iterations of the kernel needed to reach a convergent solution, which is obtained numerically at the left hand side of Fig. 16 for a fixed value of the λ parameter. On the right hand side of the figure a study of the azimuthal angular correlation of the gluon Green's function is presented at $Y = 5$. This decorrelation will directly impact the prediction for the azimuthal angular decorrelation of two jets with a large rapidity separation, in a fully inclusive jet sample (i.e. no rapidity gaps). The increase of the angular correlation when the NLL terms are included is a characteristic feature of these corrections. This study is possible using this approach because the NLL kernel is treated in full, without angular averaging, so there is no need to use a Fourier expansion in angular variables.

References

[1] S. Moch et al., *Precision predictions for deep-inelastic scattering*. These proceedings.

[2] Sterman, G., Nucl. Phys. **B281**, 310 (1987).

[3] Catani, S. and Trentadue, L., Nucl. Phys. **B327**, 323 (1989).

[4] Forte, S. and Ridolfi, G., Nucl. Phys. **B650**, 229 (2003).

[5] Vogt, A., Phys. Lett. **B497**, 228 (2001).

[6] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. **B688**, 101 (2004).

[7] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. **B726**, 317 (2005).

[8] Moch, S. and Vogt, A., *Higher-order soft corrections to lepton pair and higgs boson production*. Preprint hep-ph/0508265, 2005.

[9] Laenen, E. and Magnea, L., *Threshold resummation for electroweak annihilation from dis data*. Preprint hep-ph/0508284, 2005.

[10] Eynck, T. O. and Laenen, E. and Magnea, L., JHEP **06**, 057 (2003).

[11] Gribov, V.N. and Lipatov, L.N., Sov.J.Nucl.Phys **15**, 438 (1972).

[12] Altarelli, G. and Parisi, G., Nucl.Phys. **B126**, 298 (1977).

[13] Dokshitzer, Yu. L., JHEP **46**, 641 (1977).

[14] Lipatov, L. N., Sov. J. Nucl. Phys. **23**, 338 (1976).

[15] Kuraev, E. A. and Lipatov, L. N. and Fadin, Victor S., Sov. Phys. JETP **45**, 199 (1977).

[16] Balitsky, I. I. and Lipatov, L. N., Sov. J. Nucl. Phys. **28**, 822 (1978).

[17] Lipatov, L. N., Sov. Phys. JETP **63**, 904 (1986).

[18] Vogt, A. and Moch, S. and Vermaseren, J. A. M., Nucl. Phys. **B691**, 129 (2004).

[19] Camici, G. and Ciafaloni, M., Phys. Lett. **B412**, 396 (1997).

[20] Ciafaloni, M. and Camici, G., Phys. Lett. **B430**, 349 (1998).

[21] Ciafaloni, M. and Colferai, D. and Salam, G. P., Phys. Rev. **D60**, 114036 (1999).

[22] Ciafaloni, M. and Colferai, D. and Colferai, D. and Salam, G. P. and Stasto, A. M., Phys. Lett. **B576**, 143 (2003).

[23] Ciafaloni, M. and Colferai, D. and Salam, G. P. and Stasto, A. M., Phys. Rev. **D68**, 114003 (2003).

[24] Altarelli, G. and Ball, R. D. and Forte, S., Nucl. Phys. **B575**, 313 (2000).

[25] Altarelli, G. and Ball, R. D. and Forte, S., Nucl. Phys. **B599**, 383 (2001).

[26] Altarelli, G. and Ball, R. D. and Forte, S., Nucl. Phys. **B621**, 359 (2002).

[27] Altarelli, G. and Ball, R. D. and Forte, S., Nucl. Phys. **B674**, 459 (2003).

[28] Altarelli, G. and Ball, R. D. and Forte, S., *An improved splitting function for small x evolution*. Preprint hep-ph/0310016, 2003.

[29] Altarelli, G. and Ball, R. D. and Forte, S., Nucl. Phys. Proc. Suppl. **135**, 163 (2004).

[30] Altarelli, G. and Ball, R. D. and Forte, S., *Perturbatively stable resummed small x kernels*. Preprint CERN-PH-TH/2005-174, 2005.

[31] Catani, S. and Mangano, Michelangelo L. and Nason, Paolo and Oleari, Carlo and Vogelsang, Werner, JHEP **03**, 025 (1999).

[32] Laenen, E. and Sterman, G. and Vogelsang, W., Phys. Rev. Lett. **84**, 4296 (2000).

[33] Kidonakis, N. and Sterman, G., Nucl. Phys. **B505**, 321 (1997).

[34] Kidonakis, N. and Owens, J. F., Phys. Rev. **D63**, 054019 (2001).

[35] Laenen, E. and Moch, S., Phys. Rev. **D59**, 034027 (1999).

[36] Corcella, G. and Mitov, A. D., Nucl. Phys. **B676**, 346 (2004).

[37] A. Cooper-Sarkar, C. Gwenlan, *Comparison and combination of zeus and h1 pdf analyses*. These proceedings.

[38] S. I. Alekhin, *Towards precise determination of the nucleon pdfs*. These proceedings.

[39] Alekhin, S., *Parton distribution functions from the precise nnlo qcd fit*. Preprint hep-ph/0508248, 2005.

[40] Tzanov, M. et al., *New qcd results from nutev*. Preprint hep-ex/0306035, 2003.

[41] Naples, D. et al., *Nutev cross section and structure function measurements*. Preprint hep-ex/0307005, 2003.

[42] Arneodo, M. et al., Nucl. Phys. **B483**, 3 (1997).

[43] Benvenuti, A. C. et al., Phys. Lett. **B223**, 485 (1989).

[44] Benvenuti, A. C. et al., Phys. Lett. **B237**, 592 (1990).

[45] Forte, S. and Garrido, L. and Latorre, J. I. and Piccione, A., JHEP **05**, 062 (2002).

[46] Del Debbio, L. and Forte, S. and Latorre, J. I. and Piccione, A. and Rojo, J., JHEP **03**, 080 (2005).

[47] Corcella, G. and Magnea, L., *Soft-gluon resummation effects on parton distributions*. Preprint hep-ph/0506278, 2005.

[48] Pumplin, J. et al., JHEP **07**, 012 (2002).

[49] Fadin, V. S. and Lipatov, L. N., Phys. Lett. **B429**, 127 (1998).

[50] Ciafaloni, M. and Colferai, D., *Dimensional regularisation and factorisation schemes in the bfkl equation at subleading level*. Preprint hep-ph/0507106, 2005.

[51] Ball, R. D. and Forte, S., Phys. Lett. **B465**, 271 (1999).

[52] Catani, S. and Ciafaloni, M. and Hautmann, F., Phys. Lett. **B242**, 97 (1990).

[53] Catani, S. and Ciafaloni, M. and Hautmann, F., Nucl. Phys. **B366**, 135 (1991).

[54] Collins, J. C. and Ellis, R. K., Nucl. Phys. **B360**, 3 (1991).

[55] Bartels, J. and Gieseke, S. and Qiao, C. F., Phys. Rev. **D63**, 056014 (2001).

[56] Bartels, J. and Gieseke, S. and Kyrieleis, A., Phys. Rev. **D65**, 014006 (2002).

[57] Bartels, J. and Colferai, D. and Gieseke, S. and Kyrieleis, A., Phys. Rev. **D66**, 094017 (2002).

[58] Bartels, J. and Colferai, D. and Vacca, G. P., Eur. Phys. J. **C24**, 83 (2002).

[59] Bartels, J. and Colferai, D. and Vacca, G. P., Eur. Phys. J. **C29**, 235 (2003).

[60] Salam, G. P., JHEP **07**, 019 (1998).

[61] Ciafaloni, M. and Colferai, D., Phys. Lett. **B452**, 372 (1999).

[62] Ciafaloni, M. and Colferai, D. and Salam, G. P., JHEP **07**, 054 (2000).

[63] Ball, R. D. and Forte, S., Phys. Lett. **B405**, 317 (1997).

[64] Catani, S. and Hautmann, F., Nucl. Phys. **B427**, 475 (1994).

[65] Ball, R. D. and Forte, S., Phys. Lett. **B359**, 362 (1995).

[66] Ciafaloni, M. and Colferai, D. and Salam, G. P. and Stasto, A. M., Phys. Lett. **B587**, 87 (2004).

[67] Andersen, J. R. and Sabio Vera, A., Phys. Lett. **B567**, 116 (2003).

[68] Andersen, J. R. and Sabio Vera, A., Nucl. Phys. **B679**, 345 (2004).

[69] Sabio Vera, A., Nucl. Phys. **B722**, 65 (2005).

[70] Schmidt, C. R., Phys. Rev. **D60**, 074003 (1999).

[71] Forshaw, J. R. and Ross, D. A. and Sabio Vera, A., Phys. Lett. **B455**, 273 (1999).

[72] Chachamis, G. and Lublinsky, M. and Sabio Vera, A., Nucl. Phys. **A748**, 649 (2005).

[73] Andersen, J. R. and Sabio Vera, A., JHEP **01**, 045 (2005).

2 Working group 2: Multi-Jet final states and energy flows

List of participants in the working group:

Jeppe R. Andersen, A. Banfi, Joachim Bartels, Riccardo Bellan, Marteen Boonekamp, Konstantin Boreskov, Nick Brook, Alessia Bruni, Craig Buttar, Jonathan Butterworth, Tancredi Carli, Sascha Caron, Gianluca Cerminara, Sergei Chekanov, John Collins, Gennaro Corcella, Brian Cox, Zofia Czyczula, Albert De Roeck, Mrinal Dasgupta, Giovanna Davatz, Y. Delenda, Markus Diehl, Rick Field, Jeff Forshaw, Piotr Golonka, E. Gotsman, Guenter Grindhammer, C. Group, Gösta Gustafson, Stefan Höche, Stanislaw Jadach, Hannes Jung, Aliosha Kaidalov, Valery Khoze, Michael Klasen, Henri Kowalski, Frank Krauss, Nils Lavesson, Victor Lendermann, Eugene Levin, Leif Lönnblad, Misha Lublinsky, Stephen Magill, Michelangelo Mangano, Uri Maor, Alan Martin, Anna Mastroberardino, Stephen Maxfield, Bruce Mellado, Arthur Moraes, E. Naftali, A. Nikitenko, Risto Orava, A. Prygarin, Peter Richardson, Elzbieta Richter-Was, Eduardo Rodrigues, Christophe Royon, Misha Ryskin, Agustin Sabio Vera, Gavin Salam, Alexander Savin, Andreas Shälicke, Frank-Peter Schilling, Thomas Schoerner-Sadenius, Steffen Schumann, Michael Seymour, Torbjörn Sjöstrand, M. Skrzypek, W.J. Stirling, Paul Szczypka, Marek Tasevsky, Thomas Teubner, Niels Tuning, Jacek Turnau, Raju Venugopalan, Bennie Ward, Zbigniew Was, Ben Waugh, S.A. Yost, G. Zanderighi

Introduction to multi-jet final states and energy flows

Craig Buttar[1]*, Jon Butterworth*[2]*, Valery Khoze*[3]*, Leif Lönnblad*[4] *and Niels Tuning*[5]
[1]Department of Physics and Astronomy, University of Glasgow, UK; [2]Department of Physics and Astronomy, University College London, UK; [3]IPPP, University of Durham, UK; [4] Department of Theoretical Physics, Lund University, Sweden; [5]NIKHEF, Amsterdam, The Netherlands.

Abstract
We summarize the activities of working group 2 of the HERA/LHC Workshop dealing with multi-jet final states and energy flows. Among the more specific topics considered were Underlying event and minimum bias, Rapidity gaps and survival probabilities, Multi-jet topologies and multi-scale QCD and Parton shower/Matrix element matching.

1 Introduction

In many ways, LHC will become the best QCD machine ever built. It will allow us to study the production of hadrons and jets at unprecedented collision energies and will surely increase our understanding of QCD tremendously. Of course, some may argue that QCD already is a well understood and an integral part of the Standard Model, and the reason for building the LHC is to discover new phenomena, hopefully beyond the Standard Model.

However, the fact is that QCD is still not a completely understood theory. The qualitative aspects of asymptotic freedom and confinement may be under control, but the quantitative predictive power of the theory is still not at a satisfactory level. This is particularly true for the non-perturbative region, but also for the high-energy limit, where the hard scale of a process is much smaller than the total collision energy. The latter situation will be dominant in the bulk of events produced at the LHC. The triggers at the main LHC detectors will discard the majority of such events, but what is left will be processes with hard scales of around 100 GeV, which is still more than a hundred times smaller than the collision energy. And there will be significant amounts of minimum bias data taken as well.

Except for a handful of gold-plated signals for new physics, any such search will be plagued by huge backgrounds stemming from pure QCD or other Standard Model processes involving jets. Hence, even if the study of QCD may seem to be a mundane preoccupation, it is of the utmost importance if we are to find and understand the few needles of new physics hopefully present in the immense LHC haystack.

Although the Tevatron may seem to be the obvious place to learn about QCD processes relevant for the LHC, the triggers there are typically tuned to high-scale processes, not far from the total collision energy. This means that HERA can give important additional insight, since there the situation is in some senses closer to the LHC, with the ratio of the typical hard scale and the total energy in DIS being $\sqrt{\langle Q^2 \rangle / S} \sim 0.01$. In addition, HERA allows us to study such processes in a more controlled environment, where one side of the collision is well constrained by our relatively precise understanding of electro-weak physics.

In our working group we have studied in some detail which lessons about multi-jet final states and general hadronic energy flows can be learned from HERA when preparing for the analysis of LHC data. And in this brief summary we will in a few pages try to distill the progress made by almost a hundred physicists as reported in more than fifty talks in this workshop and also in almost twenty separate contributions to these proceedings. The work was broadly divided into four categories: Underlying events and minimum bias; Rapidity gaps and survival probabilities; Multi-jet topologies and multi-scale QCD; and Matrix element – parton shower matching.

The first category may not represent the most striking feature of HERA physics, but it will surely be of great importance for the LHC. And it turns out that there are many possibilities to gain further understanding of underlying events both in photo-production and DIS at HERA.

The study of rapidity gaps and, in particular, hard diffractive scattering gained momenta when it was observed at HERA, and the suggestion to use such processes to obtain clean signals of new physics at the LHC presents exciting prospects where the experience from HERA will be very important.

Multi-scale processes have already been presented as an important connection between HERA and the LHC. This is not least true for the LHCb experiment, where the understanding of the forward region is vital, a region which has been intensely studied at HERA. Also the recent theoretical development in QCD resummation techniques, which so far have mainly been applied to e^+e^--annihilation, may provide important tools for understanding event shapes at the LHC, and the corresponding application to HERA data will be essential for this understanding.

Finally, the more technical issue of matching fixed-order tree-level matrix elements with parton shower generators as well as other theoretical improvements of such simulation programs will surely be vital for the successful understanding of data from the LHC and also here the comparison to HERA data will be essential for the tuning and validation.

It should be noted that all of these categories, presented in more detail below, have a fairly large overlap with other working groups in this workshop. There most obvious overlaps are the working groups for Diffraction and Monte Carlo simulations, but there is also overlap with the heavy flavour and parton distributions working groups.

2 Underlying events and minimum bias

An understanding of the underlying event is an interesting physics topic in its own right but is also crucial in developing robust analyses for LHC physics. The underlying event can enhance central jet production, reducing the effectiveness of the central jet veto in analyses such as the vector boson fusion Higgs channel, or reduce the isolation of leptons resulting in reduced efficiency for identifying isolated leptons. In particular for LHCb and ALICE, where the triggers typically do not mandate high-scale processes, a good understanding of underlying events and minimum-bias events is crucial.

In this workshop there were several contributions dealing with underlying events and multiple interactions. They are all described in a joint contribution to these proceedings [1]. There the event generator models in PYTHIA [2, 3, 4, 5], HERWIG/JIMMY [6, 7, 8] and SHERPA [9] are presented together with results from tuning these and other models to available data. The contribution also includes a summary of the plenary talk by Gösta Gustafson on the theory and phenomenology underlying events and multiple scattering.

Of the models presented and studied in [1], the one implemented in PYTHIA is probably the most advanced. This model has recently been developed further, introducing a scheme for *interleaving* the multiple interaction with a transverse-momentum ordered parton shower [3]. In contrast the default underlying event model in HERWIG is a simple parameterization of UA5 data [10]. However, HERWIG is easily interfaced to the multiple-interaction model in the JIMMY program, which is similar to the PYTHIA model in spirit, although many of the details differ. Also the JIMMY program has recently been improved, making the generation of events more efficient where the signal process is different from the additional multiple scattering processes. Also the SHERPA event generator is now equipped with multiple interactions. Again, this model is similar in spirit to that in PYTHIA. One interesting aspect which differs is the attempt to incorporate the multiple scatterings in the general CKKW (see section 5 below) framework of SHERPA.

The CDF collaboration has carried out studies of the underlying event in jet processes [11, 12, 13] and this was used to provide a tuning for PYTHIA. In [1] a new analysis is presented which has extended these studies by increasing the energy range of the leading jet from around 50GeV to 450GeV using

E_T from the calorimeter as well as particle $p_\perp$ measured in the tracker, and defining two-jet topologies as a subset of the leading jet to investigate the beam-beam and radiation components of the underlying event. Both PYTHIA tune-A and HERWIG/JIMMY were found to be in good agreement with the data, although both underestimate the transverse energy. The extension to higher energy scale shows that the underlying event activity increases with leading jet $p_\perp$ ie. the hardness of the primary scatter, but by studying the maximum and minimum activity it is seen that this rise is largely due to bremsstrahlung from the primary scattering rather than secondary interactions between the beam remnants.

The CDF analysis has been carried out primarily at 1.8 TeV although also some of the early 546 GeV data has been analyzed. This has meant that there is only limited information on the energy dependence of the underlying event. To cover a wider range of energy ATLAS have used minimum bias data from the SppS and Tevatron covering 200 GeV to 1.8 TeV in addition to the CDF underlying event data to tune PYTHIA and HERWIG/JIMMY. Comparing the predictions of minimum bias and underlying event distributions at the LHC using the tuned PYTHIA, the tuned HERWIG/JIMMY and PHOJET [14] shows large variations, emphasizing the need to understand the energy dependence of these processes better. The energy dependence has been investigated further by LHCb, again using minimum bias data to fit the parameters required for the model of energy dependence in PYTHIA.

Both the ATLAS and LHCb analyses have the implicit assumption that minimum bias and the underlying event have the same physics origin. While CDF data supports this, it would be helpful to probe the underlying event directly over a larger range of energy scales. HERA is in a prime position to make such a contribution by studying jets from photoproduction in an energy range corresponding to centre-of-mass energies in the region of 200 GeV, fitting well with the low energy minimum bias data. In photoproduction, resolved photons behave like hadrons so that HERA is effectively a hadron-hadron collider. Photoproduction data shows that particle flow and multi-jet measurements require models with multiple interactions to best describe the data but detailed studies of multiple interactions have not been made. However, studies of particle and energy flow in the transverse region similar to that carried out by CDF could be made at HERA.

An interesting question is whether there is also an underlying event present in DIS at HERA. As explained in [15, 16] it is possible to relate diffraction and saturation to multiple interaction processes also for DIS using a QCD reformulation of the so-called AGK cutting rules [17]. And since diffractive processes have been clearly seen at high Q^2 at HERA, it is reasonable to expect that multiple interactions may also be present. A good place to search for such effects is in forward jet production at HERA. In [18] preliminary results are presented indicating that multiple-interaction effects may indeed give a noticeable increase the measured forward jet cross section in resolved virtual photon processes at small x and moderate Q^2.

The connection between multiple interactions, saturation and diffraction was also discussed in the plenary talk by Gösta Gustafson. He pointed out a possible problem with the qualitative AGK predictions for the hadronic multiplicity in multiple-interaction events. Taking the tuning of PYTHIA to CDF data at face value, there is an indication that the colour flows of secondary interactions are not independent from the primary scattering. Rather, the different colour flows seem to combine in a way where the total string length is minimized, resulting in a multiplicity which does not grow proportionally to the number of scatterings. Currently there is no theoretical understanding of this phenomenon. Gustafson also pointed out the problem that all multiple-interaction models discussed here rely on collinear factorization of the individual scatterings in a region where we expect $k_\perp$-factorization to be the relevant formalism. In fact, using $k_\perp$-factorization, the soft divergencies in the partonic cross section present in the conventional models may be removed, which could make the extrapolation of the model predictions to high energy more constrained.

3 Rapidity gaps and survival probabilities

A characteristic signature of diffractive processes is the existence of a large rapidity gap (LRG) in the final state, defined as a region of (pseudo-) rapidity devoid of hadronic activity. A rapidity gap may be adjacent to a leading proton or may arise between the decay products of final hadronic systems. The appearance of the rapidity gaps is intimately related to the exchange in the t-channel of objects with vacuum quantum numbers (Pomeron in the Regge theory, di-gluon Pomeron in pQCD, photon or W -mediator). The diffractive rapidity gap events have been studied in many details at the ISR, SPS, HERA and the Tevatron. The LHC is the first collider which will have enough energy to allow the events with several (n=2-4) LRGs.

The activity of our Working Group was focused mainly on the LRGs in the hard diffractive processes. For specifics of the photon and W- mediated reactions see, for example Refs. [19, 20, 21, 22].

An intensive discussion concerned the breakdown of factorization in hard hadronic diffractive processes. It is the consequence of unitarisation effects, that both hard and Regge factorization are broken. This breakdown of factorization is experimentally seen [23] as the suppression of the single diffractive dijet cross section at the Tevatron as compared to the prediction based on HERA results. The observed suppression is in a quantitative agreement with the calculations [24] where the unitarisation effects are described by multi-Pomeron exchange diagrams. The analysis of the current CDF diffractive dijet data with one or two rapidity gaps shows a good agreement with this approach. The situation with the factorization breaking in dijet photoproduction is not completely clear and further experimental and theoretical efforts are needed. A possible way to study this effect is to measure the ratio of diffractive and inclusive dijet photoproduction, see [25].

It is important to emphasize that the rapidity gap signal is very powerful but, at the same time, quite a fragile tool. We have to pay a price for ensuring such a clean environment. The gaps may easily fade away (filled by hadronic secondaries) due to various sources of QCD "radiation damage":

(i) soft or hard rescattering between the interacting hadrons (classic screening/unitarisation effects or underlying event);

(ii) bremsstrahlung induced by the 'active' partons in the hard subprocesses;

(iii) radiation originating from the small transverse distances in two-gluon Pomeron dipoles.

An essential issue in the calculation of the rate of events with LRG concerns the size of the factor W which determines the probability for the gaps to survive in the (hostile) QCD environment. As discussed in the contributions of Brian Cox [26] and Jeff Forshaw [27], this factor is a crucial ingredient for evaluation of the discovery potential of the LHC in the exclusive processes with double proton tagging.

Symbolically, the survival probability W can be written as

$$W \;=\; S^2 T^2. \tag{1}$$

S^2 is the probability that the gaps are not filled by secondary particles generated by soft rescattering, i.e. that no other interactions occur except the hard production process. Following Bjorken [28, 29], who first introduced such a factor in the context of rescattering, such a factor is often called the survival probability of LRG. The second factor, T^2, is the price to pay for not having gluon radiation in the hard production subprocess. It is related to Sudakov-suppression phenomena and is incorporated in the pQCD calculation via the skewed unintegrated parton densities. In more detail the physics of of Sudakov suppression is discussed in the contribution of Jeff Forshaw to these Proceedings [27].

In some sense the soft survival factor S^2 is the "Achilles heel" of the calculations of the rates of diffractive processes, since, in principle, S^2 could strongly depend on the phenomenological models for soft diffraction. This factor is not universal, but depends on the particular hard subprocess, as well as on the distribution of partons inside the proton in impact parameter space. It has a specific dependence on the characteristic momentum fractions carried by the active partons in the colliding hadrons [24].

However, the good news is that, as discussed in these Proceedings by Uri Maor et al. [30], the existing estimates of S^2 calculated by different groups for the same processes appear to be in a reasonably good agreement with each other. This is related to the fact that these approaches reproduce the existing data on high-energy soft interactions, and, thus, result in the similar profile of the optical density in the impact parameter space. Another reason results from the comparatively small role of the high-mass diffractive dissociation.

Note, that it is possible to check the value of S^2 by observing double-diffractive dijet production [31]. The gap survival in the Higgs production via the WW-fusion process can be probed in Z production which is driven by the same dynamics, and has a higher cross section, see Ref. [32, 33]. Let us emphasize that it is the presence of this factor which makes the calculation infrared stable, and pQCD applicable. Neglecting the Sudakov suppression would lead to a considerable overshooting of the cross section of the hard central exclusive processes at large momentum transfer.

4 Multi-jet topologies and multi-scale QCD

In this workshop work a wide range of topics regarding jet production and multi-scale processes has been presented [34]. It is of great interest what the LHC will teach us in the area of QCD, but at the same time uncertainties on the theoretical predictions for processes at the LHC should be limited as far as possible beforehand. By using the knowledge attained at HERA, our models can be sharpened and our theories can be tested.

Predictions of the event topology of $gg \to H$ at the LHC have been investigated for various parton shower models - such as PYTHIA, HERWIG and ARIADNE, that have proven their validity at HERA - and uncertainties in the event selection have been estimated [35, 36]. In the parton cascade as implemented in some of these programs, the parton emissions are calculated using the DGLAP approach, with the partons ordered in virtuality. DGLAP accurately describes high-energy collisions of particles at moderate values of the Bjorken-x by resummation of the leading log terms of transverse momenta ($\alpha_s \ln Q^2$). However, to fixed order, the QCD scale used in the ladder is not uniquely defined. There are many examples were more than one hard scale plays a role in the hard scatter, such as the virtuality Q, the transverse momentum E_T of the jet, or the mass of a produced object. Also, at low values of Bjorken-x large logarithms appear ($\alpha_s \ln 1/x$), leading to large corrections.

The CCFM formalism takes this into account, describing the evolution in an angular ordered region of phase space, while reproducing DGLAP and BFKL in the appropriate asymptotic limits. The CASCADE program has implemented the CCFM formalism, describing the low-x F_2 data and forward jet data at HERA. The predictions for the jet production at the LHC have been studied, both in the context of a $gg \to H$, as well as in the context of the forward event topology at LHCb [37].

In order to get reliable predictions for exclusive final-state processes, unintegrated parton density function $f(x, Q^2, k_\perp)$ (uPDF) become indispensable. For example, in the small x regime, when the transverse momenta of the partons are of the same order as their longitudinal momenta, the collinear approximation is no longer appropriate and $k_\perp$-factorization has to be applied, with the appropriate CCFM evolution equations. In this workshop various parameterizations for unintegrated gluon densities matched to HERA F_2 data were compared to each-other [38]. It is however still questionable if these densities are constrained enough for reliable predictions for Higgs production cross section. Final state measurements like photoproduction of D^*+jet events could however constrain these uPDFs further. It is argued that it is important to reformulate perturbative QCD in terms of fully unintegrated parton densities, since neglecting parton transverse momentum leads to wrong results. The HERA F_2 data has also been fitted using non-linear BFKL evolution, expressed with a universal dipole cross section, which in turn can be related to the unintegrated gluon distribution. Finally, a theoretical description of hard diffractive processes at HERA can provide information on the so-called generalized, or skewed, gluon distribution (depending on the x of the emitted and absorbed gluon), providing for a theoretical description for diffractive Higgs production at the LHC.

The role of HERA is also emphasized in the area of resummed calculations, obtaining accurate QCD parameters such as the strong coupling, quark masses and parton distribution functions, which are vital inputs for predictions at the LHC. For example, event shape distributions at HERA led to the finding of non-global logarithms, influencing observables at the LHC such as energy flows away from jets. Additionally, HERA data seem to confirm $1/Q$ power corrections (arising from gluon emission with transverse momentum $\sim \Lambda_{QCD}$), demonstrating that these corrections are not affected by the presence of the initial state proton. HERA data is also used to study dijet E_T and angular spectra, in order to test NLL perturbative predictions. Finally, it has been discussed whether additional small x terms are needed to accommodate HERA DIS data, which at LHC energies would result in a broadening of the vector boson p_T spectrum.

5 Parton shower/ME matching

LHC is, of course, mainly a machine for discovering new physics. But irrespective of what new phenomena may exist, we know for sure that the LHC events will contain huge numbers of hadrons, and that a large fraction of these events will have many hard jets produced by standard QCD processes. Such events are interesting in their own right, but they are also important backgrounds for almost any signal of new physics. Unfortunately the standard Parton Shower (PS) based event generators of today are not well suited to describe events with more than a couple of hard jets. The alternative is to use matrix element (ME) generator programs, which typically can generate up to six hard partons according to the exact fixed-order tree-level matrix elements. But these generators are not well suited for describing the conversion of these hard partons into jets of hadrons.

To get properly generated events it is therefore important to interface the ME generators to realistic hadronization models, which requires that also soft and collinear partons are generated according PS models to get reliable predictions for the intra- and inter-jet structure. When adding a PS to an event from a ME generator, it is important to avoid double counting. Hence the PS must be *vetoed* to avoid generating parton emissions above the cutoff needed to avoid divergences in the ME generator. In addition the PS assumes that the emissions are ordered in some evolution variable (scale) and uses Sudakov form factors to ensure that there was no additional emission with a scale between two generated emissions. This also generates the virtual corrections to the splittings. The ME generators, of course, have no such ordering since all diagrams are added coherently. However, there is still a need for a cutoff in some scale to regulate soft and collinear divergencies, and to naively add a PS to events from a ME generator will therefore give a strong dependence on this cutoff.

A solution to this problem was presented by Catani et al. [39]. This so-called CKKW procedure is based on using a jet reconstruction algorithm on the ME generated event to define an ordering of the emissions and then reweight the event according to Sudakov form factors obtained from the reconstructed scales. In this way it was shown that the dependence on the ME cutoff cancels to NLL accuracy. The procedure was originally developed for e^+e^- annihilation where it was further developed in [40], but lately it has also been applied to hadron-hadron collisions in [41, 42, 43, 44, 45] using several different parton shower models. In addition an alternative procedure, called MLM, was developed by Mangano [46, 47] which is similar in spirit to CKKW, but which has a simpler interface between the ME and PS program.

There was some hope that during this workshop there would also be developed an implementation of CKKW for DIS. This would be interesting, not least because the procedure would then be tested in a small-x environment, and comparing with such HERA data as well as with high scale Tevatron data should then give a more reliable understanding about the uncertainties when extrapolating to LHC. Although some progress has been made on the application to DIS [48] there was not enough time to make a proper implementation. Instead the activities were focused on comparing the predictions of some of the programs (SHERPA [9] and MADGRAPH/MADEVENT [49]+ARIADNE [50] using CKKW, and ALPGEN [51]+PYTHIA [4] using MLM) for the case of W+jets production at the Tevatron and the LHC. This process is very interesting in its own right, but is also an important background for almost any signal of new physics at the LHC. The results are presented in these proceedings [52] and it was found that the models give fairly similar predictions for jet rates, but some differences were found eg. for the rapidity correlation between jets and the W. The latter may be related to the fact that W production, especially at the LHC, can be considered to be a small-x process ($m_W/\sqrt{S} \sim x \sim 0.005$) and we know that there are large differences between parton shower models in this region. This emphasizes again the importance of confronting the ME+PS matching procedures also with HERA DIS data.

Possible improvements of the QCD PS approach were discussed in three other contributions to these proceedings. All of these are based on experience of Monte Carlo programs for QED resummation. One of these contributions [53] describes a new algorithm for forward evolution of the initial-state parton cascade in which the type and energy of the final parton is predefined/constrained. Contrary to the widely used backward-evolution algorithms [54] this algorithm is similar to the one used in the LDCMC generator [55] and does not need a fully evolved PDF parameterization as input.

Another contribution [56] describes, using an operator formalism, what we can learn about QCD parton showers from the popular PHOTOS generator, which combines in a clever way soft photon resummation and hard collinear photon resummation in QED. Finally there is a contribution [57] which describes a more ambitious attempt to combine ME+PS calculations for both QCD and QED, preserving the proper soft gluon limit and the standard factorization of collinear singularities. All of these contributions represents work which is still in a rather early stage. Nevertheless, they signal important efforts which may lead to interesting new Monte Carlo tools for the LHC era.

6 Conclusions and outlook

In this summary we hope to have made it clear that there is a rich flora of interesting topics relating to jets and hadronic energy flows where the understanding of results from HERA will be important for the upcoming analysis of LHC data. It should also be clear that although substantial progress has been made during this workshop, we have only started to botanize among these topics. Hence, as we now thank the participants of our working group for all the work they have contributed to the workshop, we would also like to remind them, and also other readers of these proceedings, that there is very much work still to be done.

References

[1] C. Buttar *et al.*, *Underlying events*. These Proceedings.

[2] T. Sjöstrand and M. van Zijl, Phys. Rev. **D36**, 2019 (1987).

[3] T. Sjöstrand and P. Z. Skands, Eur. Phys. J. **C39**, 129 (2005). `hep-ph/0408302`.

[4] T. Sjštrand, and others, Comput. Phys. Commun. **135**, 238 (2001). `arXiv:hep-ph/0010017`.

[5] T. Sjöstrand, L. Lönnblad, S. Mrenna, and P. Skands (2003). `hep-ph/0308153`.

[6] J. Butterworth *et al.* `http://hepforge.cedar.ac.uk/jimmy/`.

[7] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, Z. Phys. **C72**, 637 (1996). `hep-ph/9601371`.

[8] G. Corcella *et al.*, JHEP **01**, 010 (2001). `hep-ph/0011363`.

[9] T. Gleisberg *et al.*, JHEP **02**, 056 (2004). `hep-ph/0311263`.

[10] UA5 Collaboration, G. J. Alner *et al.*, Nucl. Phys. **B291**, 445 (1987).

[11] CDF Collaboration, J. Huston, Int. J. Mod. Phys. **A16S1A**, 219 (2001).

[12] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. **D65**, 092002 (2002).

[13] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. **D70**, 072002 (2004). `hep-ex/0404004`.

[14] R. Engel, *Phojet program and manual*. `http://www-ik.fzk.de/~engel/phojet.html`.

[15] H. Kowalski, *Multiple Interactions in DIS*. These Proceedings.

[16] J. Bartels, *Multiple Scattering at HERA and at LHC - Remarks on the AGK Rules*. These Proceedings.

[17] V. A. Abramovsky, V. N. Gribov, and O. V. Kancheli, Yad. Fiz. **18**, 595 (1973).

[18] J. Turnau and L. Lönnblad, *Forward Jets and Multiple Interactions*. These Proceedings.

[19] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C23**, 311 (2002). `hep-ph/0111078`.

[20] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C21**, 99 (2001). `hep-ph/0104230`.

[21] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C18**, 167 (2000). `hep-ph/0007359`.

[22] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C24**, 459 (2002). `hep-ph/0201301`.

[23] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **84**, 5043 (2000).

[24] A. B. Kaidalov, V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C21**, 521 (2001). `hep-ph/0105145`.

[25] A. B. Kaidalov, V. A. Khoze, A. D. Martin, and M. G. Ryskin, Phys. Lett. **B567**, 61 (2003). `hep-ph/0306134`.

[26] B. Cox *et al.*, *Diffractive Higgs production: experiment*. These Proceedings.

[27] J. Forshaw, *Diffractive Higgs production: theory*. These Proceedings.

[28] J. D. Bjorken, Int. J. Mod. Phys. **A7**, 4189 (1992).

[29] J. D. Bjorken, Phys. Rev. **D47**, 101 (1993).

[30] U. Maor *et al.*, *Gap-survival and factorization breaking*. These Proceedings.

[31] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C14**, 525 (2000). `hep-ph/0002072`.

[32] H. Chehime and D. Zeppenfeld, Phys. Rev. **D47**, 3898 (1993).

[33] V. A. Khoze, M. G. Ryskin, W. J. Stirling, and P. H. Williams, Eur. Phys. J. **C26**, 429 (2003). `hep-ph/0207365`.

[34] Z. Czyczula *et al.*, *Multi-Jet Production and Multi-Scale QCD*. These Proceedings.

[35] Z. Czyczula and E. Richter-Was, *MSSM Higgs production with the Yukawa bbH coupling induced mechanisms*. In [34], These Proceedings.

[36] G. Davatz and A. Nikitenko, *$gg \to H$ at the LHC: Uncertainty due to a Jet Veto*. In [34], These Proceedings.

[37] E. Rodrigues and N. Tuning, *Forward Studies with CASCADE at LHC Energies*. In [34], These Proceedings.

[38] J. Collins *et al.*, *Unintegrated parton density functions*. These proceedings.

[39] S. Catani *et al.*, JHEP **11**, 063 (2001). `hep-ph/0109231`.

[40] L. Lönnblad, JHEP **05**, 046 (2002). `hep-ph/0112284`.

[41] F. Krauss, JHEP **08**, 015 (2002). `hep-ph/0205283`.

[42] S. Mrenna and P. Richardson, JHEP **05**, 040 (2004). `hep-ph/0312274`.

[43] F. Krauss, A. Schalicke, S. Schumann, and G. Soff, Phys. Rev. **D70**, 114009 (2004). `hep-ph/0409106`.

[44] N. Lavesson and L. Lönnblad, JHEP **07**, 054 (2005). `hep-ph/0503293`.

[45] F. Krauss, A. Schaelicke, S. Schumann, and G. Soff (2005). `hep-ph/0503280`.

[46] M. Mangano, *The so-called mlm prescription for me/ps matching*. `http://www-cpd.fnal.gov/personal/mrenna/tuning/nov2002/mlm.pdf`. Talk presented at the Fermilab ME/MC Tuning Workshop, October 4, 2002.

[47] M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. **B632**, 343 (2002). `hep-ph/0108069`.

[48] C. Åberg, *Correcting the colour dipole cascade with fixed order matrix elements in deep inelastic scattering*. Diploma thesis, LU-TP 04-25.

[49] F. Maltoni and T. Stelzer, JHEP **02**, 027 (2003). `hep-ph/0208156`.

[50] L. Lïnblad, Comput. Phys. Commun. **71**, 15 (1992).

[51] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07**, 001 (2003). `hep-ph/0206293`.

[52] S. Hoeche *et al.*, *Matching Parton Showers and Matrix Elements*. These Proceedings.

[53] S. Jadach and M. Skrzypek, *Constrained non-Markovian Monte Carlo modeling of the evolution equation in QCD*. These Proceedings.

[54] M. Bengtsson and T. Sjöstrand, Z. Phys. **C37**, 465 (1988).

[55] H. Kharraziha and L. Lönnblad, JHEP **03**, 006 (1998). `hep-ph/9709424`.

[56] P. Golonka and Z. Was, *PHOTOS as a pocket parton shower: flexibility tests for the algorithm, howpublished = These Proceedings*.

[57] B. Ward and S. Yost, *QED$\otimes$QCD Exponentiation and Shower/ME Matching at the LHC*. These Proceedings.

The Underlying Event

C.M.Buttar[1], *J.M.Butterworth*[2], *R.D.Field*[3], *C.Group*[3], *G.Gustafson*[4], *S.Hoeche*[5], *F.Krauss*[5], *A.Moraes*[1], *M.H.Seymour*[6,7], *A.Schalicke*[8], *P.Szczypka*[9], *T.Sjöstrand*[4,7]
[1]Dept. of Physics and Astronomy, University of Glasgow, UK
[2]Dept. of Physics and Astronomy, University College London, UK
[3]Dept. of Physics, University of Florida, USA
[4]Dept. of Theoretical Physics, Lund University, Sweden
[5]Institute for Theoretical Physics, Dresden University of Technology, FRG
[6]School of Physics and Astronomy, University of Manchester, UK
[7]CERN, Switzerland
[8]DESY Zeuthen, FRG
[9]Dept. of Physics, University of Bristol, UK

Abstract

The contributions to working group II: "Multi-jet final states and energy flows" on the underlying event are summarized. The study of the underlying event in hadronic collisions is presented and Monte Carlo tunings based on this are described. New theoretical and Monte Carlo methods for describing the underlying event are also discussed.

1 Introduction

The underlying event is an important element of the hadronic environment within which all physics at the LHC, from Higgs searches to physics beyond the standard model, will take place. Many aspects of the underlying event will be constrained by LHC data when they arrive. However, the physics is so complex, spanning non-perturbative and perturbative QCD and including sensitivities to multi-scale and very low-x physics, that even after LHC switch-on many uncertainties will remain. For this reason, and also for planning purposes, it is critical to have to hand sensible models containing our best physical knowledge and intuition, tuned to all relevant available data.

In this summary of several contributions to the workshop, we first outline the available models in section 2, most of which are in use at HERA and/or the Tevatron. Recent improvements, some of which were made during the workshop, are also discussed.

Next, current work on tuning these to data is discussed. The underlying event has been extensively studied by CDF and the latest results are presented in section 3 and compared to predictions from the PYTHIA and HERWIG+JIMMY Monte Carlo generators. The CDF tunings are compared to other tunings based on CDF data and minimum bias data and used to predict the level of underlying events at the LHC in sections 4 and 5. These reports are very much a snapshot of ongoing work, which will be continued in the follow-up meetings of this workshop and the TeV4LHC workshop.

One major issue in extrapolating the underlying event (UE) to LHC energies is the possible energy dependence of the transverse momentum cut-off between hard and soft scatters, $\hat{p}_T^{\min}$. The need for such a cut-off may be avoided by using the $k_\perp$ factorization scheme as discussed in section 6, where soft emissions do not contribute to the total cross-section or to the parton density functions (PDFs), but do contribute to the properties of the event. The cross-section for a chain of partonic emission can be extracted from HERA data and can be used to predict the minijet rate or multiple interaction rate in pp or $\mathrm{p\bar{p}}$ collisions. The running of α_s still introduces a cut-off scale between soft and hard chains; however it has been shown that the total cross-section is insensitive to this cut-off and predictions for the mini-jet rate at the LHC are stable. The hadron multiplicity observed in the CDF underlying event data indicates that the string connections in the underlying event are made to minimise the string length. This is the opposite to what is observed in e^+e^- collisions. The implications for this on the AGK cutting rules is discussed further in section 6.

This summary ends with a section on conclusions and suggestions for future work.

2 Underlying event models

Several underlying event models are available, at varying stages of development and use. In this section we review the status of thosed discussed during the workshop.

2.1 Multiple Interactions in PYTHIA

The basic implementation of multiple interactions in PYTHIA is almost 20 years old, and many of the key aspects have been confirmed by comparisons with data. In recent years the model has been gradually improved, with junction-string topologies, with flavour-correlated multiparton densities, and with transverse-momentum-ordered showers interleaved with the multiple interactions. However, the "correct" description of colour flow still remains to be found.

The traditional PYTHIA [1, 2] model for multiple interactions (MI) [3] is based on a few principles:

1. The naive perturbative QCD $2 \to 2$ cross section is divergent like $\mathrm{d}p_\perp^2/p_\perp^4$ for transverse momenta $p_\perp \to 0$. Colour screening, from the fact that the incoming coloured partons are confined in colour singlet states, should introduce a dampening of this divergence, e.g. by a factor $p_\perp^4/(p_{\perp 0}^2 + p_\perp^2)^2$, where $p_{\perp 0}$ is a free parameter, which comes out to be of the order of 2 GeV.
2. From the thus regularized integrated interaction rate $\sigma_{\mathrm{int}}(E_{\mathrm{cm}}, p_{\perp 0})$ and the nondiffractive cross section $\sigma_{\mathrm{nd}}(E_{\mathrm{cm}})$, the average number of interactions per event can be derived as $\langle n_{\mathrm{int}} \rangle = \sigma_{\mathrm{int}}/\sigma_{\mathrm{nd}}$. With no impact-parameter dependence, the actual number of interactions is given by a Poissonian with mean as above (modulo some corrections coming from $n_{\mathrm{int}} = 0$).
3. More realistically, since hadrons are extended objects, there should be more (average) activity in central collisions than in peripheral ones. By introducing a matter distribution inside a hadron, the overlap between the two incoming hadrons can be calculated as a function of impact parameter b. The number of interactions is now a Poissonian for each b separately, with a mean proportional to the overlap. All events are required to contain at least one interaction; thereby the cross section is automatically dampened for large b. Empirically, the required hadronic impact parameter profile is more peaked at small b than in a Gaussian distribution.
4. It is natural to consider the interactions in an event in order of decreasing $p_\perp$ values. Such a $p_\perp$ ordering has a natural interpretation in terms of formation-time arguments. The generation procedure can conveniently be written in a language similar to that used for parton showers, with the equivalent of a Sudakov form factor being used to pick the next smaller $p_\perp$, given the previous ones. It allows the hardest interaction to be described in terms of conventional PDFs, whereas subsequent ones have to be based on modified PDFs, at the very least reduced by energy–momentum conservation effects. This also reduces the tail of events with very many interactions.
5. Technical limitations lead to several simplifications, such that only the hardest interaction was allowed to develop initial- and final state interactions, and have flavours selected completely freely.
6. Colour correlations between different scatterings cannot be predicted by perturbation theory, but have a direct consequence on the structure of events. One of the most senstive quantities is $\langle p_\perp \rangle (n_{\mathrm{charged}})$. Data here suggest a very strong colour correlation, where the total string length is essentially minimized in the final state.

For a long period of time, only one significant change was made to this scenario:

7. Originally the $p_{\perp 0}$ parameter had been assumed energy-independent. In the wake of the HERA

data [4], which led to newer PDF parametrizations having a steeper small-x behaviour than previously assumed, it became necessary to let $p_{\perp 0}$ increase with energy to avoid too steep a rise of the multiplicity. Such an energy dependence can be motivated by colour screening effects [5]. A functional form $p_{\perp 0} \propto s^{\epsilon}$ with $\epsilon \sim 0.08$ is suggested by Pomeron arguments.

Several studies have been presented based on this framework. Some of the recent tuning activities are described elsewhere in this report. The PYTHIA Tune A [6] is a standard reference for much of the current Tevatron underlying-event and minimum-bias physics studies.

In recent years, an effort has been made to go beyond the framework outlined above. Several new or improved components have been introduced.

1. The fragmentation of junction-string topologies has been implemented [7] . Such topologies must be considered when at least two valence quarks are kicked out of an incoming proton beam particle. Here a proton is modelled as a Y-shaped topology, where each valence quarks sits at the end of one of the three legs going out from the middle, the junction. When some ends of this Y are kicked out, also the junction is set in motion. The junction carries no energy or momentum of its own, but it is around the junction that the baryon inheriting the original baryon number will be formed. The junction rest frame is defined by having 120° between the three jets. A number of technical problems have to be overcome in realistic situations, where also gluons may be colour-connected on the three legs, thus giving more complicated space–time evolution patterns.
2. PDFs are more carefully modelled, to take into account the flavour structure of previous interactions [8], not only the overall energy–momentum constraints. Whenever a valence quark is kicked out, the remaining valence PDF of this flavour is rescaled to the new remaining number. When a sea quark is kicked out, an extra "companion" antiquark distribution contribution is inserted, thereby increasing the likelihood that also the antiquark is kicked out.
3. Also remnant flavours are more carefully considered, along with issues such as primordial $k_{\perp}$ values and remnant longitudinal momentum sharing.
4. A few further impact-parameter possibilities are introduced.
5. New transverse-momentum-ordered showers are introduced, both for initial- and final-state radiation (ISR and FSR) [9]. On the one hand, this appears to give an improved description of (hard) multijet production. On the other hand, it allows all evolution to be viewed in terms of a common "time" ordering given by decreasing $p_{\perp}$ values. This is especially critical for the description of MI and ISR, which are in direct competition, in the sense that both mechanisms take momentum out of the incoming beams and thereby require a rescaling of PDF's at later "times". This approach, with interleaved MI and ISR, is illustrated in Fig. 1.

Currently we still make use of two simplifications to the new $p_{\perp}$-ordered framework: (a) the inclusion of FSR is deferred until the MI and ISR have been considered in full, and (b) there is no intertwining, in which two seemingly separate higher-virtuality parton chains turns out to have a common origin when studied at lower $p_{\perp}$ scales. Fortunately there are good reasons why neither of those omitted aspects should be so important.

There is one big remaining unsolved issue in this model, however, namely that of colour flow. If colours are only connected via the fact that the incoming beam remnants are singlets, the correct $\langle p_{\perp} \rangle (n_{\mathrm{charged}})$ behaviour cannot be reproduced whatever variation is tried. It appears necessary to assume that some final-state colour reconnection mechanism tends to reduce the total string length almost to the minimal possible, as was required for Tune A. The most physically reasonable approach, that is yet not too time-consuming to implement, remains to be found. It is possible that also diffractive topologies will need to become a part of this game.

Apart from this big colour issue, and the smaller ones of a complete interleaving/intertwining, PYTHIA now contains a very consistent and complete picture of both minimum-bias and underlying-event physics. It will be interesting to see how this framework fares in comparisons with data. However,

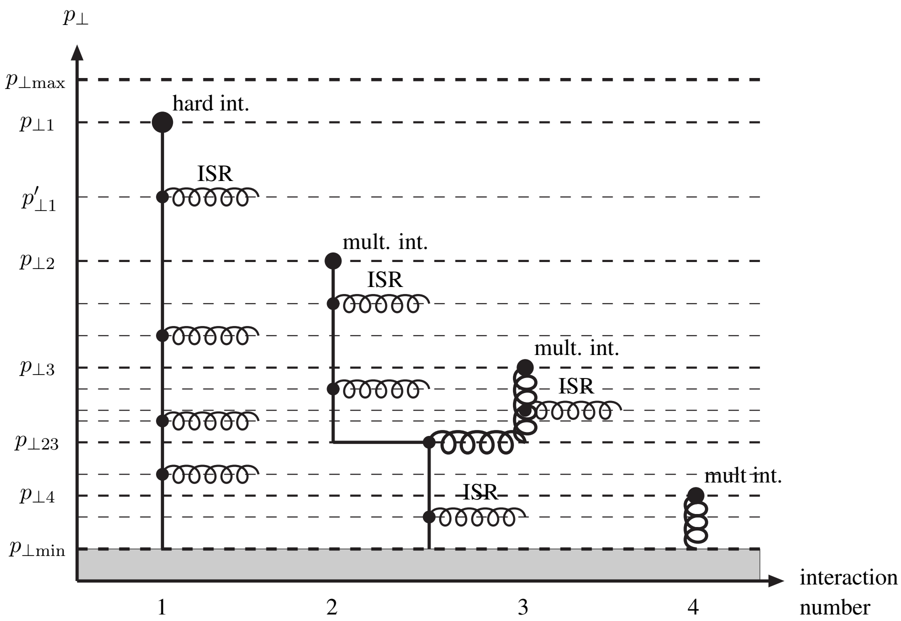


Fig. 1: Schematic figure illustrating one incoming hadron in an event with a hard interaction occurring at $p_{\perp 1}$ and three further interactions at successively lower $p_\perp$ scales, each associated with (the potentiality of) initial-state radiation, and further with the possibility of two interacting partons (2 and 3 here) having a common ancestor in the parton showers. Full lines represent quarks and spirals gluons. The vertical $p_\perp$ scale is chosen for clarity rather than realism; most of the activity is concentrated to small $p_\perp$ values.

if the models appears complex, this complexity is driven by necessity: all of the issues already brought up must be included in the "definitive" description, in one form or other, plus possibly some more not yet brought to light.

2.2 JIMMY

The basic ideas of the eikonal model implemented in JIMMY are discussed elsewhere [10]. The model derives from the observation that for partonic scatters above some minimum transverse momentum, $\hat{p}_T^{\min}$, the values of the hadronic momentum fraction x which are probed decrease as the centre-of-mass energy, s, increases, and since the proton structure function rises rapidly at small x [4], high parton densities are probed. Thus the perturbatively-calculated cross section grows rapidly with s. However, at such high densities, the probability of more than one partonic scattering in a single hadron-hadron event may become significant. Allowing such multiple scatters reduces the total cross section, and increases the activity in the final state of the collisions.

2.2.1 Model Assumptions

The JIMMY model assumes some distribution of the matter inside the hadron in impact parameter (b) space, which is independent of the momentum fraction, x. The multiparton interaction rate is then calculated using the cross section for the hard subprocess, the conventional parton densities, and the area overlap function, $A(b)$. No assumption about the behaviour of the *total* cross section is used. For cross sections other than QCD $2 \to 2$ scatters, JIMMY makes use of approximate formulae, valid when all cross sections except QCD $2 \to 2$ are small, which is true in most cases of interest. This approximation is described in detail elsewhere [11].

2.2.2 *Standard* JIMMY

The starting point for the multiple scattering model is the assertion that, at fixed impact parameter, b, different scatters are independent and so obey Poisson statistics. It is then straightforward to show that the cross section for events in which there are n scatters of type a is given by

$$\sigma_n = \int \mathrm{d}^2b \, \frac{(A(b)\sigma_\mathrm{a})^n}{n!} \, \mathrm{e}^{-A(b)\sigma_\mathrm{a}}, \tag{1}$$

where σ_a is the parton–parton cross section and $A(b)$ is the matter density distribution, obeying

$$\int \mathrm{d}^2b \, A(b) = 1. \tag{2}$$

It is straightforward to show that the inclusive cross section for scatters of type a is σ_a and the total cross section for events with at least one scatter of type a is

$$\sigma_\mathrm{tota} = \int \mathrm{d}^2b \left(1 - \mathrm{e}^{-A(b)\sigma_\mathrm{a}}\right). \tag{3}$$

These can then be combined to give the probability that an event has exactly n scatters of type a, given that it has at least 1 scatter of type a,

$$P_n = \frac{\int \mathrm{d}^2b \, \frac{(A(b)\sigma_\mathrm{a})^n}{n!} \, \mathrm{e}^{-A(b)\sigma_\mathrm{a}}}{\int \mathrm{d}^2b \left(1 - \mathrm{e}^{-A(b)\sigma_\mathrm{a}}\right)}, \qquad n \geq 1. \tag{4}$$

This is the probability distribution pretabulated (as a function of $\sqrt{s}$) by Jimmy.

Jimmy's procedure can then be summarized as:

1. Give all events cross section σ_tota.
2. In a given event, choose n according to Eq. (4).

It is interesting to note that Jimmy's procedure, despite integrating over b once-and-for-all at initialization time, correctly reproduces the correlation between different scatters, whose physical origin is a b-space correlation: small cross section scatters are more likely to come from events with a large overlap and hence be accompanied by a larger-than-average number of large cross section scatters.

2.2.3 *Two Different Scattering Types*

We consider the possibility that there are two different scattering types, but that the cross section for the second type, σ_b, is small enough that events with more than one scatter of type b are negligible. The probability distribution for number of scatters of type a, n, in events with at least one of type b is given by [11]

$$P(n|m \geq 1) = \frac{\int \mathrm{d}^2b \, \frac{(A(b)\sigma_\mathrm{a})^n}{n!} \, \mathrm{e}^{-A(b)\sigma_\mathrm{a}} \left(1 - \mathrm{e}^{-A(b)\sigma_\mathrm{b}}\right)}{\int \mathrm{d}^2b \left(1 - \mathrm{e}^{-A(b)\sigma_\mathrm{b}}\right)}, \qquad n \geq 0. \tag{5}$$

Since σ_b is small, we can expand the exponentials and obtain

$$P(n|m \geq 1) \approx \int \mathrm{d}^2b \, A(b) \, \frac{(A(b)\sigma_\mathrm{a})^n}{n!} \, \mathrm{e}^{-A(b)\sigma_\mathrm{a}}, \qquad n \geq 0. \tag{6}$$

Note that this expression is independent of σ_b. It is therefore ideal for implementing into JIMMY. It is useful to rewrite this equation, as follows. We redefine n to be the total number of scatters, including the one of type b (i.e. "new n"="old n"+1) and rewrite, to obtain

$$P_n \approx \frac{\int \mathrm{d}^2b \, n \, \frac{(A(b)\sigma_\mathrm{a})^n}{n!} \, \mathrm{e}^{-A(b)\sigma_\mathrm{a}}}{\sigma_\mathrm{a}}, \qquad n \geq 1. \tag{7}$$

Note the similarity with Eq. (4), making this form even easier to implement into Jimmy.

The Monte Carlo implementation of this procedure is straightforward:

1. Give all events cross section $\sigma_{\rm b}$.
2. In a given event choose n according to Eq. (7).
3. Generate 1 scatter of type $\rm b$ and $n-1$ of type $\rm a$.

There is one important difference between the cases in which $\rm b$ is distinct from $\rm a$ and $\rm b$ is a subset of $\rm a$: some of the $n-1$ scatters of type $\rm a$ could also be of type $\rm b$. Although this is a small fraction of the total, it can be phenomenologically important. As each scatter of type $\rm a$ is generated, a check is made as to whether it is also of type $\rm b$. The mth scatter of type $\rm b$ generated so far is rejected with probability $1/(m+1)$. This ensures that the proposed algorithm is continuous at the boundary of $\rm b$.

When using JIMMY at the LHC, the tuneable parameters are those described previously [10], with the obvious exception of those parameters which only concern the photon. Those remaining are therefore the minimum transverse momentum of a hard scatter, the proton structure, and the effective radius of the proton. Details on how to adjust these parameters can be found elsewhere [11].

2.3 Simulation of Multiple Interactions in `Sherpa`

Given the studies presented in the following sections, and references therein, current multi-purpose event generators rely heavily on the implementation of multiple parton interaction models to describe the final state in hadronic collisions. To allow `Sherpa` to provide a complete description of hadronic events, the module `AMISIC++` has been developed to simulate multiple parton interactions. This module is capable of simulating multiple scatterings according to the formalism initially presented in [3] and in its current implementation acts as a benchmarking tool to cross-check new multiple interaction models [12].

The basic assumption of the multiple interaction formalism according to T. Sjöstrand and M. van Zijl is, that the differential probability $\mathcal{P}(p_\perp^{\rm out})$ to get a (semi-)hard scattering in the underlying event is given by $\mathcal{P}(p_\perp^{\rm out}) = \sigma_{\rm hard}(p_\perp^{\rm out})/\sigma_{\rm ND}$, where $p_\perp^{\rm out}$ is the transverse momentum of the outgoing partons in the scattering. Since $\sigma_{\rm hard}$ is dominated by $2 \to 2$ processes, the definition of $p_\perp^{\rm out}$ is unambiguous. The specific feature of `AMISIC++` is, that it allows for an independent Q^2-evolution of initial and final state partons in each (semi-)hard scattering via an interface to `Sherpa`'s parton shower module `APACIC++` [13, 14]. The key point here is, that the parton shower must then respect the initial $p_\perp^{\rm out}$ distribution of each (semi-)hard scattering. In particular, it must not radiate partons with $p_\perp > p_\perp^{\rm out}$. The appropriate way to incorporate this constraint is in fact identical to the realisation of the highest multiplicity treatment in the CKKW approach [15, 16, 17, 18]. Our proposed algorithm works as follows:

1. Create a hard scattering process according to the CKKW approach.
   Employ a K_T jet finding algorithm in the E-scheme to define final state jets.
   Stop the jet clustering as soon as there remains only one QCD node to be clustered.
   Set the starting scale of the multiple interaction evolution to $p_\perp$ of this node.
2. Select $p_\perp$ of the next (semi-)hard interaction according to [3].
   If done for the first time in the event, select the impact parameter b of the collision.
3. Set the jet veto scale of the parton shower to the transverse momentum $p_\perp$, selected in 2.
   Start the parton shower at the QCD hard scale $\mu^2_{\rm QCD} \;=\; 2\,stu/\left(s^2+t^2+u^2\right)$.
4. Return to step 2.

The above algorithm works for pure QCD hard matrix elements as well as for electroweak processes in the hard scattering. In the QCD case the selected starting scale for the determination of the first additional interaction reduces to $p_\perp^{\rm out}$ and is thus equal to the original ordering parameter. In the case of electroweak core processes, like single W- or Z-boson production there is no such unique identification. On the other hand the multiple scatterings in the underlying event must not spoil jet topologies described by the hard event through, e.g., using multi-jet matrix elements. However, since the electroweak bosons may be regarded to have been radiated off QCD partons during the parton shower evolution of a hard

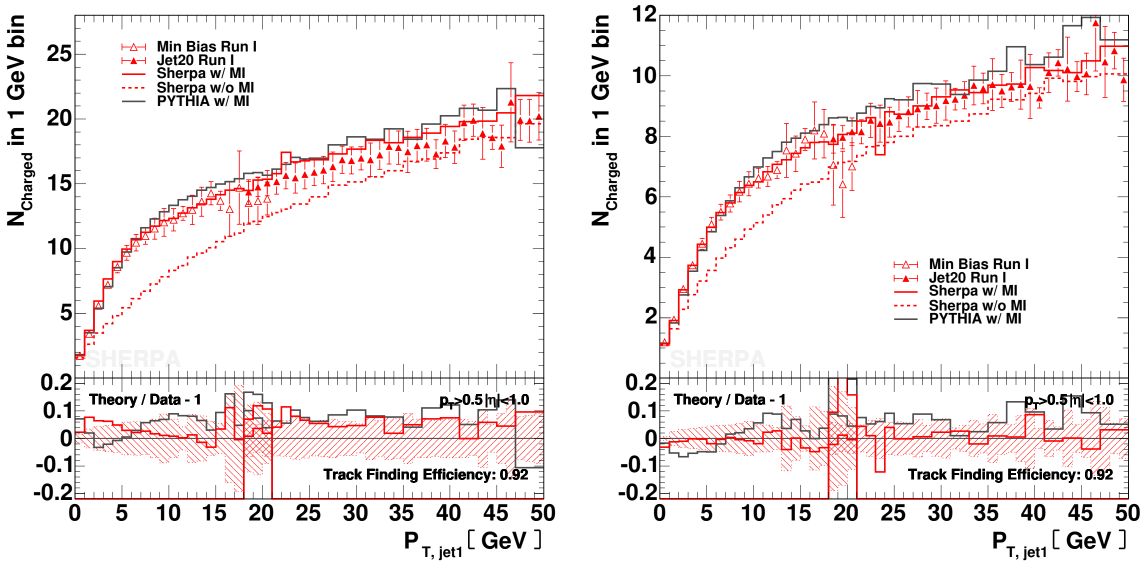


Figure 2: Charged particle multiplicity as a function of P_T of the leading charged particle jet. The left figure shows the total charged particle multiplicity in the selected p_T- and η-range, the right one displays the same in the "Toward" region (for definitions, see section 3 and [20]).

QCD event, it is appropriate to reinterprete the hard matrix element as such a QCD+EW process, where the simplest is a 1-jet process.

An important question in conjunction with the simulation of underlying events is the assignment of colours to final state particles. In the `Sherpa` framework, colour connections in any hard $2 \to 2$ QCD process are chosen according to the kinematics of the process. In particular the most probable colour configuration is selected. Additionally, initial state hadrons are considered to be composed from QCD partons in such a way that the colour string lengths in the final state are minimized. In cases, where it is impossible to realise this constraint, the colour configurations of the hard matrix elements are kept but the configuration of the beam remnants is shuffled until a suitable solution is found.

Figures 1-4 show some preliminary results obtained with the above algorithm, implemented in the current `Sherpa` version, `Sherpa-1.0.6`. We compare the `Sherpa` prediction including multiple interactions to the one without multiple interactions and to the result obtained with PYTHIA 6.214, also including multiple interactions and employing the parameters of PYTHIA Tune A [6]. Shown are hadron-level predictions, which are uncorrected for detector acceptance, except for a uniform track finding efficiency as given in [19]. Data were taken at the Fermilab Tevatron during Run I [20]. Good agreement between the simulations and data is observed only if multiple interactions are included. The mean interaction number in `Sherpa`, including the hard scattering, in this case is $<N_{\rm hard}> = 2.08$, while for `PYTHIA 6.214` it is $<N_{\rm hard}> = 7.35$. The lower interaction number in `Sherpa` can easily be understood, as a decrease of parton multiplicity in the (semi-)hard scatterings due to a rise of the parton multiplicity in the parton showers. PYTHIA 6.214 does not allow for parton showers in the (semi-)hard scatterings in the underlying event. This feature has, however, been added in PYTHIA 6.3 (see section 2.1), and is also present in JIMMY(section 2.2).

2.4 PHOJET

The physics model used in the MC event generator PHOJET combines the ideas of the DPM [21] with perturbative QCD to give an almost complete picture of high-energy hadron collisions [22].

PHOJET is formulated as a two-component model containing contributions from both soft and hard interactions. The DPM is used to describe the dominant soft processes and perturbative QCD is applied

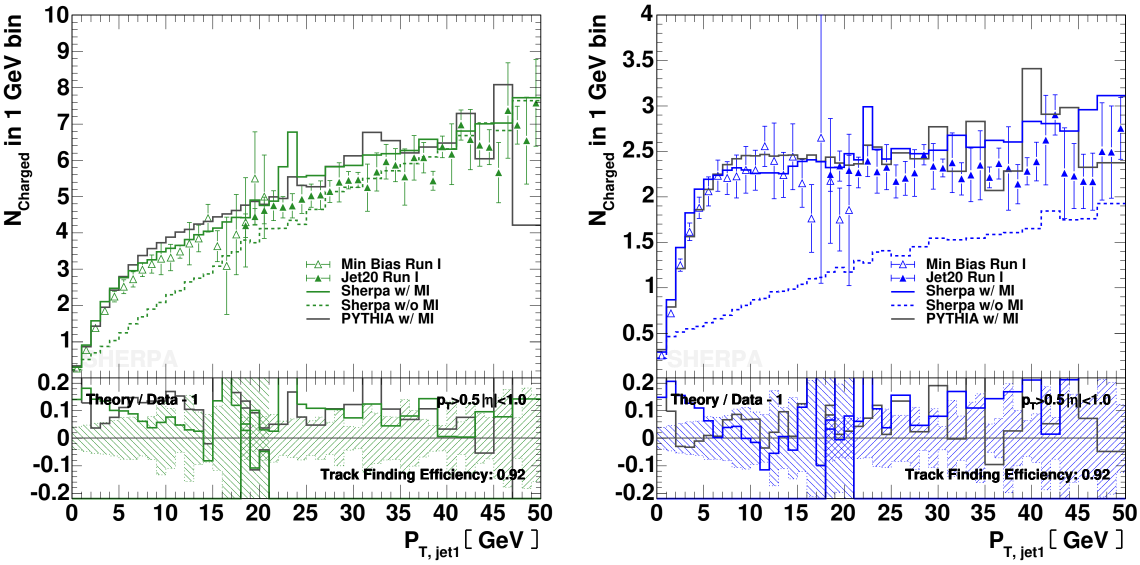


Figure 3: Charged particle multiplicity as a function of P_T of the leading charged particle jet. The left figure shows results for the “Away” side region, the right one displays results for the “Transverse” region.

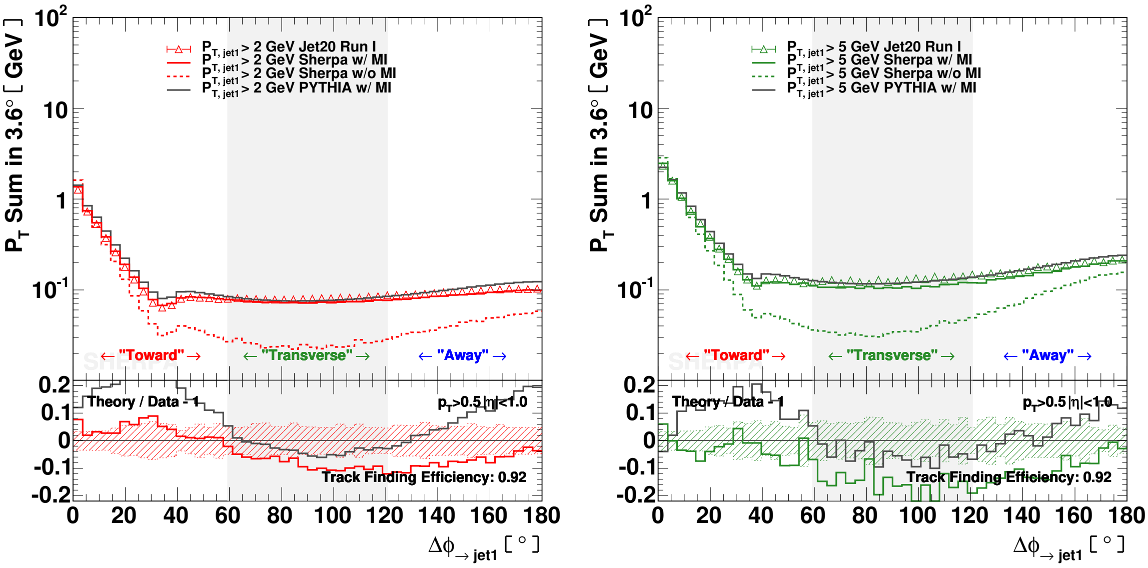


Figure 4: Scalar P_T sum as a function of the azimuthal angle relative to the leading charged particle jet. The left figure shows results for $P_{T,jet1} > 2\,\mathrm{GeV}$, the right one displays results for $P_{T,jet1} > 5\,\mathrm{GeV}$.

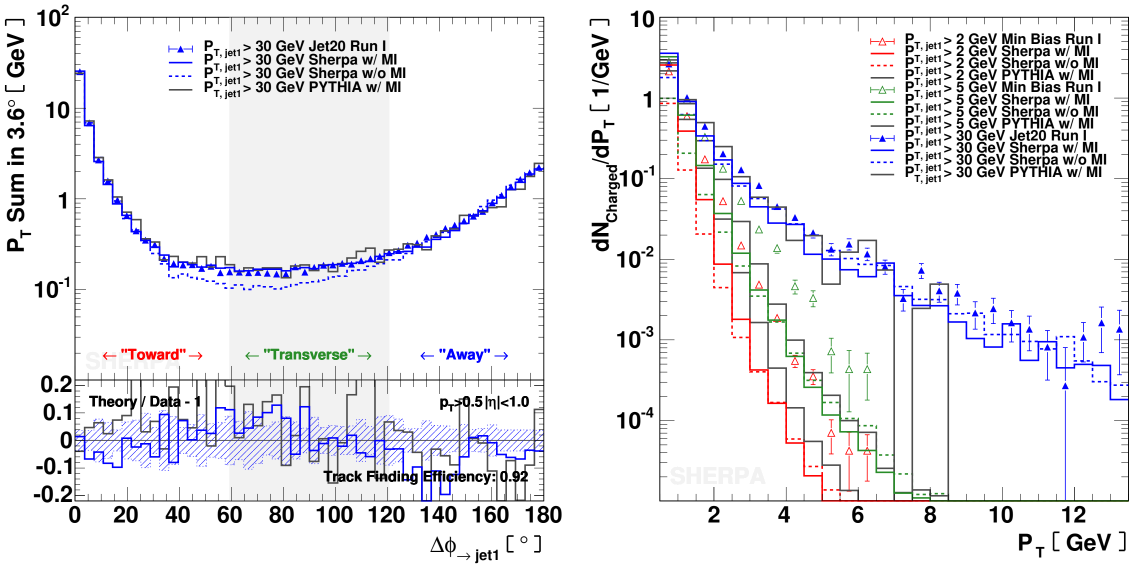


Figure 5: Left: Scalar P_T sum as a function of the azimuthal angle relative to the leading charged particle jet for $P_{T,jet1} > 30\,\text{GeV}$.
Right: Charged particle multiplicity as a function of P_T in the "Transverse" region.

to generate hard interactions.

There has been very little development on PHOJET for the last few years, although it is used quite widely in minimum bias and cosmic ray physics. A major disadvantage for the LHC is that it is not part of a general purpose generator, and therefore cannot be used to generate underlying events to low cross section processes.

3 Tuning PYTHIA and HERWIG/JIMMY in Run 2 at CDF

The behaviour of the charged particle ($p_T > 0.5\,\text{GeV/c}$, $|\eta| < 1$) and energy ($|\eta| < 1$) components of the UE in hard scattering proton-antiproton collisions at 1.96 TeV has been studied at CDF. The goal is to produce data on the UE that is corrected to the particle level, so that it can be used to tune the QCD Monte-Carlo models using tools such as those described in the contributions from Group 5 of this workshop without requiring a simulation of the CDF detector. Unlike the previous CDF Run 2 UE analysis which used JetClu to define "jets" and compared uncorrected data with the QCD Monte-Carlo models after detector simulation (*i.e.*, CDFSIM), this analysis uses the midpoint jet algorithm and corrects the observables to the particle level. The corrected observables are then compared with the QCD Monte-Carlo models at the particle level (*i.e.*, generator level). The QCD Monte-Carlo models include PYTHIA Tune A, HERWIG and a tuned version of JIMMY.

One can use the topological structure of hadron-hadron collisions to study the UE [19, 23, 24]. The direction of the leading calorimeter jet is used to isolate regions of η-ϕ space that are sensitive to the UE. As illustrated in Fig. 6, the direction of the leading jet, jet#1, is used to define correlations in the azimuthal angle, $\Delta\phi$. The angle $\Delta\phi = \phi - \phi_{\text{jet\#1}}$ is the relative azimuthal angle between a charged particle (or a calorimeter tower) and the direction of jet#1. The "transverse" region is perpendicular to the plane of the hard 2-to-2 scattering and is therefore very sensitive to the UE. We restrict ourselves to charged particles in the range $p_T > 0.5\,\text{GeV/c}$ and $|\eta| < 1$ and calorimeter towers with $E_T > 0.1\,\text{GeV}$ and $|\eta| < 1$, but allow the leading jet that is used to define the "transverse" region to have $|\eta(jet\#1)| < 2$. Furthermore, we consider two classes of events. We refer to events in which there are no restrictions placed on the second and third highest P_T jets (jet#2 and jet#3) as "leading jet" events. Events with at least two jets with $P_T > 15\,\text{GeV/c}$ where the leading two jets are nearly "back-to-back" ($|\Delta\phi| > 150°$) with $P_T(jet\#2)/P_T(jet\#1) > 0.8$ and $P_T(jet\#3) < 15\,\text{GeV/c}$ are referred to as "back-to-back" events. “Back-to-back” events are a subset of the “leading jet” events. The idea is to suppress hard initial and final-state radiation thus increasing the sensitivity of the “transverse” region to the “beam-beam remnants” and the multiple parton scattering component of the “underlying event”.

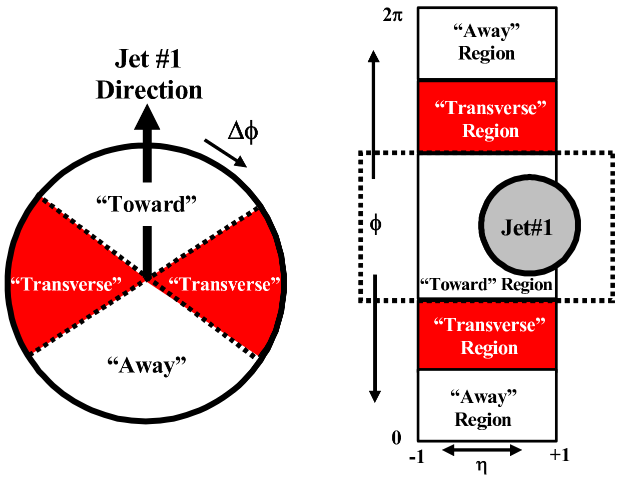


Fig. 6: Illustration of correlations in azimuthal angle ϕ relative to the direction of the leading jet (MidPoint, $R = 0.7$, $f_{merge} = 0.75$) in the event, jet#1. The angle $\Delta\phi = \phi - \phi_{\rm jet1}$ is the relative azimuthal angle between charged particles and the direction of jet#1. The “transverse” region is defined by $60^\circ < |\Delta\phi| < 120^\circ$ and $|\eta| < 1$. We examine charged particles in the range $p_T > 0.5$ GeV/c and $|\eta| < 1$ and calorimeter towers with $|\eta| < 1$, but allow the leading jet to be in the region $|\eta(\text{jet}\#1)| < 2$.

As illustrated in Fig. 7, we define a variety of MAX and MIN “transverse” regions which help to separate the “hard component” (initial and final-state radiation) from the “beam-beam remnant” component. MAX (MIN) refer to the “transverse” region containing largest (smallest) number of charged particles or to the region containing the largest (smallest) scalar PTsum of charged particles or the region containing the largest (smallest) scalar ETsum of particles. Since we will be studying regions in η-ϕ space with different areas, we will construct densities by dividing by the area. For example, the number density, $dN_{chg}/d\phi d\eta$, corresponds to the number of charged particles ($p_T > 0.5$ GeV/c) per unit η-ϕ the PTsum density, $dPT_{sum}/d\phi d\eta$, corresponds to the amount of charged particle ($p_T > 0.5$ GeV/c) scalar PTsum per unit η-ϕ, and the transverse energy density, $dET_{sum}/d\phi d\eta$, corresponds the amount of scalar ETsum of all particles per unit η-ϕ. One expects that the “transMAX” region will pick up the hardest initial or final-state radiation while both the “transMAX” and “transMIN” regions should receive “beam-beam remnant” contributions. Hence one expects the “transMIN” region to be more sensitive to the “beam-beam remnant” component of the “underlying event”, while the “transMAX” minus the “transMIN” (*i.e.*, “transDIF”) is very sensitive to hard initial and final-state radiation. This idea, was first suggested by Bryan Webber and Pino Marchesini [25], and implemented in a paper by Jon Pumplin [26]. This was also studied by Valeria Tano in her CDF Run 1 analysis of maximum and minimum transverse cones [27].

Our previous Run 2 UE analysis [28] used JetClu to define jets and compared uncorrected data with PYTHIA Tune A [6] and HERWIG after detector simulation (*i.e.*, CDFSIM). This analysis uses the MidPoint jet algorithm ($R = 0.7$, $f_{merge} = 0.75$) and corrects the observables to the particle level. The corrected observables are then compared with the QCD Monte-Carlo models at the particle level (*i.e.*, generator level). The models includes PYTHIA Tune A, HERWIG and HERWIG with a tuned version of JIMMY [10]. In addition, for the first time we study the transverse energy density in the “transverse” region.

Fig. 8 compares the data on the density of charged particles and the charged PTsum density in the “transverse” region corrected to the particle level for “leading jet” and “back-to-back” events with

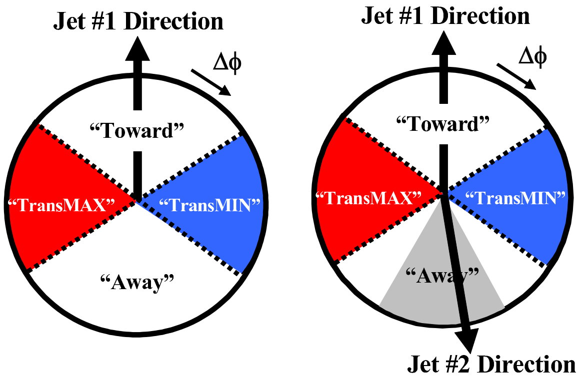


Fig. 7: Illustration of correlations in azimuthal angle ϕ relative to the direction of the leading jet (highest P_T jet) in the event, jet#1. The angle $\Delta\phi = \phi - \phi_{\text{jet}\#1}$ is the relative azimuthal angle between charged particles and the direction of jet#1. On an event by event basis, we define "transMAX" ("transMIN") to be the maximum (minimum) of the two "transverse" regions, $60^\circ < \Delta\phi < 120^\circ$ and $60^\circ < -\Delta\phi < 120^\circ$. "transMAX" and "transMIN" each have an area in η-ϕ space of $\Delta\eta\Delta\phi = 4\pi/6$. The overall "transverse" region defined in Fig. 6 contains both the "transMAX" and the "transMIN" regions. Events in which there are no restrictions placed on the second and third highest p_T jets (jet#2 and jet#3) are referred to as "leading jet" events (*left*). Events with at least two jets with $p_T > 15\,\text{GeV/c}$ where the leading two jets are nearly "back-to-back" ($|\Delta\phi| > 150^\circ$) with $p_T(\text{jet}\#2)/p_T(\text{jet}\#1) > 0.8$ and $p_T(\text{jet}\#3) < 15\,\text{GeV/c}$ are referred to as "back-to-back" events (*right*).

PYTHIA Tune A and HERWIG at the particle level. As expected, the "leading jet" and "back-to-back" events behave quite differently. For the "leading jet" case the "transMAX" densities rise with increasing $P_T(\text{jet}\#1)$, while for the "back-to-back" case they fall with increasing $P_T(\text{jet}\#1)$. The rise in the "leading jet" case is, of course, due to hard initial and final-state radiation, which has been suppressed in the "back-to-back" events. The "back-to-back" events allows a closer look at the "beam-beam remnant" and multiple parton scattering component of the UE. PYTHIA Tune A, which includes multiple parton interactions, does a better job of describing the data than HERWIG which does not have multiple parton interactions.

The "transMIN" densities are more sensitive to the "beam-beam remnant" and multiple parton interaction component of the "underlying event". The "back-to-back" data show a decrease in the "transMIN" densities with increasing $P_T(\text{jet}\#1)$ which is described fairly well by PYTHIA Tune A (with multiple parton interactions) but not by HERWIG (without multiple parton interactions). The decrease of the "transMIN" densities with increasing $P_T(\text{jet}\#1)$ for the "back-to-back" events is very interesting and might be due to a "saturation" of the multiple parton interactions at small impact parameter. Such an effect is included in PYTHIA Tune A but not in HERWIG (without multiple parton interactions).

Fig. 9(left) compares the data on average p_T of charged particles in the "transverse" region corrected to the particle level for "leading jet" and "back-to-back" events with PYTHIA Tune A and HERWIG at the particle level. Again the "leading jet" and "back-to-back" events behave quite differently.

Fig. 9(right) shows the data corrected to the particle level for the scalar $ET\text{sum}$ density in the "transverse" region for "leading jet" and "back-to-back" events compared with PYTHIA Tune A and HERWIG. The scalar $ET\text{sum}$ density has been corrected to correcpond to all particles (all p_T, $|\eta| < 1$). Neither PYTHIA Tune A nor HERWIG produce enough energy in the "transverse" region. HERWIG has more "soft" particles than PYTHIA Tune A and does slightly better in describing the energy density in the "transMAX" and "transMIN" regions.

Fig. 10(left) shows the difference of the "transMAX" and "transMIN" regions ("transDIF" = "transMAX" minus "transMIN") for "leading jet" and "back-to-back" events compared with PYTHIA Tune A and HERWIG. "TransDIF" is more sensitive to the hard scattering component of the UE (*i.e.*, initial and final state radiation). Both PYTHIA Tune A and HERWIG underestimate the energy density

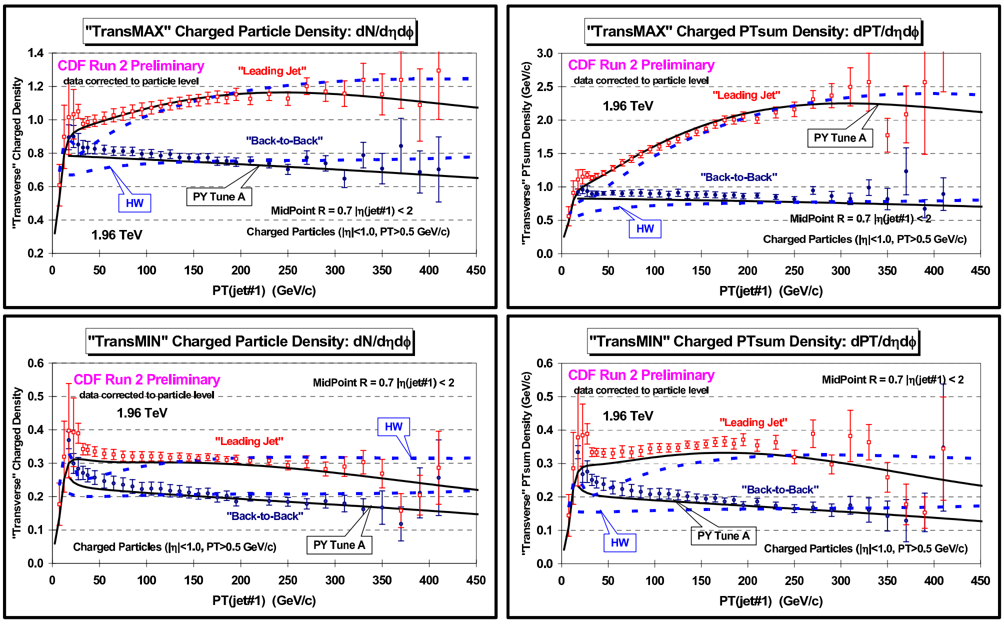


Fig. 8: Data at 1.96 TeV on (left) the density of charged particles $dN_{chg}/d\phi d\eta$ and (right) on the scalar PTsum density of charged particles, with $p_T > 0.5\,\mathrm{GeV/c}$ and $|\eta| < 1$ in the "transMAX" region (*top*) and the "transMIN" region (*bottom*) for "leading jet" and "back-to-back" events defined in Fig. 7 as a function of the leading jet P_T compared with PYTHIA Tune A and HERWIG. The data are corrected to the particle level (with errors that include both the statistical error and the systematic uncertainty) and compared with the theory at the particle level (*i.e.*, generator level).

in the "transMAX" and "transMIN" regions (see Fig. 9). However, they both fit the "transDIF" energy density. This indicates that the excess energy density seen in the data probably arises from the "soft" component of the UE (*i.e.*, beam-beam remnants and/or multiple parton interactions).

JIMMY is a model of multiple parton interaction which can be combined with HERWIG to enhance the UE thereby improving the agreement with data. Fig. 10(right) and Fig. 11(left) show the energy density and charged PTsum density, respectively, in the "transMAX" and "transMIN" regions for "leading jet" and "back-to-back" events compared with PYTHIA Tune A and a tuned version of JIMMY˙JIMMY was tuned to fit the "transverse" energy density in "leading jet" events ($PTJIM = 3.25\,\mathrm{GeV/c}$). The default JIMMY ($PTJIM = 2.5\,\mathrm{GeV/c}$) produces too much energy and too much charged PTsum in the "transverse" region. Tuned JIMMY does a good job of fitting the energy and charged PTsum density in the "transverse" region (although it produces slightly too much charged PTsum at large $P_T(\mathrm{jet}\#1)$). However, the tuned JIMMY produces too many charged particles with $p_T > 0.5\,\mathrm{GeV/c}$ (see Fig. 11(right)). The particles produced by this tune of JIMMY are too soft. This can be seen clearly in Fig. 12 which shows the average charge particle p_T in the "transverse" region.

The goal of this analysis is to produce data on the UE that is corrected to the particle level so that it can be used to tune the QCD Monte-Carlo models without requiring CDF detector simulation. Comparing the corrected observables with PYTHIA Tune A and HERWIG at the particle level (*i.e.*, generator level) leads to the same conclusions as we found when comparing the uncorrected data with the Monte-Carlo models after detector simulation [28]. PYTHIA Tune A (with multiple parton interactions) does a better job in describing the UE (*i.e.*, "transverse" regions) for both "leading jet" and "back-to-back" events than does HERWIG (without multiple parton interactions). HERWIG does not have enough activity in the UE for $P_T(\mathrm{jet}\#1)$ less than about 150 GeV/c, which was also observed in our published Run 1 analysis [19].

This analysis gives our first look at the energy in the UE (*i.e.*, the "transverse" region). Neither

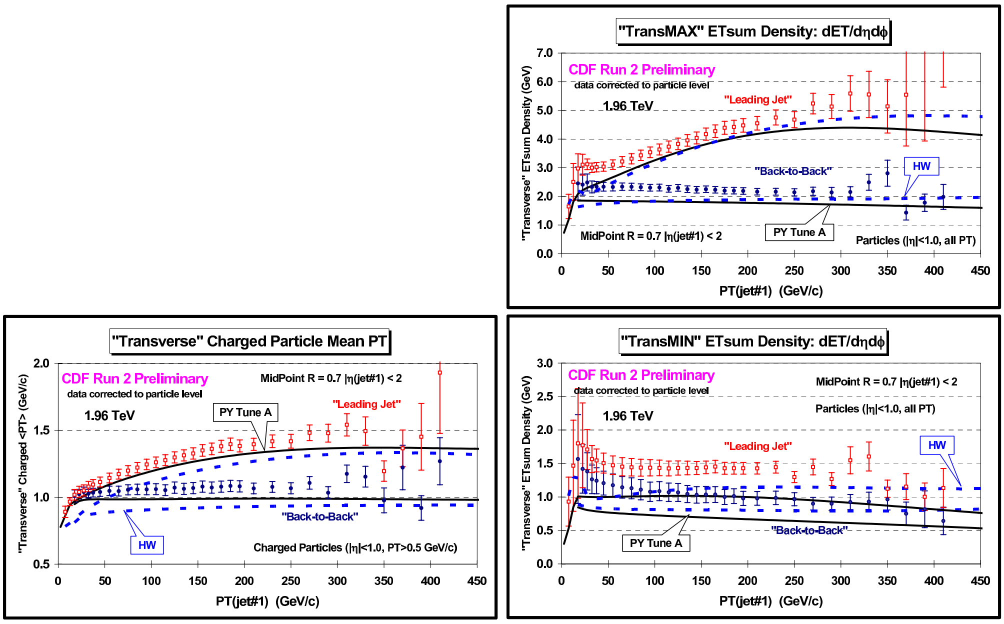


Fig. 9: On the left, data at 1.96 TeV on average transverse momentum, $\langle p_T \rangle$, of charged particles $|\eta| < 1$ in the with with $p_T > 0.5\,\mathrm{GeV/c}$ and $|\eta| < 1$ in the "transverse" region. On the right, scalar ETsum density, $dET_{sum}/d\phi d\eta$, for particles. with $p_T > 0.5\,\mathrm{GeV/c}$ and $|\eta| < 1$ in the "transMAX" region or the "transMIN" region. The "leading jet" and "back-to-back" events are defined in Fig. 7, and the data are shown as a function of the leading jet P_T and compared with PYTHIA Tune A and HERWIG. The data are corrected to the particle level (with errors that include both the statistical error and the systematic uncertainty) and compared with the theory at the particle level (*i.e.*, generator level).

PYTHIA Tune A nor HERWIG produce enough transverse energy in the "transverse" region. However, they both fit the "transDIF" energy density ("transMAX" minus "transMIN"). This indicates that the excess energy density seen in the data probably arises from the "soft" component of the UE (*i.e.*, beam-beam remnants and/or multiple parton interactions). HERWIG has more "soft" particles than PYTHIA Tune A and does slightly better in describing the energy density in the "transMAX" and "transMIN" regions. Tuned JIMMY does a good job of fitting the energy and charged PTsum density in the "transverse" region (although it produces slightly too much charged PTsum at large $P_T(\mathrm{jet}\#1)$). However, the tuned JIMMY produces too many charged particles with $p_T > 0.5\,\mathrm{GeV/c}$ indicating that the particles produced by this tuned JIMMY are too soft.

In summary, we see an interesting dependence of the UE on the transverse momentum of the leading jet (*i.e.*, the Q^2 of the hard scattering). For the "leading jet" case the "transMAX" densities rise with increasing $P_T(\mathrm{jet}\#1)$, while for the "back-to-back" case they fall with increasing $P_T(\mathrm{jet}\#1)$. The rise in the "leading jet" case is due to hard initial and final-state radiation with $p_T > 15\,\mathrm{GeV/c}$, which has been suppressed in the "back-to-back" events. The "back-to-back" data show a decrease in the "transMIN" densities with increasing $P_T(\mathrm{jet}\#1)$. The decrease of the "transMIN" densities with increasing $P_T(\mathrm{jet}\#1)$ for the "back-to-back" events is very interesting and might be due to a "saturation" of the multiple parton interactions at small impact parameter. Such an effect is included in PYTHIA Tune A (with multiple parton interactions) but not in HERWIG (without multiple parton interactions). PYTHIA Tune A does predict this decrease, while HERWIG shows an increase (due to increasing initial and final state radiation).

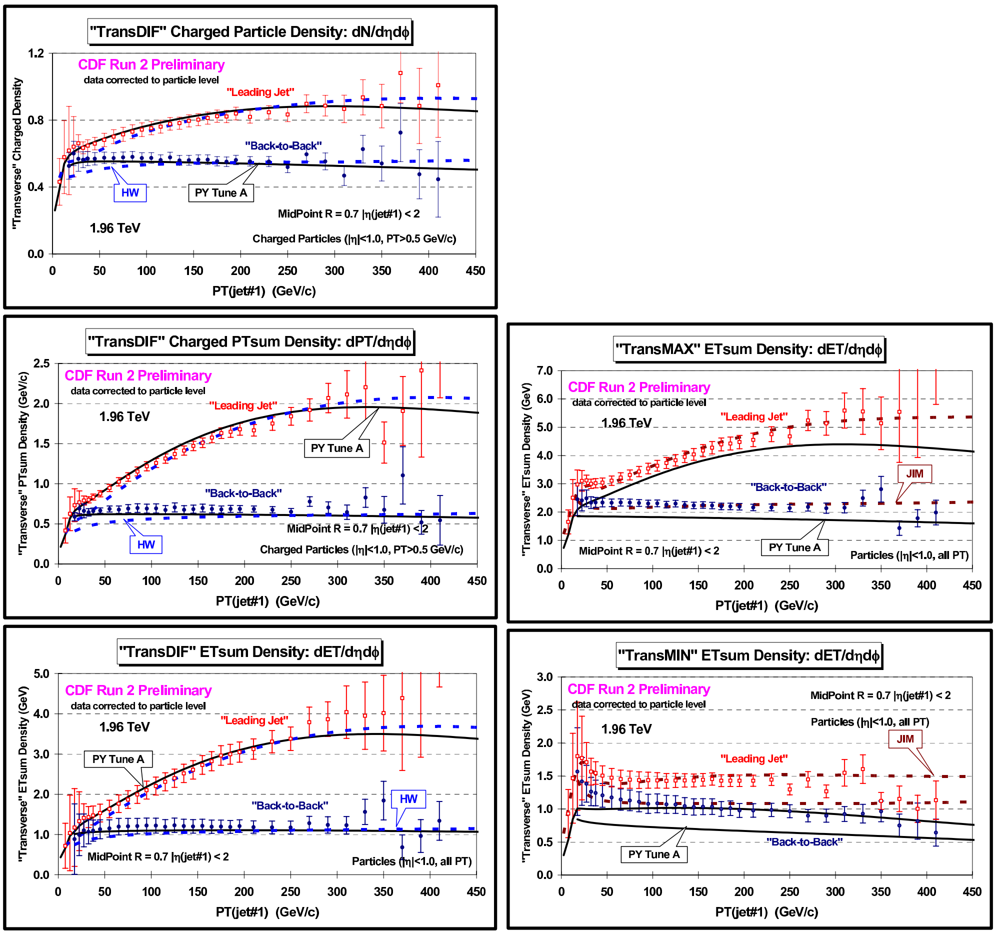


Fig. 10: Left: Data at 1.96 TeV on the difference of the "transMAX" and "transMIN" regions ("transDIF" = "transMAX"-"transMIN") for "leading jet" and "back-to-back" events defined in Fig. 7 as a function of the leading jet P_T compared with PYTHIA Tune A and HERWIG.
Right: Data on scalar ETsum density, $dET_{sum}/d\phi d\eta$, for particles with $|\eta| < 1$ in the "transMAX" region (*top*) and the "transMIN" region (*bottom*) for "leading jet" and "back-to-back" events defined in Fig. 7 as a function of the leading jet P_T compared with PYTHIA Tune A and tuned JIMMY. JIMMY was tuned to fit the "transverse" energy density in "leading jet" events ($PTJIM = 3.25$ GeV/c). The data are corrected to the particle level (with errors that include both the statistical error and the systematic uncertainty) and compared with the theory at the particle level (*i.e.*, generator level).

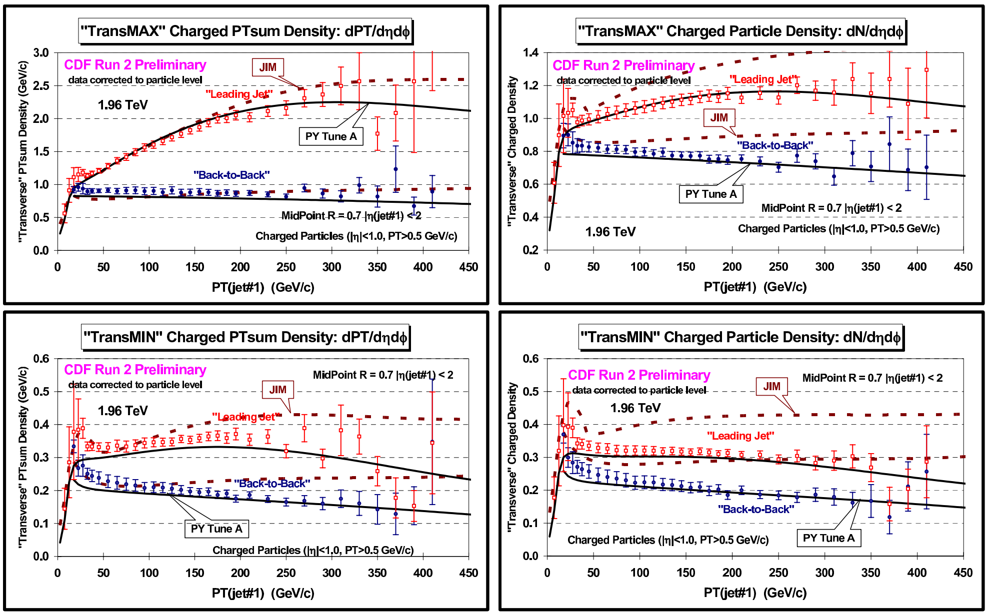


Fig. 11: Left: Data at 1.96 TeV on scalar PTsum density of charged particles, $dPT_{sum}/d\phi d\eta$, with $p_T > 0.5\,\mathrm{GeV/c}$ and $|\eta| < 1$ in the "transMAX" region (*top*) and the "transMIN" region (*bottom*) for "leading jet" and "back-to-back" events defined in Fig. 7 as a function of the leading jet P_T compared with PYTHIA Tune A and tuned JIMMY. JIMMY was tuned to fit the "transverse" energy density in "leading jet" events ($PTJIM = 3.25$ GeV/c). Right: Data on the density of charged particles, $dN_{chg}/d\phi d\eta$, with $p_T > 0.5\,\mathrm{GeV/c}$ and $|\eta| < 1$ in the "transMAX" region (*top*) and the "transMIN" region (*bottom*) for "leading jet" and "back-to-back" events defined in Fig. 2 as a function of the leading jet P_T compared with PYTHIA Tune A and tuned JIMMY. JIMMY was tuned to fit the "transverse" energy density in "leading jet" events ($PTJIM = 3.25$ GeV/c). The data are corrected to the particle level (with errors that include both the statistical error and the systematic uncertainty) and compared with the theory at the particle level (*i.e.*, generator level).

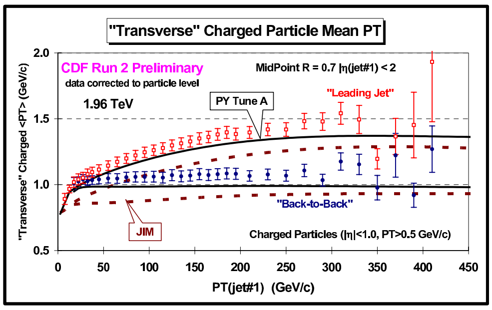


Fig. 12: Data at 1.96 TeV on average transverse momentum, $\langle p_T \rangle$, of charged particles with $p_T > 0.5\,\mathrm{GeV/c}$ and $|\eta| < 1$ in the "transverse" region for "leading jet" and "back-to-back" events defined in Fig. 7 as a function of the leading jet P_T compared with PYTHIA Tune A and tuned JIMMY. JIMMY was tuned to fit the "transverse" energy density in "leading jet" events ($PTJIM = 3.25$ GeV/c). The data are corrected to the particle level (with errors that include both the statistical error and the systematic uncertainty) and compared with the theory at the particle level (*i.e.*, generator level).

4 Extrapolation to LHC energies

The LHCb experiment [29] is designed to measure CP violation in the B-quark sector at the LHC and expand the current studies underway at the B-factories (BaBar, Belle) and at the Tevatron (CDF, D0). At $\sqrt{s}$=1.8 TeV, 28% of all of the primary produced B-mesons in $p\bar{p}$ collisions are produced in L=1 excited states [30]. These excited states decay via the emission of a charged hadron, allowing the possibility of same-side-tagging (SST) studies. As such, it is important to simulate the production of B mesons as accurately as possible.

The production of primary produced excited meson states are not included in the default PYTHIA [31] settings and including them increases the average multiplicity of an event. An attempt to reproduce the HFAG [32] values whilst retaining the spin counting rule for B** states has been made. This note covers a preliminary re-tuning [33] of PYTHIA v6.224 including these settings.

4.1 Method

The main parameter of the multiple-interaction model in PYTHIA v6.224 is the $\hat{p}_T^{\min}$ parameter, which defines the minimum transverse momentum of the parton-parton interactions. This effectively controls the number of parton-parton collisions and hence the average track multiplicity.

The charged particle density measured at $\eta = 0$ in the range of centre-of-mass energies, 52 GeV $< \sqrt{s} <$ 1800 GeV, [34] [35] is used to tune the $\hat{p}_T^{\min}$ parameter of PYTHIA. We define $\rho = \frac{1}{N_{ev}} \frac{dN_{ch}}{d\eta}|_{\eta=0}$ and measure ρ for a range of $\hat{p}_T^{\min}$ values at each $\sqrt{s}$. The quantity $\delta = \rho_{MC} - \rho_{Data}$ is plotted against $\hat{p}_T^{\min}$ and a linear fit performed. In Fig. 13, the re-tuned value of $\hat{p}_T^{\min}$ at $\sqrt{s} = 900$ GeV is taken to be the point at which the fit crosses the $\hat{p}_T^{\min}$ axis. To extrapolate $\hat{p}_T^{\min}$ to LHC energy, a fit is performed (Figure 14) using the form suggested by PYTHIA:

$$\hat{p}_T^{\min} = \hat{p}_T^{\min}(LHC)\Big(\frac{\sqrt{s}}{14TeV}\Big)^{2\epsilon} \tag{8}$$

4.2 Results

Extrapolating to 14 TeV using the tuned values of $\hat{p}_T^{\min}(\sqrt{s})$ and (8), we obtain $\hat{p}_T^{\min}(LHC) = 3.34 \pm 0.13$, with $\epsilon = 0.079 \pm 0.0006$ with a corresponding central multiplicity of $\rho = 6.45 \pm 0.25$. Comparing the output of the re-tuned settings (dashed line) to the old LHCb settings (solid line), *Fig.* 15, 16 and 17, we find that the re-tuned settings produce a slightly lower multiplicity which affects the other distributions accordingly. Note: both the fragmentation parameters and the $\hat{p}_T^{\min}$ parameter affect the multiplicity of a generated event. This re-tuning method varies the $\hat{p}_T^{\min}$ parameter only i.e. it does not alter the fragmentation parameters in any fashion. Further investigations into re-tuning the fragmentation parameters using data from LEP are underway.

4.3 Conclusions

The central multiplicity values measured at CDF and UA5 are accurately reproduced using the re-tuned values for $\hat{p}_T^{\min}$ at several $\sqrt{s}$. An extrapolation of $\hat{p}_T^{\min}$ to LHC energies using a form implemented in PYTHIA gives $\hat{p}_T^{\min}(LHC) = 3.34 \pm 0.13$, with $\epsilon = 0.079 \pm 0.0006$ with a corresponding central multiplicity of $\rho_{LHC} = 6.45 \pm 0.25$ in non-diffractive events.

5 Tuned models for the underlying event and minimum bias interactions

In this section we compare tuned MC generator models for the underlying event and minimum bias interactions. The aim of this study is to predict the event activity of minimum bias and the underlying event at the LHC. The models investigated correspond to tuned versions of PYTHIA, PHOJET and JIMMY.

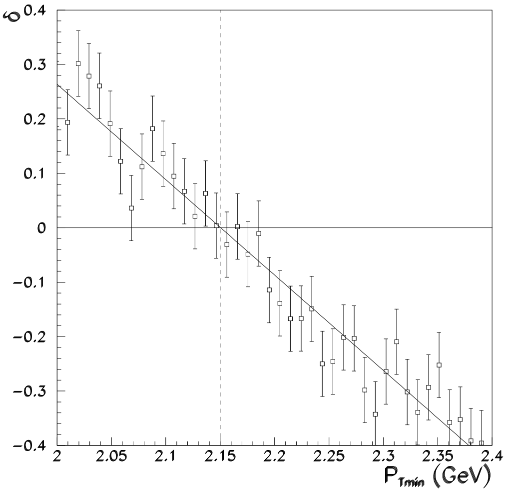


Fig. 13: Determining the value of $\hat{p}_T^{\min}(\sqrt{s} = 900GeV)$, the dashed line shows the point at which $|\delta|$ is minimised.

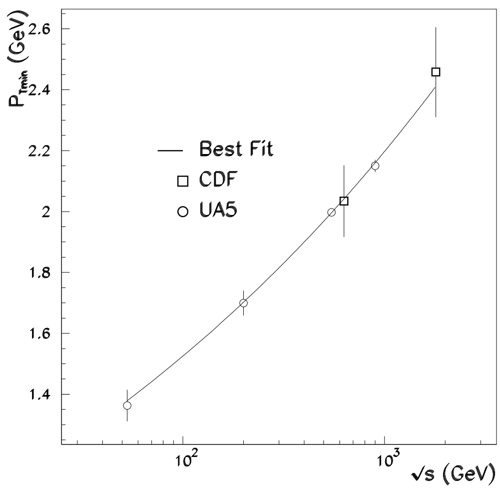


Fig. 14: The $\sqrt{s}$ dependance of $\hat{p}_T^{\min}$. The curve is the result of a fit assuming the functional form of (8).

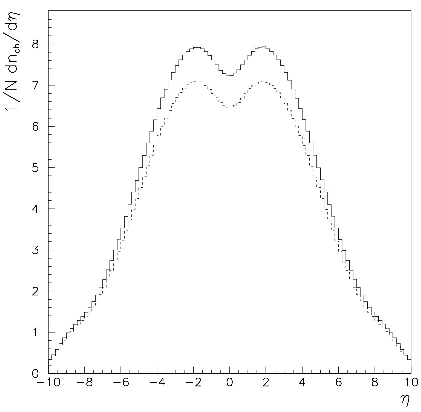


Fig. 15: The η distribution at 14 TeV using the extrapolated value of $P_{T_{Min}}$

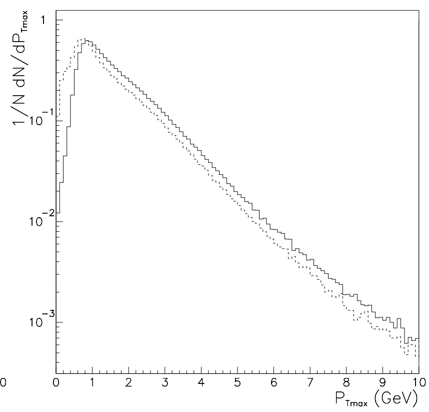


Fig. 16: The $p_{\perp_{max}}$ distribution in the LHCb acceptance

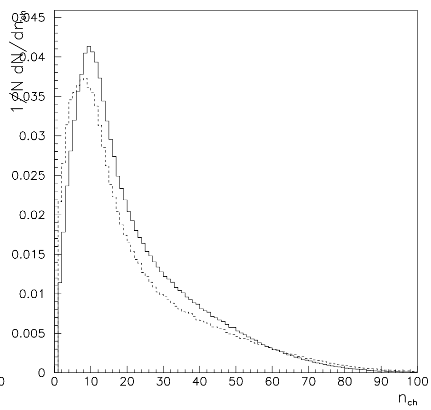


Fig. 17: Charged-stable multiplicity distribution in the LHCb acceptance.

5.1 Tuned models for the underlying event and minimum bias interactions

The starting point for the event generation in PYTHIA and JIMMY is the description of multiple hard interactions in the hadronic collision described in section 2.1 (for PYTHIA 6.2), section 2.2 for JIMMY and section 2.4 for PHOJET.

PYTHIA and PHOJET have been shown to describe both minimum bias and underlying event data reasonably well when appropriately tuned [3, 36, 37, 6]. JIMMY is limited to the description of the underlying event; again, it has been shown capable of describing this rather well [38].

5.2 PYTHIA tunings

Several minimum bias and underlying event (UE) tunings for PYTHIA have been proposed in recent years. Ref. [37] describes how the current ATLAS tuning for PYTHIA was obtained after extensive comparisons to a variety of experimental measurements made at different colliding energies. Similar work has been done by the CDF Collaboration, although their PYTHIA tuning, CDF tune A [6], is primarily based on the description of the underlying event in jet events measured for $p\overline{p}$ at $\sqrt{s}$ = 1.8 TeV.

Default [31]	**ATLAS** [37]	**CDF tune A** [6]	**Comments**
MSTP(51)=7	MSTP(51)=7	MSTP(51)=7	CTEQ5L - selected p.d.f.
MSTP(81)=1	MSTP(81)=1	MSTP(81)=1	multiple interactions
MSTP(82)=1	MSTP(82)=4	MSTP(82)=4	complex scenario + double Gaussian matter distribution
PARP(67)=1	PARP(67)=1	PARP(67)=4	parameter regulating initial state radiation
PARP(82)=1.9	PARP(82)=1.8	PARP(82)=2.0	$p_{t_{\min}}$ parameter
PARP(84)=0.2	PARP(84)=0.5	PARP(84)=0.4	hadronic core radius (only for MSTP(82)=4)
PARP(85)=0.33	PARP(85)=0.33	PARP(85)=0.9	probab. for gluon production with colour connection to nearest neighbours
PARP(86)=0.66	PARP(86)=0.66	PARP(86)=0.95	probab. to produce gluons either either as in PARP(85) or as a closed gluon loop
PARP(89)=1.0	PARP(89)=1.0	PARP(89)=1.8	energy scale (TeV) used to calculate $p_{t_{\min}}$
PARP(90)=0.16	PARP(90)=0.16	PARP(90)=0.25	power of the energy dependence of $p_{t_{\min}}$

Table 1: PYTHIA 6.214 default, ATLAS and CDF tune A parameters for minimum bias and the underlying event.

Table 1 displays the relevant parameters tuned to the data as proposed by the ATLAS [37] and CDF [6] collaborations. For the purpose of comparison, the corresponding default values [31] are also shown in the table.

5.3 PHOJET

The parameters used in PHOJET to describe minimum bias and the underlying event can be found in Ref. [22] and are currently set as default in PHOJET1.12, which is used in this study.

5.4 JIMMY tunings

We have tuned JIMMY to describe the UE as measured by CDF [19] and the resulting sets of parameters are shown in table 2. Figure 18 shows JIMMY predictions for the UE compared to CDF data for the average charged particle multiplicity (a) and the average p_t sum in the underlying event (b). In Fig.18 we compare JIMMY - default parameters to “Tuning A” and “Tuning B”. Note that for the default parameters JIMMY does not give a correct description of the data. The other two distributions, generated with tuning A and B parameters, agree fairly well with the data.

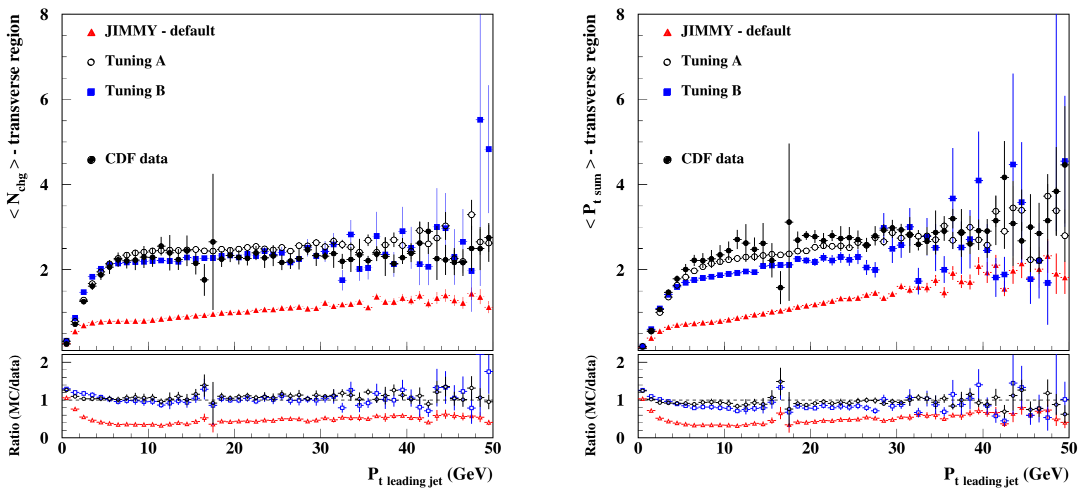
In this study, JIMMY - tuning A and B will only be used to generate LHC predictions for the underlying event associated to jet events.

5.5 Minimum bias interactions at the LHC

Throughout this report, minimum bias events will be associated with non-single diffractive inelastic interactions, following the experimental trend (see Ref. [37] and references therein).

Default	Tuning A	Tuning B	Comments
JMUEO=1	JMUEO=0	JMUEO=0	multiparton interaction model
PTMIN=10.0	PTMIN=3.0	PTMIN=2.0	minimum p_T in hadronic jet production
PTJIM=3.0	-	-	minimum p_T of secondary scatters when JMUEO=1 or 2
JMRAD(73)=0.71	JMRAD(73)=2.13	JMRAD(73)=0.71	inverse proton radius squared
PRSOF=1.0	PRSOF=0.0	PRSOF=0.0	probability of a soft underlying event

Table 2: JIMMY 4.1 default, tunings A and B parameters for the underlying event.



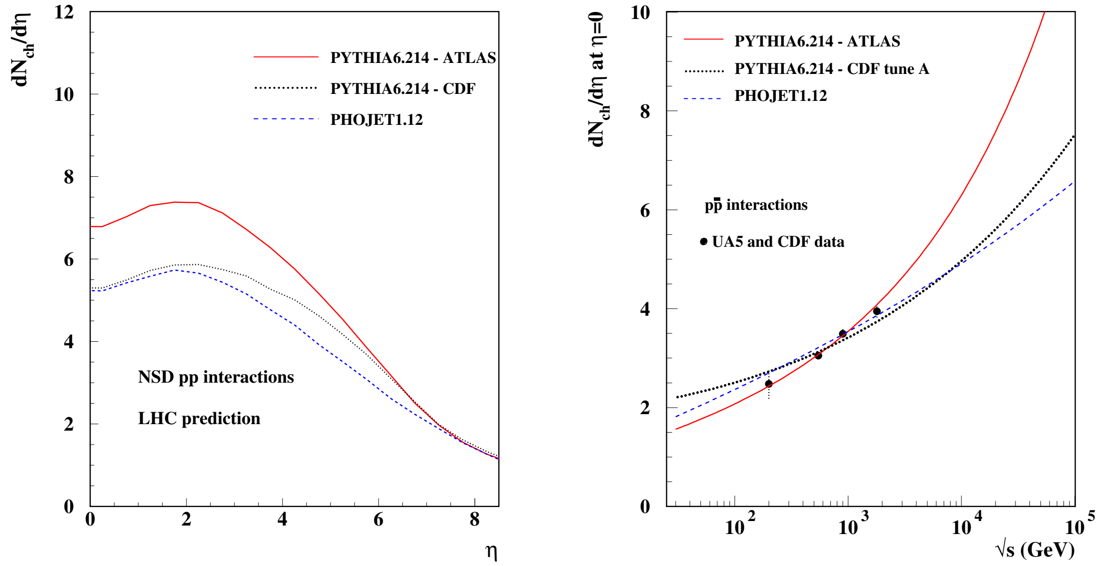
(a) (b)

Fig. 18: JIMMY predictions for the UE compared to CDF data. (a) Average charged particles multiplicity in the UE and (b) average p_t sum in the UE.

For LHC collisions (pp collisions at $\sqrt{s}$ = 14 TeV) the minimum bias cross-section estimated by PYTHIA 6.214, regardless of which tuning is used, is σ_{nsd} = 65.7 mb while PHOJET1.12 predicts σ_{nsd} = 73.8 mb, 12.3% greater than the former. Hence, for the same luminosity PHOJET1.12 generates more minimum bias pp collisions than PYTHIA 6.214 - tuned. We shall however, focus on the general properties per pp collision not weighted by cross-sections. The results per pp collision can later be easily scaled by the cross-section and luminosity.

Figure 19(a) shows charged particle density distributions in pseudorapidity for minimum bias pp collisions at $\sqrt{s}$ = 14 TeV generated by PHOJET1.12 and PYTHIA 6.214 - ATLAS and CDF tune A. The charged particle density generated by PHOJET1.12 and PYTHIA 6.214 - CDF tune A and ATLAS at $\eta = 0$ is 5.1, 5.3 and 6.8, respectively. Contrasting to the agreement seen in previous studies for $p\bar{p}$ collisions at $\sqrt{s}$ = 200 GeV, 546 GeV, 900 GeV and 1.8 TeV in Ref. [37], at the LHC PYTHIA 6.214 - ATLAS generates $\sim$ 25% more charged particle density in the central region than PYTHIA 6.214 - CDF tune A and PHOJET1.12.

Compared to the charged particle density $dN_{ch}/d\eta$ measured by the CDF experiment at 1.8 TeV



(a) (b)

Fig. 19: (a) Charged particle density distributions, $dN_{ch}/d\eta$, for NSD pp collisions at $\sqrt{s}$ = 14 TeV. (b) $dN_{ch}/d\eta$ at $\eta = 0$ for a wide range of $\sqrt{s}$. Predictions generated by PYTHIA 6.214 - ATLAS and CDF tune A and PHOJET1.12 .

[39], PYTHIA 6.214 - ATLAS indicates a plateau rise of $\sim 70\%$ at the LHC in the central region while PHOJET1.12 and PYTHIA 6.214 - CDF tune A suggest a smaller rise of $\sim 35\%$.

Figure 19(b) displays $dN_{ch}/d\eta$ at $\eta = 0$ plotted as a function of $\sqrt{s}$. For centre-of-mass energies greater than $\sim$ 1 TeV, the multiparton interaction model employed by PYTHIA and the DPM used by PHOJET lead to multiplicity distributions with different rates of increase with the energy. PYTHIA suggests a rise dominated by the $\ln^2(s)$ term while PHOJET predicts that the dominant term gives a $\ln(s)$ rise for $dN_{ch}/d\eta$ at $\eta = 0$. The ATLAS tuning for PYTHIA gives a steeper rise than CDF tune A and PHOJET (Fig. 19(b)) indicating a faster increase in the event activity at the partonic level in the ATLAS tuning when compared to CDF tune A and PHOJET. The average charged particle multiplicity in LHC minimum bias collisions, $< n_{ch} >$, is 69.6, 77.5 and 91.0 charged particles as predicted by PHOJET1.12, PYTHIA 6.214 - CDF tune A and ATLAS, respectively.

The $< p_t >$ at $\eta = 0$ for charged particles in LHC minimum bias collisions predicted by PHOJET1.12 and PYTHIA 6.214 - ATLAS and CDF tune A models is 0.64 GeV, 0.67 GeV and 0.55 GeV, respectively. Generating less particles in an average minimum bias collision at the LHC, PHOJET1.12 predicts that the average p_t per particle at $\eta = 0$ is greater (or harder) than the corresponding prediction from PYTHIA 6.214 - ATLAS. However, amongst the three models, PYTHIA 6.214 - CDF tune A gives the hardest $< p_t >$ at $\eta = 0$. The main reason for this is the increased contribution of harder parton showers used to make the model agree with the p_t spectrum of particles in the UE, and obtained by setting PARP(67)=4 [6].

5.6 The underlying event

Based on CDF measurements, we shall use their definition for the UE, i.e. the angular region in ϕ which is transverse to the leading charged particle jet as described in section 3 and shown in figure 6. Figure 20(a) displays PYTHIA 6.214 - ATLAS and CDF tune A, and PHOJET1.12 predictions for the average particle multiplicity in the UE for pp collisions at the LHC (charged particles with $p_T > 0.5$ GeV and $|\eta| < 1$). The distributions generated by the three models are fundamentally different. Except for events with $p_{t_{\rm ljet}} \lesssim 3$ GeV, PYTHIA 6.214 - ATLAS generates greater multiplicity in the UE than the other models

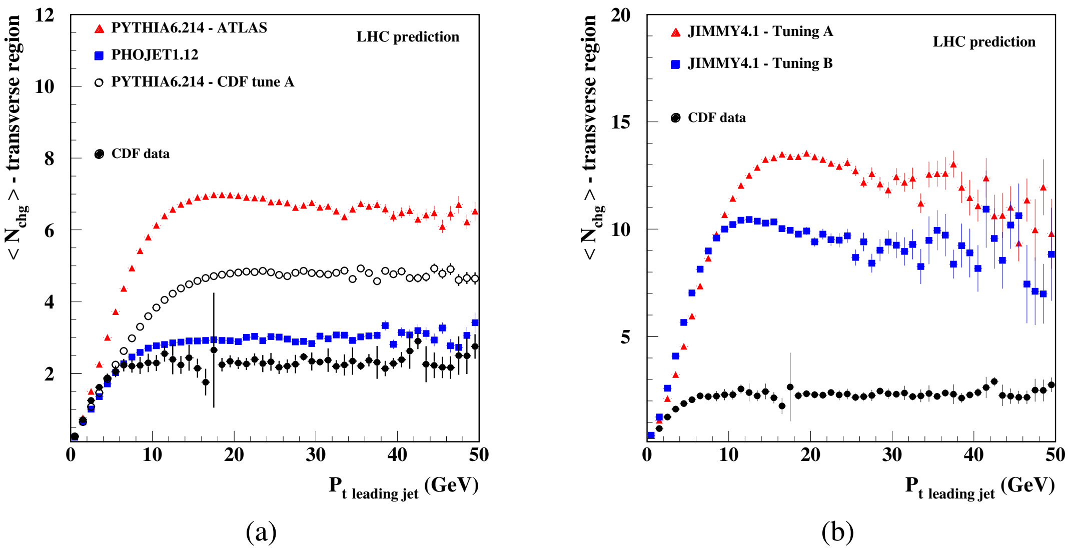


Fig. 20: (a) PYTHIA 6.214 (ATLAS and CDF tune A), PHOJET1.12 and (b) JIMMY 4.1 (tunings A and B) predictions for the average multiplicity in the UE for LHC pp collisions.

shown in Fig. 20(a).

A close inspection of predictions for the UE given in Fig. 20(a), shows that the average multiplicity in the UE for $P_{t_{\rm ljet}} > 10$ GeV reaches a plateau at ~ 6.5 charged particles according to PYTHIA 6.214 - ATLAS, ~ 5 for CDF tune A and ~ 3.0 according to PHOJET1.12. Compared to the underlying event distributions measured by CDF at 1.8 TeV, PYTHIA 6.214 - ATLAS indicates a plateau rise of $\sim 200\%$ at the LHC while PYTHIA 6.214 - CDF tune A predicts a rise of $\sim 100\%$ and PHOJET1.12 suggests a much smaller rise of $\sim 40\%$.

In Fig. 20(b) we show JIMMY 4.1 - Tuning A and B predictions for the average particle multiplicity in the UE for LHC collisions. The average multiplicity in the UE for $P_{t_{\rm ljet}} > 10$ GeV reaches a plateau at ~ 12 charged particles according to JIMMY 4.1 - Tuning A, and ~ 9.0 according to JIMMY 4.1 - Tuning B. Note that, for both JIMMY tunings, the plateau rise for the average multiplicity in the UE is much greater than the ones predicted by any of the PYTHIA tunings or by PHOJET as shown in Figs. 20(a) and (b). Once again, compared to the underlying event distributions measured by CDF at 1.8 TeV, JIMMY 4.1 - Tuning A indicates a five-fold plateau rise at the LHC while JIMMY 4.1 - Tuning B - CDF suggests a four-fold rise.

5.7 Conclusion

The minimum bias and underlying event predictions for the LHC generated by models which have been tuned to the available data have been compared. In previous studies, these models have been shown to be able to describe the data distributions for these two classes of interactions. However, in this article, it has been shown that for the models detailed in tables 1 and 2, there can be dramatic disagreements in their predictions at LHC energies. This is especially evident in the distributions for the average multiplicity in the UE (Fig. 20) where, for example, PHOJET1.12 predicts that the distribution's plateau will be at ~ 3 charged particles while JIMMY 4.1 - Tuning A predicts for the same distribution, a plateau at ~ 12.

Even though models tuned to the data have been used in this study, uncertainties in LHC predictions for minimum bias and the underlying event are still considerable. Improved models for the soft component of hadronic collisions are needed as well as more experimental information which may be used to tune current models. Future studies should focus on tuning the energy dependence for the event activity in both minimum bias and the underlying event, which at the moment seems to be one of the least understood aspects of all the models investigated in this study.

6 Can the final state at LHC be determined from ep data at HERA?

6.1 Jets and $E_\perp$-flow

A phenomenological fit for a soft-cutoff, $\hat{p}_T^{\rm min}$, and an extrapolation to LHC energies, was discussed in sections 4.1 and 5.2. However, in the $k_\perp$-factorization formalism the soft divergence is avoided, and it is possible to predict minijets and $E_\perp$-flow from HERA data alone. Thus it is not necessary to rely on a purely phenomenological fit using $\mathrm{p\bar{p}}$ collision data. This gives a better dynamical insight, and avoids the uncertainties associated with the extrapolation to higher energies.

High $p_\perp$ jets are well described by conventional *collinear factorization*, but in this formalism the minijet cross section diverges, $\sigma_{jet} \propto 1/p_\perp^4$. This implies that the total $E_\perp$ also diverges, and therefore a cutoff $\hat{p}_T^{\rm min}$ is needed. Fits to data give $\hat{p}_T^{\rm min} \sim 2$ GeV growing with energy [8, 9]. There is no theoretical basis for the extrapolation of $\hat{p}_T^{\rm min}$ from the Tevatron to LHC, which induces an element of uncertainty in the predictions for LHC.

In the *$k_\perp$-factorization* formalism the off shell matrix element for the hard subcollision $k_1 + k_2 \to q_1 + q_2$ does not blow up, when the momentum exchange $k_\perp^2$ is smaller than the incoming virtualities $k_{\perp 1}^2$ and $k_{\perp 2}^2$. The unintegrated structure functions $\mathcal{F}(x, k_\perp^2, Q^2)$ are also suppressed for small $k_\perp$, and as a result the total $E_\perp$ is not divergent but stays finite. An "effective cutoff" increases with energy, but the increase is less steep for larger energies [40].

At high energy $\sigma_{\rm jet}$ is larger than $\sigma_{\rm tot}$, which implies that there usually are *multiple hard subcollisions* in a single event. The experimental evidence for multiple collisions has been discussed in previous sections. It includes multijet events, forward-backward correlations, the pedestal effect, and associated particles in jet events. The data also indicate that the hard subcollisions are not independent. Central collisions contain more, and peripheral collisions fewer, minijets, and the results are well described by a double Gaussian distribution in impact parameter, as suggested in ref. [3].

At high energies the pdfs needed to calculate the minijet cross section have to be evaluated in the BFKL domain of small x and low $k_\perp$. This implies that non-$k_\perp$-ordered parton chains are important. For a γ^*p collision a single local $k_\perp$-maximum corresponds to a resolved photon interaction. Similarly several local maxima in a single chain correspond to correlated hard subcollisions.

In the BFKL formalism the gluon links in the t-channel correspond to reggeized gluons, which means that soft emissions are compensated by virtual corrections. These soft emissions *do not* contribute to the parton distributions or total cross sections, but they *do* contribute to the properties of final states, and should then be added with Sudakov form factors. The CCFM model [41, 42] interpolates between DGLAP and BFKL. Here some soft emissions are included in the initial state radiation, which implies that they must be suppressed by non-eikonal form factors. The Linked Dipole Chain (LDC) model [43] is a reformulation and generalization of CCFM, in which more emissions are treated as final state emissions, in closer agreement with the BFKL picture. In the LDC formalism the chain formed by the initial state radiation is *fully symmetric* with respect to the photon end and the proton end of the ladder. This symmetry implies that the formalism is also directly applicable to hadron-hadron collisions. Thus a fit to DIS data will also give the cross section for a *parton chain* in pp collisions [44].

A potential problem is due to the fact that with a running α_s, the enhancement of small $k_\perp$ implies that the result depends on a necessary cutoff Q_0. Good fits to DIS data are possible with different Q_0, if the input distribution $f_0(x, Q_0^2)$ is adjusted accordingly. However, although a larger cutoff gives fewer hard chains, it also implies a larger number of soft chains, in which no link has a $k_\perp$ larger than Q_0. Thus the total number of chains in pp scattering is independent of Q_0, and therefore well determined by the fit to DIS data.

When the fit to HERA data in this way is applied to $\mathrm{p\bar{p}}$ scattering at the Tevatron, the predictions for *e.g.* jet multiplicity and the pedestal effect are very close to CDF's tune A, described in section 3. The

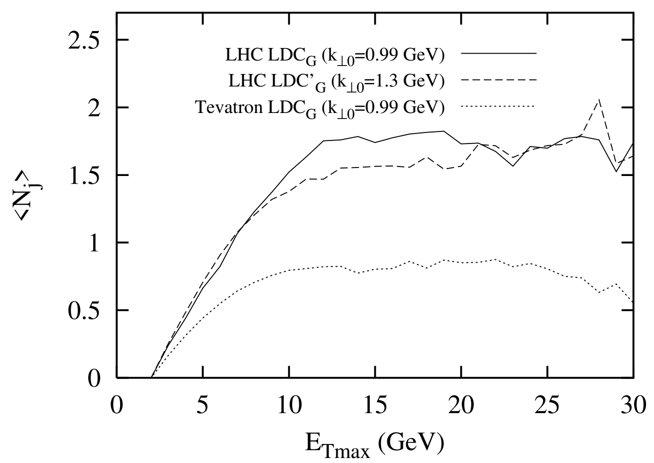


Fig. 21: The average number of minijets per event in the "minimum azimuth region", as a function of transverse energy of the trigger jet, $E_{\perp max}$. The figure shows the result for 1.8 TeV and for LHC. The two LHC curves correspond to different values for Q_0, showing the stability with respect to the soft cutoff.

result is insensitive to the soft cutoff Q_0, which implies that the extrapolation to LHC energies is stable, and does not depend on an uncertain extrapolation of the low-$p_\perp$ cutoff needed in a collinear formalism. As an example fig. 21 shows a prediction for the average number of minijets per event within 60° in azimuth perpendicular to a trigger jet, on the side with minimum activity.

As the LDC model is fully symmetric with respect to an interchange of the projectile and the target, the parton chains have to combine at one end at the same rate as they multiply at the other. Therefore the formalism should be suitable for studies of gluon recombination and *saturation*. This work is in progress, and some preliminary results from combining the LDC model with Mueller's dipole formulation in transverse coordinate space [45, 46, 47] are presented in ref. [48].

6.2 Hadron multiplicities

The hadron multiplicity is much more sensitive to non-perturbative effects. This implies larger uncertainties, and models differ by factors 3-4 in their predictions for LHC (see section 5). The CDF data also show that the data are best fitted if colours rearrange so that secondary hard scatterings give minimum extra string length, i.e. minimum extra multiplicity. This is very different from the case in e^+e^- annihilation.

In pp collisions the multiplicity of final state hadrons depends very sensitively on the colour connections between the produced partons. This implies that the result depends on soft non-perturbative effects. Multiple interactions are related to multiple pomeron exchange, which is expected to obey the Abramovskyĭ-Gribov-Kancheli cutting rules [49]. These rules are derived for a multiperipheral model, but a multiperipheral chain has important similarities with a gluonic chain. An essential feature is the dominance of small momentum exchanges at each vertex. The colour structure of QCD gives, however, some extra complications as discussed by J. Bartels (see the contribution by Bartels to working group 4).

The pomeron is identified by two gluon exchange, and multiple chains correspond to multi-pomeron exchange. For the example of two pomeron exchange, the AGK rules give the relative weights $1 : -4 : 2$ for cutting 0, 1 or 2 pomerons. These ratios imply that the two-pomeron diagram contributes to the multiplicity *fluctuations*, but has no effect on the number of produced particles, determined by $\sum n\sigma_n$. This result can also be generalized to the exchange of more pomerons.

Similar cutting rules apply to a diagram with two pomerons attached to one proton and one pomeron to the other, connected by a central triple-pomeron coupling. In ref. [49] this and similar diagrams are, however, expected to give smaller contributions.

A hard gg $\to$ gg subcollision will imply that the two proton remnants carry colour octet charges. This is expected to give two colour triplet strings, or two cluster chains, connecting the two remnants and the two final state gluons. In the string model the strings are stretched between the remnants, with the gluons acting as kinks on the strings. These kinks can either be on different strings or both on the first or both on the second string, with equal probabilities for the three possibilities (see ref. [50]). Including initial state radiation will give extra kinks, which due to colour coherence will be connected so as to result in minimal extra string length.

Multiple collisions with two independent gg $\to$ gg scatterings would be expected to correspond to two cut pomerons, with four triplet strings stretched between the proton remnants. This would give approximately a doubled multiplicity, in accordance with the AGK cutting rules. However, the CDF data show that this is *far from reality*.

CDF's successful tune A [6] is a fit using an early PYTHIA version. Already in the analysis in ref. [3] it was realized that four strings would give too high multiplicity. Therefore in this early PYTHIA version there are three possible string connections for a secondary hard subcollision. 1) An extra closed string loop between the two final state gluons. 2) A single string between the scattered partons, which are then treated as a $q\bar{q}$ system. 3) The new hard gluons are inserted as extra kinks among the initial state radiations, in a way which corresponds to minimum extra string length. In the successful tune A the last possibility is chosen in 90% of the cases, which corresponds to *minimal extra multiplicity*. The default PYTHIA tune, which contained equal probabilities for the three cases, does not give a good fit. A more advanced treatment of pp collisions [8, 9] is implemented in a new PYTHIA version (6.3) [2] (see section 2.1). This model does, however, not work as well as Field's tune A of the older model.

Consequently two independent hard collisions do not correspond to two cut pomeron ladders stretched between the proton remnants. It also does not correspond to a cut pomeron loop in the centre. Instead it looks like a single ladder, with a higher density of gluon rungs in the central region.

How can this be understood? It raises a set of important questions: What does it imply for the AGK rules and the diffractive gap survival probability? Do rescattering and unitarity constraints (and AGK) work in the initial perturbative phase? If so, does this correspond to an initial hard collision inside a confining bag, with the final state partons colour connected in a later non-perturbative phase?

We can compare with the situation in e^+e^--annihilation. If two gluons are emitted from the quark or antiquark legs, these gluons form a colour singlet with probability $\sim 1/N_c^2$. They could then hadronize as a separate system. Analyses of data from LEP indicate that such isolated systems are suppressed even more than by a factor $\sim 1/N_c^2$.

In conclusion we have following important questions:

- Why do the strings make the shortest connections in $\approx 100\%$ in pp and almost never in e^+e^-?
- How do multiplicity fluctuations and the relation diffraction *diffraction* and *high multiplicity events* reflect features of AGK in ep, γp, and pp?
- Do unitarity effects and AGK cutting rules work as expected in an initial perturbative phase, and the colours recombine in a subsequent nonperturbative soft phase?
- Or is the pomeron a much more complicated phenomenon than the simple ladder envisaged by Abramovskyĭ-Gribov-Kancheli?

7 Conclusions and the potential for HERA data

This was a very active area of discussion during the workshop. In fact, the area remains so active that firm conclusions are hard to make, and likely to be superceded on a very short timescale. Nevertheless there are some things which do seem clear.

- The underlying event is clearly an topic of substantial importance for the LHC.

- The dominant input data for understanding the underlying event comes at present from the Tevatron, with HERA data primarily featuring indirectly, though importantly, via the parton densities.
- The data strongly indicate that multiple hard scatters are required to adequately describe the final state in high energy hadron collisions.
- The UE depends on the measurement being made as demonstrated by difference between the UE in the CDF leading jet and back-to-back jet analysis.
- The colour structure of the final parton state is an unsolved problem. The CDF data indicate that 'short strings' are strongly favoured.
- There are large uncertainties associated with extrapolating the available models to LHC energies.

As far as the future impact of HERA data on this area goes, some ideas have been discussed in the previous section. In addition, it is worth noting that most of the models discussed here have also been used in high energy photoproduction at HERA [51], where they also improve the description of the data. No study comparable to those carried out at pp or $\mathrm{p\bar{p}}$ experiments is currently available. The benefits of such a study would be that (a) HERA could add another series of points in energy (around 200 GeV) to help pin down the energy dependence of the underlying event, (b) it is possible to select regions of phase space where resolved (*i.e.*, hadronic) or direct (*i.e.*, pointlike) photons dominate, thus effectively switching on or off the photon PDF (and thus presumably multiparton interactions) and allowing comparison between the two cases, (c) the photon is a new particle with which the physics assumptions of underlying event models can be confronted. The last of these points however also implies that a slew of new parameters will be introduced, and one may learn more about the photon this way than about underlying events themselves. Either way, it is to be hoped that such a study will be carried out before HERA finishes and LHC switches on.

References

[1] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001). `hep-ph/0010017`.

[2] T. Sjöstrand *et al.*, LU TP 03-38 (2003). `hep-ph/0308153`.

[3] T. Sjöstrand and M. van Zijl, Phys. Rev. **D36**, 2019 (1987).

[4] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C21**, 443 (2001);
H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C21**, 33 (2001).

[5] J. Dischler and T. Sjöstrand, Eur. Phys. J. direct **C3**, 2 (2001). `hep-ph/0011282`.

[6] R. Field, *Min-bias and the underlying event at the tevatron and the lhc*.
http://www.phys.ufl.edu/~rfield/cdf/tunes/py_tuneA.html.

[7] T. Sjöstrand and P. Z. Skands, Nucl. Phys. **B659**, 243 (2003). `hep-ph/0212264`.

[8] T. Sjöstrand and P. Z. Skands, JHEP **03**, 053 (2004). `hep-ph/0402078`.

[9] T. Sjöstrand and P. Z. Skands, Eur. Phys. J. **C39**, 129 (2005). `hep-ph/0408302`.

[10] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, Z. Phys. C **72**, 637 (1996).
(Hep-ph/9601371).

[11] J. M. Butterworth and M. H.Seymour, In preparation see
`http://www.cedar.ac.uk/hepforge/jimmy`.

[12] S. Höche and F. Krauss, in preparation.

[13] R. Kuhn, F. Krauss, B. Ivanyi, and G. Soff, Comput. Phys. Commun. **134**, 223 (2001). `hep-ph/0004270`.

[14] F. Krauss, A. Schälicke, and G. Soff, hep-ph/0503087.

[15] S. Catani *et al.*, JHEP **11**, 063 (2001). `hep-ph/0109231`.

[16] F. Krauss, JHEP **08**, 015 (2002). `hep-ph/0205283`.

[17] F. Krauss *et al.*, Phys. Rev. **D70**, 114009 (2004). `hep-ph/0409106`.

[18] A. Schälicke and F. Krauss, JHEP **07**, 018 (2005). `hep-ph/0503281`.

[19] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. **D65**, 092002 (2002).

[20] R. D. Field, http://www.phys.ufl.edu/~rfield/cdf/chgjet/chgjet_intro.html.

[21] A. Capella *et al.*, Phys. Rep. **236**, 225 (1994).

[22] R. Engel, *PHOJET manual (program version 1.05)*, June 1996.

[23] R. Field, Int. J. Mod. Phys. **A16S1A**, 250 (2001).

[24] J. Huston, Int. J. Mod. Phys. **A16S1A**, 219 (2001).

[25] G. Marchesini and B. R. Weber, Phys. Rev **D38**, 3419 (1988).

[26] J. Pumplin, Phys. Rev. **D57**, 5787 (1998).

[27] D. Acosta *et al.*, Phys. Rev. **D70**, 072002 (2004).

[28] R. Field, Acta Phys. Polon. **B36**, 167 (2005).

[29] LHCb Collaboration. CERN-LHCC-2003-030.

[30] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. **D64**, 072002 (2001).

[31] T. Sjöstrand, L. Lönnblad, and S. Mrenna, *Pythia 6.2: Physics and manual.* Preprint hep-ph/0108264, 2001.

[32] Heavy Flavor Averaging Group (HFAG) (2005). `hep-ex/0505100`.

[33] LHCb Collaboration, P. Bartalini *et al.* (1999).

[34] UA5 Collaboration, G. J. Alner *et al.*, Z. Phys. **C33**, 1 (1986).

[35] CDF Collaboration, F. Abe *et al.*, Phys. Rev. **D41**, 2330 (1990).

[36] R. Engel, *Hadronic Interactions of Photons at High Energies*, 1997. Siegen (PhD Thesis).

[37] A. Moraes, C. Buttar, and I. Dawson, p. 46 (2005). ATL-PHYS-PUB-2005-007.

[38] J. M. Butterworth and S. Butterworth, Comput. Phys. Commun. **153**, 164 (2003). See results available at http://www.cedar.ac.uk/jetweb, `hep-ph/0210404`.

[39] F. Abe *et al.*, Phys. Rev. **D41**, 2330 (1990).

[40] G. Gustafson and G. Miu, Phys. Rev. **D63**, 034004 (2001). `hep-ph/0002278`.

[41] S. Catani, F. Fiorani, and G. Marchesini, Phys. Lett. **B234**, 339 (1990).

[42] M. Ciafaloni, Nucl. Phys. **B296**, 49 (1988).

[43] B. Andersson, G. Gustafson, and J. Samuelsson, Nucl. Phys. **B467**, 443 (1996).

[44] G. Gustafson, L. Lönnblad, and G. Miu, Phys. Rev. **D67**, 034020 (2003). `hep-ph/0209186`.

[45] A. H. Mueller, Nucl. Phys. **B415**, 373 (1994).

[46] A. H. Mueller and B. Patel, Nucl. Phys. **B425**, 471 (1994). `hep-ph/9403256`.

[47] A. H. Mueller, Nucl. Phys. **B437**, 107 (1995). `hep-ph/9408245`.

[48] E. Avsar, G. Gustafson, and L. Lönnblad, JHEP **07**, 062 (2005). `hep-ph/0503181`.

[49] V. A. Abramovsky, V. N. Gribov, and O. V. Kancheli, Yad. Fiz. **18**, 595 (1973).

[50] G. Gustafson, Z. Phys. **C15**, 155 (1982).

[51] J. M. Butterworth and M. Wing, Rep. Prog. Phys. **68**, 2773 (2005). `hep-ex/0509018`.

Forward Jets and Multiple Interactions

Jacek Turnau[1], *Leif Lönnblad*[2]
[1]Institute of Nuclear Physics, Polish Academy of Sciences, Cracow, Poland;
[2]Department of Theoretical Physics, Lund University, Sweden.

Abstract
HERA provides a unique possibility to investigate the dependence of multiple interactions on transverse interaction sizes through variation of the photon virtuality Q^2. In order to observe effects of multiple interactions at Q^2 substantially different from zero we have to look into regions of phase space where resolved processes dominate over direct ones. The forward jet production at small values of Bjorken x is one example. PYTHIA and RAPGAP have been employed to estimate contribution of the multiple interactions to forward jet production cross section.

Comparisons of HERA photoproduction data with QCD NLO calculations for high transverse momentum jets revealed that the observed jets are not well described by the calculations. The energy flow adjacent to jets - the underlying event or jet pedestal - was found to be far above QCD expectations [1]. Similar excess of underlying energy was observed in $p\bar{p}$ data, see [2] and [3] for recent studies. It appears that both HERA and TEVATRON data can be described by adding beam remnant interactions, from soft to hard, as first proposed in ref. [4]. The remnant beam-beam interactions can result in multiple hard parton interactions (MI) thus creating additional pairs of jets. Therfore the presence of four high transverse momentum objects in the hadronic final state (e.g. four jets or prompt photon and three jets) allows searches for signatures of multi-parton interactions in a region of phase space where their effects may be maximized. The evidence of MI coming from 4-jet studies is more explicit and is not complicated by initial/final state radiation and soft beam-remnant components of the underlying event. Both ZEUS [5] and CDF [6] observed explicite double parton interactions in rough agreement with PYTHIA [7, 4] simulations.

The very interesting aspect of measurements at HERA is that variation of the photon virtuality Q^2 provides information about transverse interaction sizes. Observation of the dependence of MI on Q^2 could be important from the phenomenological point of view. In order to see MI at photon virtuality substantially different from zero we have to look into regions of phase space for deep inelastic scattering where the resolved virtual photon processes dominate over direct ones. The forward jet production at small values of Bjorken x is one example. Here one could expect that additional interactions between the remnants of the proton and resolved virtual photon would produce extra hadron multiplicity in an underlying event. Although the transverse momentum of these hadrons would be limited, they could still give a substantial effect on the rate of forward jets which have a steeply falling $p_\perp$ spectrum.

The forward jet cross-section is especially interesting since it has been notoriously difficult to reproduce by standard DGLAP-based parton shower event generators. It has been shown that the description of the forward jet cross section can be improved by adding resolved virtual photon component in eg. the RAPGAP Monte Carlo [8], but the jet rates produced in the simulations are still a bit too low in the small-x region. In order to check if MI can give measurable contribution to this process we have performed a study in which we estimate MI effect using both PYTHIA 6.2 and RAPGAP 3.1. We use PYTHIA since the MI model there has been shown to be able to give a good description of underlying events and jet pedestal effects in hadron-hadron collisions and in photoproduction, and it is fairly easy to apply the same model to the resolved part of the $\gamma^\star - p$ collisions. However, PYTHIA does not describe correctly the transverse energy flow in in DIS at HERA above $Q^2 \approx 5$ GeV^2. We can still use PYTHIA

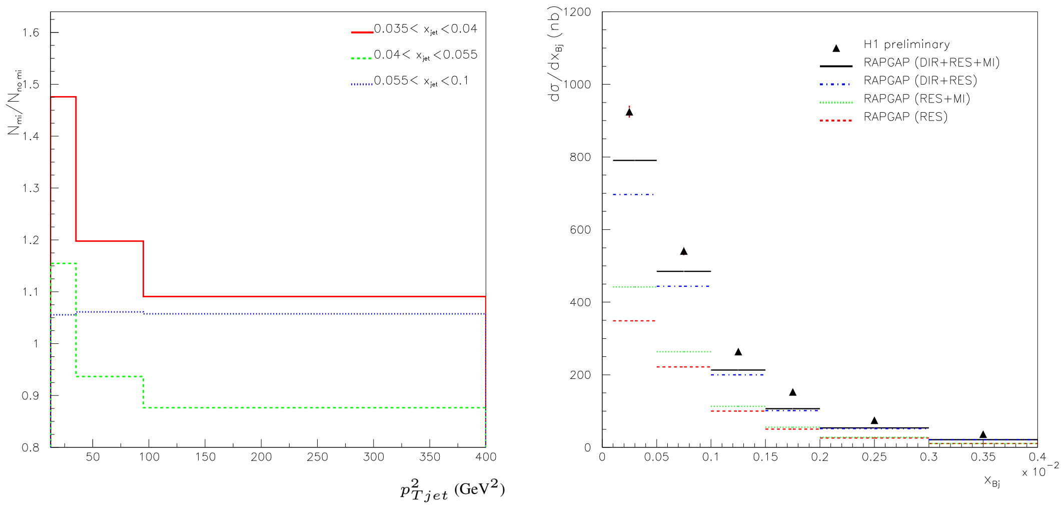


Fig. 1: **Left:** Ratio of forward jets with and without multiple interactions as a function of jet transverse momentum squared for three regions of proton momentum fraction carried by jet **Right:** The H1 forward jet cross section data compared with RAPGAP 3.1 simulation. Multiple interactions are included as x, Q^2, x_{jet} and p^2_{Tjet} dependent weights to resolved component, calculated using PYTHIA 6.2

to estimate the relative effect of MI and we have generated forward jet cross section with H1 cuts [9]: ($p_{Tjet} > 3.5$ GeV,$x_{jet} > 0.035$, $20° > \Theta_{jet} > 7°$ and $0.5 < p^2_{Tjet}/Q^2 < 5$) using PYTHIA 6.2 with default settings in γp mode (MI in mode 2) with $\gamma^\star$ momentum corresponding to several values of x and Q^2 within DIS kinematical phase space $0.0001 < x < 0.004$ and $5 < Q^2 < 85$ GeV^2.

In Fig. 1 (left) we show example of the ratio of number of the forward jets with and without MI, here for $x = 0.0004$ and $Q^2 = 8$ GeV^2, as a function of p^2_{Tjet}. It can be seen that effect of MI is quite substantial in the lowest p^2_{Tjet} bin. Treating the above mentioned ratios as weights depending on x, Q^2, x_{jet} and p^2_{Tjet}, we have generated inclusive forward jet cross section using RAPGAP 3.1 within above mentioned H1 cuts. The Fig. 1 (right) shows the result of this calculation. The inclusive forward jet cross section is enhanced by MI for about 15% in the lowest x bin, in fact improving description of the data. The effect of MI diminishes quickly with increasing x as result of decreasing contribution of the resolved photon component.

This very preliminary study suggests that Q^2 dependence of multiple interactions can be studied at HERA. This will require large statistics and an improved understanding of the underlying QCD evolution in forward jet production.

References

[1] H1 Collaboration, I. Abt *et al.*, Phys. Lett. **B314**, 436 (1993);
ZEUS Collaboration, M.Derrick *et al.*, Phys. Lett. **B342**, 417 (1995);
H1 Collaboration, T. Ahmed *et al.*, Nucl. Phys. **B445**, 195 (1995);
G. Knies, J. Phys. G. Nucl. Part. Phys. **G19**, 1523 (1993).

[2] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. **D65**, 092002 (2002).

[3] CDF Collaboration, R. Field, Acta Physica Polonica **B36**, 167 (2005).

[4] T. Sjostrand and M. van Zijl, Phys. Rev. **D36**, 2019 (1987).

[5] ZEUS Collaboration, C. Gwenlen, Acta Physica Polonica **B33**, 3123 (2002).

[6] CDF Collaboration, F. Abe *et al.*, Phys. Rev. **D56**, 3811 (1997).

[7] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna, and E. Norrbin, Comput. Phys. Commun. **135**, 238 (2001). `hep-ph/0010017`.

[8] ZEUS Collaboration, J. Breitweg *et al.*, Phys. Lett. **B479**, 37 (2000);
H1 Collaboration, C. Adloff *et al.*, EPJ **C13**, 397 (2000);
L. J. H. Jung and H. Kuester, EPJ **C9**, 383 (1999).

[9] H1 Collaboration, A. Aktas *et al.*, submitted to EPJ (2005). `hep-ex/0508055`.

Survival Probability of Large Rapidity Gaps

E. Gotsman, E. Levin, U. Maor, E. Naftali, A. Prygarin
HEP Department, School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Science, Tel Aviv University, Tel Aviv, 69978, ISRAEL.

Abstract
Our presentation centers on the consequences of s-channel unitarity, manifested by soft re-scatterings of the spectator partons in a high energy diffractive process, focusing on the calculations of gap survival probabilities. Our emphasis is on recent estimates relevant to exclusive diffractive Higgs production at the LHC. To this end, we critically re-examine the comparison of the theoretical estimates of large rapidity gap hard di-jets with the measured data, and remark on the difficulties in the interpretation of HERA hard di-jet photoproduction.

1 Introduction

A large rapidity gap (LRG) in an hadronic, photo or DIS induced final state is experimentally defined as a large gap in the $\eta - \phi$ lego plot devoid of produced hadrons. LRG events were suggested [1, 2, 3, 4] as a signature for Higgs production due to a virtual $W - W$ fusion subprocess. An analogous pQCD process, in which a colorless exchange ("hard Pomeron") replaces the virtual W, has a considerably larger discovery potential as it leads also to an exclusive $p + H + p$ final state. Assuming the Higgs mass to be in the range of $100 - 150\,GeV$, the calculated rates for this channel, utilizing proton tagging are promissing. Indeed, LRG hard di-jets, produced via the same production mechanism, have been observed in the Tevatron [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17] and HERA [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29]. The experimental LRG di-jets production rates are much smaller than the pQCD (or Regge) estimates. Following Bjorken [3, 4], the correcting damping factor is called "LRG survival probability".

The present summary aims to review and check calculations of the survival probability as applied to the HERA-Tevatron data and explore the consequences for diffractive LRG channels at LHC with a focus on diffractive Higgs production.

We distinguish between three configurations of di-jets (for details see Ref. [13, 14, 15, 16, 17]):
1) A LRG separates the di-jets system from the other non diffractive final state particles. On the partonic level this is a single diffraction (SD) Pomeron exchange process denoted GJJ.
2) A LRG separates between the two hard jets. This is a double diffraction (DD) denoted JGJ.
3) Centrally produced di-jets are separated by a LRG on each side of the system. This is a central diffraction (CD) two Pomeron exchange process denoted GJJG. This mechanism also leads to diffractive exclusive Higgs production.

We denote the theoretically calculated rate of a LRG channel by F_{gap}. It was noted by Bjorken [3, 4] that we have to distinguish between the theoretically calculated rate and the actual measured rate f_{gap}

$$f_{gap} = \langle \mid S \mid^2 \rangle \cdot F_{gap}. \tag{1}$$

The proportionality damping factor [30, 31, 32, 33] is the survival probability of a LRG. It is the probability of a given LRG not to be filled by debris (partons and/or hadrons). These debris originate from the soft re-scattering of the spectator partons resulting in a survival probability denoted $\mid S_{spec}(s) \mid^2$, and/or from the gluon radiation emitted by partons taking part in the hard interaction with a corresponding survival probability denoted $\mid S_{brem}(\Delta y) \mid^2$,

$$\langle \mid S(s, \Delta y) \mid^2 \rangle = \langle \mid S_{spec}(s) \mid^2 \rangle \cdot \langle \mid S_{brem}(\Delta y) \mid^2 \rangle. \tag{2}$$

s is the c.m. energy square of the colliding particles and Δy is the large rapidity gap. Gluon radiation from the interacting partons is strongly suppressed by the Sudakov factor [34]. However, since this suppression is included in the perturbative calculation (see **4.3**) we can neglect $\langle \mid S_{brem}(\Delta y) \mid^2 \rangle$ in our calculations. In the following we denote $\langle \mid S_{spec} \mid^2 \rangle = S^2$. It is best defined in impact parameter space (see **2.1**)). Following Bjorken [3, 4], the survival probability is determined as the normalized integrated product of two quantities

$$S^2 = \frac{\int d^2b \mid M^H(s,b) \mid^2 P^S(s,b)}{\int d^2b \mid M^H(s,b) \mid^2}. \tag{3}$$

$M^H(s,b)$ is the amplitude for the LRG diffractive process (soft or hard) of interest. $P^S(s,b)$ is the probability that no inelastic soft interaction in the re-scattering eikonal chain results in inelasticity of the final state at (s,b).

The organization of this paper is as follows: In Sec.2 we briefly review the role of s-channel unitarity in high energy soft scattering and the eikonal model. The GLM model [30, 31, 32, 33] and its consequent survival probabilities [35, 36, 37] are presented in Sec.3, including a generalization to a multi channel re-scattering model [38, 39]. The KKMR model [40, 41, 42, 43, 44] and its survival probabilities is presented in Sec.4. A discussion and our conclusions are presented in Sec.5. An added short presentation on Monte Carlo calculations of S^2 is given in an Appendix.

2 Unitarity

Even though soft high energy scattering has been extensively studied experimentally over the last 50 years, we do not have, as yet, a satisfactory QCD framework to calculate even the gross features of this impressive data base. This is just a reflection of our inability to execute QCD calculations in the non-perturbative regime. High energy soft scattering is, thus, commonly described by the Regge-pole model [45, 46]. The theory, motivated by S matrix approach, was introduced more than 40 years ago and was soon after followed by a very rich phenomenology.

The key ingredient of the Regge pole model is the leading Pomeron, whose linear t-dependent trajectory is given by

$$\alpha_{I\!P}(t) = \alpha_{I\!P}(0) + \alpha'_{I\!P} t. \tag{4}$$

A knowledge of $\alpha_{I\!P}(t)$ enables a calculation of σ_{tot}, σ_{el} and $\frac{d\sigma_{el}}{dt}$, whose forward elastic exponential slope is given by

$$B_{el} = 2B_0 + 2\alpha'_{I\!P} ln \left(\frac{s}{s_0}\right). \tag{5}$$

Donnachie and Landshoff (DL) have vigorously promoted [47, 48] an appealing and very simple Regge parametrization for total and forward differential elastic hadron-hadron cross sections in which they offer a global fit to all available hadron-hadron and photon-hadron total and elastic cross section data. This data, above $P_L = 10\,GeV$, is excellently fitted with universal parameters. We shall be interested only in the DL Pomeron with an intercept $\alpha_{I\!P}(0) = 1 + \epsilon$, where $\epsilon = 0.0808$, which accounts for the high energy growing cross sections. Its fitted [49] slope value is $\alpha'_{I\!P} = 0.25\,GeV^{-2}$.

2.1 S-channel unitarity

The simple DL parametrization is bound to violate s-channel unitarity at some energy since σ_{el} grows with energy as $s^{2\epsilon}$, modulu logarithmic corrections, while σ_{tot} grows only as s^{ϵ}. The theoretical problems at stake are easily identified in an impact b-space representation.

The elastic scattering amplitude is normalized so that

$$\frac{d\sigma_{el}}{dt} = \pi \mid f_{el}(s,t) \mid^2, \tag{6}$$

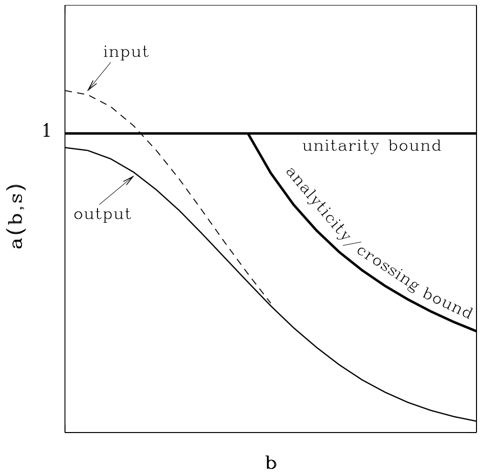


Fig. 1: *A pictorial illustration of a high energy b-space elastic amplitude bounded by unitarity and analyticity/crossing. In the illustration we have an input amplitude which violates the eikonal unitarity bound and an output amplitude obtained after a unitarization procedure.*

$$\sigma_{tot} \;=\; 4\pi Im f_{el}(s,0). \tag{7}$$

The elastic amplitude in b-space is defined as

$$a_{el}(s,b) \;=\; \frac{1}{2\pi}\int d\mathbf{q}e^{-i\mathbf{q}\cdot\mathbf{b}} f_{el}(s,t), \tag{8}$$

where $t \;=\; -\mathbf{q}^2$. In this representation

$$\sigma_{tot} \;=\; 2\int d^2b\, Im[a_{el}(s,b)], \tag{9}$$

$$\sigma_{el} \;=\; \int d^2b \mid a_{el}(s,b) \mid^2, \tag{10}$$

$$\sigma_{in} \;=\; \sigma_{tot} \;-\; \sigma_{el}. \tag{11}$$

As noted, a simple Regge pole with $\alpha_{I\!P}(0) \;>\; 1$ will eventually violate s-channel unitarity. The question is if this is a future problem to be confronted only at far higher energies than presently available, or is it a phenomena which can be identified through experimental signatures observed within the available high energy data base. It is an easy exercise to check that the DL model [47, 48], with its fitted global parameters, will violate the unitarity black bound (see **2.2**) at very small b, just above the present Tevatron energy. Indeed, CDF reports [50] that $a_{el}(b=0,\sqrt{s}=1800) = 0.96 \pm 0.04$. A pictorial illustration of the above is presented in Fig.1. Note that the energy dependence of the experimental SD cross section [13, 14, 15, 16, 17] in the ISR-Tevatron energy range is much weaker than the power dependences observed for σ_{el}. Diffractive cross sections are not discussed in the DL model.

2.2 The eikonal model

The theoretical difficulties, pointed out in the previous subsection, are eliminated once we take into account the corrections necessitated by unitarity. The problem is that enforcing unitarity is a model dependent procedure. In the following we shall confine ourselves to a Glauber type eikonal model

[51]. In this approximation, the scattering matrix is diagonal and only repeated elastic re-scatterings are summed. Accordingly, we write

$$a_{el}(s,b) \;=\; i\left(1 - e^{-\Omega(s,b)/2}\right). \tag{12}$$

Since the scattering matrix is diagonal, the unitarity constraint is written as

$$2Im[a_{el}(s,b)] \;=\; \mid a_{el}(s,b) \mid^2 \;+\; G^{in}(s,b), \tag{13}$$

with

$$G^{in} \;=\; 1 \;-\; e^{-\Omega(s,b)}. \tag{14}$$

The eikonal expressions for the soft cross sections of interest are

$$\sigma_{tot} \;=\; 2\int d^2b\left(1 \;-\; e^{-\Omega(s,b)/2}\right), \tag{15}$$

$$\sigma_{el} \;=\; \int d^2b\left(1 \;-\; e^{-\Omega(s,b)/2}\right)^2, \tag{16}$$

$$\sigma_{in} \;=\; \int d^2b\left(1 \;-\; e^{-\Omega(s,b)}\right), \tag{17}$$

and

$$B_{el}(s) \;=\; \frac{\int \, d^2b \; b^2 \; \left(1 \;-\; e^{-\Omega(s,b)/2}\right)}{2\,\int \, d^2b\; \left(1 \;-\; e^{-\Omega(s,b)/2}\right)}. \tag{18}$$

From Eq.(14) it follows that $P^S(s,b) \;=\; e^{-\Omega(s,b)}$ is the probability that the final state of the two initial interacting hadrons is elastic, regardless of the eikonal rescattering chain. It is identified, thus, with $P^S(s,b)$ of Eq.(3).

Following our implicit assumption that, in the high energy limit, hadrons are correct degrees of freedom, i.e. they diagonalize the interaction matrix, Eq.(12) is a general solution of Eq.(13) as long as the input opacity Ω is arbitrary. In the eikonal model Ω is real and equals the imaginary part of the iterated input Born amplitude. The eikonalized amplitude is imaginary. Its analyticity and crossing symmetry are easily restored. In a Regge language we substitute, to this end, $s^{\alpha_{I\!P}} \;\rightarrow\; s^{\alpha_{I\!P}} e^{-\frac{1}{2}i\pi\alpha_{I\!P}}$.

In the general case, Eq.(13) implies a general bound, $\mid a_{el}(s,b) \mid \leq 2$, obtained when $G^{in} \;=\; 0$. This is an extreme option in which asymptotically $\sigma_{tot} \;=\; \sigma_{el}$ [52]. This is formally acceptable but not very appealing. Assuming that a_{el} is imaginary, we obtain that the unitarity bound coincides with the black disc bound, $\mid a_{el}(s,b) \mid \leq 1$. Accordingly,

$$\frac{\sigma_{el}}{\sigma_{tot}} \leq \frac{1}{2}. \tag{19}$$

3 The GLM Model

The GLM screening correction (SC) model [30, 31, 32, 33] is an eikonal model originally conceived so as to explain the exceptionally mild energy dependence of soft diffractive cross sections. It utilized the observation that s-channel unitarization enforced by the eikonal model operates on a diffractive amplitude in a different way than it does on the elastic amplitude. The GLM diffractive damping factor is identical to Bjorken's survival probability.

3.1 The GLM SC model

In the GLM model, we take a DL type Pomeron exchange amplitude input in which $\alpha_{I\!P}(0) = 1+\Delta > 0$. The simplicity of the GLM SC model derives from the observation that the eikonal approximation with a central Gaussian input, corresponding to an exponential slope of $\frac{d\sigma_{el}}{dt}$, can be summed analytically. This is, clearly, an over simplification, but it reproduces the bulk of the data well, i.e. the total and the forward elastic cross sections. Accordingly, the eikonal DL type b-space expression for $\Omega(s,b)$ is:

$$\Omega(s,b) \,=\, \nu(s)\,\Gamma^S(s,b), \tag{20}$$

where,

$$\nu(s) \,=\, \sigma(s_0)\left(\frac{s}{s_0}\right)^{\Delta}, \tag{21}$$

$$R^2(s) \,=\, 4R_0^2 \,+\, 4\alpha'_{I\!P} ln(\frac{s}{s_0}), \tag{22}$$

and the soft profile is defined

$$\Gamma^S(s,b) \,=\, \frac{1}{\pi R^2(s)}\, e^{-\frac{b^2}{R^2(s)}}. \tag{23}$$

It is defined so as to keep the normalization $\int d^2b\,\Gamma^S(s,b) \,=\, 1$.

One has to distinguish between the eikonal model input and output. The key element is that the power Δ, and ν, are input information, not bounded by unitarity, and should not be confused with DL effective power ϵ and the corresponding total cross section. Since the DL model reproduces the forward elastic amplitude, in the ISR-HERA-Tevatron range, well, we require that the eikonal model output will be compatible with the DL results. Obviously, $\Delta > \epsilon$. In a non screened DL type model with a Gaussian profile the relation $B_{el} = \frac{1}{2}R^2(s)$ is exact. In a screened model, like GLM, $B_{el} > \frac{1}{2}R^2(s)$ due to screening.

With this input we get

$$\sigma_{tot} \,=\, 2\pi R^2(s)\left[ln\left(\frac{\nu(s)}{2}\right) + C - Ei\left(-\frac{\nu(s)}{2}\right)\right] \,\propto\, ln^2(s), \tag{24}$$

$$\sigma_{el} \,=\, \pi R^2(s)\left[ln\left(\frac{\nu(s)}{4}\right) + C - 2Ei\left(-\frac{\nu(s)}{2}\right) + Ei\left(-\nu(s)\right)\right] \,\propto\, \frac{1}{2}ln^2(s), \tag{25}$$

$$\sigma_{in} \,=\, \pi R^2(s)\{ln[\nu(s)] + C - Ei[-\nu(s)]\} \,\propto\, \frac{1}{2}ln^2(s). \tag{26}$$

$Ei(x) = \int_{-\infty}^{x} \frac{e^t}{t}dt$, and $C = 0.5773$ is the Euler constant. An important consequence of the above is that the ratio $\frac{\sigma_{el}}{\sigma_{tot}}$ is a single variable function of $\nu(s)$. In practice it means that given the experimental value of this ratio at a given energy we can obtain an "experimental" value of ν which does not depend on the adjustment of free parameters.

The formalism presented above is extended to diffractive channels through the observation, traced to Eqs.(3) and (14), that $P^S(s,b) = e^{-\Omega(s,b)}$. Accordingly, a screened non elastic diffractive cross section is obtained by convoluting its b-space amplitude square with the probability P^S.

The above has been utilized [30, 31, 32, 33] to calculate the soft integrated single diffraction cross section. To this end, we write, in the triple Regge approximation [53], the double differential cross section $\frac{M^2 d\sigma_{sd}}{dM^2 dt}$, where M is the diffracted mass. We, then, transform it to b-space, multiply by $P^S(s,b)$ and integrate. The output $\frac{M^2 d\sigma_{sd}}{dM^2 dt}$, changes its high energy behaviour from $s^{2\Delta}$ modulu $ln(\frac{s}{s_0})$ (which is identical to the behaviour of a DL elastic cross section) to the moderate behaviour of $ln(\frac{s}{s_0})$. Note also a major difference in the diffractive b-space profile which changes from an input central Gaussian to an output peripheral distribution peaking at higher b. Consequently, the GLM model is compatible with the Pumplin bound [54, 55].

$$\frac{\sigma_{el}(s,b)\,+\,\sigma_{diff}(s,b)}{\sigma_{tot}(s,b)}\;\leq\;\frac{1}{2}. \tag{27}$$

3.2 Extension to a multi channel model

The most serious deficiency of a single channel eikonal model is inherent, as the model considers only elastic rescatterings. This is incompatible with the relatively large diffractive cross section observed in the ISR-Tevatron energy range. To this we add a specific problematic feature of the GLM model. Whereas, σ_{tot}, σ_{el} and B_{el} are very well fitted, the reproduction of σ_{sd}, in the available ISR-Tevatron range, is poorer. A possible remedy to these deficiencies is to replace the one channel with a multi channel eikonal model, in which inelastic diffractive intermediate re-scatterings are included as well [38, 39, 56]. However, we have to insure that a multi channel model does improve the diffractive (specifically SD) predictions of the GLM model, while maintaining, simultaneously, its excellent reproductions [30, 31, 32, 33] of the forward elastic amplitude, as well as its appealing results on LRG survival probabilities [35, 36, 37] to be discussed in **3.3**.

In the simplest approximation we consider diffraction as a single hadronic state. We have, thus, two orthogonal wave functions

$$\langle \Psi_h \mid \Psi_d \rangle \;=\; 0. \tag{28}$$

Ψ_h is the wave function of the incoming hadron, and Ψ_d is the wave function of the outgoing diffractive system initiated by the incoming hadron. Denote the interaction operator by **T** and consider two wave functions Ψ_1 and Ψ_2 which are diagonal with respect to **T**. The amplitude of the interaction is given by

$$A_{i,k} = \langle \Psi_i \Psi_k \mid \mathbf{T} \mid \Psi_{i'} \Psi_{k'} \rangle = a_{i,k}\,\delta_{i,i'}\,\delta_{k,k'}. \tag{29}$$

In a 2×2 model $i, k \;=\; 1, 2$. The amplitude $a_{i,k}$ satisfies the diagonal unitarity condition (see Eq.(13))

$$2Im\,a_{i,k}(s,b) \;= \mid a_{i,k}(s,b) \mid^2 \;+ G^{in}_{i,k}(s,b), \tag{30}$$

for which we write the solution

$$a_{i,k}(s,b) \;=\; i\left(1\;-\;e^{-\frac{\Omega_{i,k}(s,b)}{2}}\right), \tag{31}$$

and

$$G^{in}_{i,k} \;=\; 1 - e^{-\Omega_{i,k}(s,b)}. \tag{32}$$

$\Omega_{i,k}(s,b)$ is the opacity of the (i,k) channel with a wave function $\Psi_i \,\times\, \Psi_k$.

$$\Omega_{i,k} \;=\; \nu_{i,k}(s)\,\Gamma^S_{i,k}(s,b) \tag{33}$$

where

$$\nu_{i,k} \;=\; \sigma^{S0}_{i,k}\left(\frac{s}{s_0}\right)^{\Delta}. \tag{34}$$

The factorizable radii are given by

$$R^2_{i,k}(s) \;=\; 2R^2_{i,0} \,+\, 2R^2_{0,k} \,+\, 4\alpha'_{I\!P} ln(\frac{s}{s_0}). \tag{35}$$

$\Gamma^S_{i,k}(s,b)$ is the soft profile of the (i,k) channel. The probability that the final state of two interacting hadron states, with quantum numbers i and k, will be elastic regardless of the intermediate rescatterings is

$$P^S_{i,k}(s,b) \;=\; e^{-\Omega_{i,k}(s,b)} \;=\; \{1 \,-\, a_{i,k}(s,b)\}^2. \tag{36}$$

In the above diagonal representation, Ψ_h and Ψ_d can be written as

$$\Psi_h \;=\; \alpha\Psi_1 \;+\; \beta\Psi_2, \tag{37}$$

$$\Psi_d \;=\; -\beta\Psi_1 \;+\; \alpha\Psi_2. \tag{38}$$

Ψ_1 and Ψ_2 are orthogonal. Since $\mid \Psi_h \mid^2 = \; 1$, we have

$$\alpha^2 \;+\; \beta^2 \;=\; 1. \tag{39}$$

The wave function of the final state is

$$\begin{aligned}\Psi_f \;=& \mid \mathbf{T} \mid \Psi_h \times \Psi_h \rangle \;=\\ &\alpha^2 a_{1,1}\{\Psi_1 \times \Psi_1\} \;+\; \alpha\beta a_{1,2}\{\Psi_1 \times \Psi_2 \;+\; \Psi_2 \times \Psi_1\} \;+\\ &\beta^2 a_{2,2}\{\Psi_2 \times \Psi_2\}.\end{aligned} \tag{40}$$

We have to consider 4 possible re-scattering processes. However, in the case of a $\bar{p}p$ (or pp) collision, single diffraction at the proton vertex equals single diffraction at the antiproton vertex. i.e., $a_{1,2} \;=\; a_{2,1}$ and we end with three channels whose b-space amplitudes are given by

$$a_{el}(s,b) \;=\; \langle \Psi_h \times \Psi_h \mid \Psi_f \rangle \;=\; \alpha^4 a_{1,1} \;+\; 2\alpha^2\beta^2 a_{1,2} \;+\; \beta^4 a_{2,2}, \tag{41}$$

$$a_{sd}(s,b) \;=\; \langle \Psi_h \times \Psi_d \mid \Psi_f \rangle \;=\; \alpha\beta\{\alpha^2 a_{1,1} + (\alpha^2 - \beta^2)a_{1,2} + \beta^2 a_{2,2}\}, \tag{42}$$

$$a_{dd}(s,b) \;=\; \langle \Psi_d \times \Psi_d \mid \Psi_f \rangle \;=\; \alpha^2\beta^2\{a_{1,1} \;-\; 2a_{1,2} \;+\; a_{2,2}\}. \tag{43}$$

In the numeric calculations one may further neglect the double diffraction channel which is exceedingly small in the ISR-Tevatron range. This is obtained by setting $a_{2,2} \;=\; 2a_{1,2} \;-\; a_{1,1}$. Note that in the limit where $\beta \;<<\; 1$, we reproduce the single channel model.

As in the single channel, we simplify the calculation assuming a Gaussian b-space distribution of the input opacities soft profiles

$$\Gamma^S_{i,k}(s,b) \;=\; \frac{1}{\pi R^2_{i,k}(s)}\, e^{-\frac{b^2}{R^2_{i,k}(s)}}. \tag{44}$$

The opacity expressions, just presented, allow us to express the physical observables of interest as functions of $\nu_{1,1}$, $\nu_{1,2}$, $R^2_{1,1}$, $R^2_{1,2}$ and β, which is a constant of the model. The determination of these variables enables us to produce a global fit to the total, elastic and diffractive cross sections as well as the elastic forward slope. This has been done in a two channel model, in which σ_{dd} is neglected [38]. The main conclusion of this study is that the extension of the GLM model to a multi channel eikonal results with a very good overall reproduction of the data. The results maintain the b-space peripherality of the diffractive output amplitudes and satisfy the Pumplin bound [54, 55]. Note that since different experimental groups have been using different algorithms to define diffraction, the SD experimental points are too scattered to enable a tight theoretical reproduction of the diffractive data, see Fig.2.

3.3 Survival probabilities of LRG in the GLM model

The eikonal model simplifies the calculation of the survival probability, Eq.(3), associated with the soft re-scatterings of the spectator partons. We can, thus, eliminate the nominator and denominator terms in $\mid M^H(s,b) \mid^2$ which depend exclusively on s. In the GLM model we assume a Gaussian b-dependence for $\mid M^H(s,b) \mid^2$ corresponding to a constant hard radius ${R^H}^2$. This choice enables an analytic solution of Eq.(3). More elaborate choices, such as dipole or multi poles distributions, require a numerical evaluation of this equation.

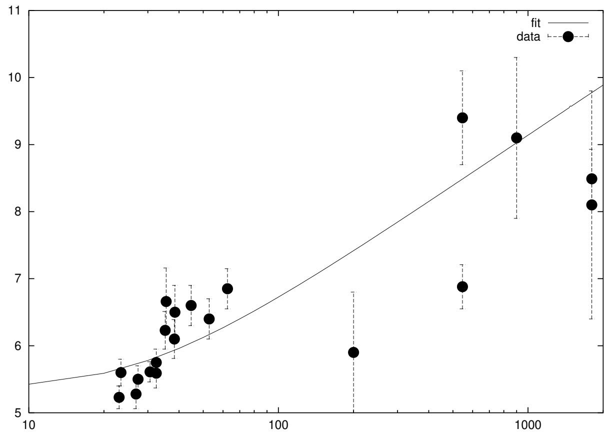


Fig. 2: *Integrated SD data and a two channel model fit.*

Define,

$$a_H(s) \;=\; \frac{R^2(s)}{R^{H^2}(s)} > 1. \tag{45}$$

$a_H(s)$ grows logarithmically with s. As stated, Eq.(3) can be analytically evaluated with our choice of Gaussian profiles and we get

$$S^2 \;=\; \frac{a_H(s)\gamma[a_H(s),\nu(s)]}{[\nu(s)]^{a_H(s)}}, \tag{46}$$

where $\gamma(a,\nu)$ denotes the incomplete Euler gamma function

$$\gamma(a,x) \;=\; \int_0^x z^{a-1}e^{-z}dz. \tag{47}$$

The solution of Eq.(46), at a given s, depends on the input values of R^{H^2}, R^2 and $\nu(s)$. In the GLM approach, R^{H^2} is estimated from the excellent HERA data [57, 58, 59] on $\gamma + p \to J/\Psi + p$. The values of $\nu(s)$ and $R^2(s)$ are obtained from the experimental $\bar{p}p$ data. This can be attained from a global fit to the soft scattering data [38]. Alternatively, we can obtain ν from the ratio $\frac{\sigma_{el}}{\sigma_{tot}}$ and then obtain the value of R^2 from the explicit expressions given in Eqs.(24,25,26). LHC predictions presently depend on model calculations with which this information can be obtained. Once we have determined $\nu(s)$ and $a_H(s)$, the survival probability is calculated from Eq.(46).

In the GLM three channel model we obtain for central hard diffraction of di-jets or Higgs a survival probability,

$$S^2_{CD}(s) \;=\; \frac{\int d^2b\,\left(\alpha^4\,P^S_{1,1}\,{\Omega^H_{1,1}}^2 \;+\; 2\alpha^2\beta^2\,P^S_{1,2}\,{\Omega^H_{1,2}}^2 \;+\; \beta^4 P^S_{2,2}\,{\Omega^H_{2,2}}^2\right)}{\int\, d^2b\left(\alpha^4\,{\Omega^H_{1,1}}^2 \;+\; 2\alpha^2\beta^2\,{\Omega^H_{1,2}}^2 \;+\; \beta^4\,{\Omega^H_{2,2}}^2\right)}. \tag{48}$$

The hard diffractive cross sections in the (i,k) channel are calculated using the multi particle optical theorem [53]. They are written in the same form as the soft amplitudes

$${\Omega^H_{i,k}}^2 \;=\; \nu^H_{i,k}(s)^2\,\Gamma^H_{i,k}(b), \tag{49}$$

where,

$$\nu^H_{i,k} \;=\; \sigma^{H0}_{i,k}\left(\frac{s}{s_0}\right)^{\Delta_H}. \tag{50}$$

As in the single channel calculation we assume that $\Gamma^H_{i,k}(b)$ is Gaussian,

$$\Gamma^H_{i,k}(b) \;=\; \frac{2}{\pi R^2_{i,k}}\, e^{-\frac{2\,b^2}{R^2_{i,k}}}. \tag{51}$$

Note, that the hard radii ${R^H_{i,k}}^2$ are constants derived from HERA J/Ψ photo and DIS production [57, 58, 59].

As it stands, a three channel calculation is not useful since σ_{dd} is very small and the 3'd channel introduces additional parameters which can not be constraint by the meager experimental information on σ_{dd} [13, 14, 15, 16, 17]. In a two channel model Eq.(48) reduces to

$$S^2_{CD}(s) \;=\; \frac{\int d^2b\,\left(P^S_{1,1}\,{\Omega^H_{1,1}}^2 \,-\, 2\beta^2\,(P^S_{1,1}\,{\Omega^H_{1,1}}^2 \,-\, P^S_{1,2}\,{\Omega^H_{1,2}}^2)\right)}{\int\, d^2b\left({\Omega^H_{1,1}}^2 \,-\, 2\beta^2\,({\Omega^H_{1,1}}^2 \,-\, {\Omega^H_{1,2}}^2)\right)}. \tag{52}$$

A new, unpublished yet, model [60], offers an explicit S^2 calculation for the exclusive $NN \to N + LRG + 2J + LRG + N$ final state, both in one and two channel eikonal models. We shall comment on its output in the next subsection.

3.4 GLM S^2 predictions

Following are a few general comments on the GLM calculations of S^2, after which we discuss the input/output features of the single and two channel models. Our objective is to present predictions for LHC.

The only available experimental observable with which we can check the theoretical S^2 predictions is the hard LRG di-jets data obtained in the Tevatron and Hera. A comparison between data and our predictions is not immediate as the basic measured observable is f_{gap} and not S^2. The application of the GLM models to a calculation of f_{gap} depends on an external input of a hard diffractive LRG cross section which is then corrected by S^2 as presented above. Regardless of this deficiency, the introduction of a survival probability is essential so as to understand the huge difference between the pQCD calculated F_{gap} and its experimental value f_{gap}. A direct test of the GLM predictions calls for a dedicated experimental determination of S^2. The only direct S^2 information from the Tevatron is provided by a JGJ ratio measured by D0 [5, 6, 7] in which $\frac{S^2(\sqrt{s}=630)}{S^2(\sqrt{s}=1800)} \;=\; 2.2 \pm 0.8$. This is to be compared with a GLM ratio of $1.2 - 1.3 \pm 0.4$ presented below.

The survival probabilities of the CD, SD and DD channels are not identical. The key difference is that each of the above channels has a different hard radius. A measure of the sensitivity of S^2 to changes in ν and a_H is easy to identify in a single channel calculation which is presented in Fig.3. Indeed, preliminary CDF GJJG data [17] suggest that f_{gap} measured for this channel is moderately smaller than the rate measured for the GJJ channel.

GLM soft profile input is a central Gaussian. This is over simplified, and most models assume a power like dipole or multipole b-dependence of $\Gamma^S(s,b)$ and $\Gamma^H(s,b)$. Explicit comparisons [60] of S^2 obtained with different input profiles shows a diminishing difference between the survival probability outputs, provided their effective radii are compatible.

Regardless of the attractive simplicity of the single channel model, one should add a cautious reminder that the single channel model does not reproduce σ_{sd} well since its survival probabilities are over-estimated. Consequently, we are inclined to suspect that the S^2 values presented in the table below are over-estimated as well.

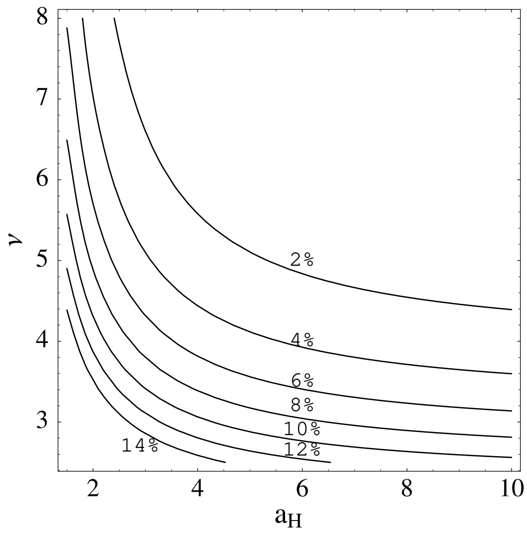


Fig. 3: *A contour plot of $S^2(1C)$ against $\nu(s)$ and $a_H(s)$.*

$\sqrt{s}$ (GeV)	$S^2_{CD}(F1C)$	$S^2_{CD}(D1C)$	$S^2_{SD_{incl}}(F1C)$	$S^2_{SD_{incl}}(D1C)$	$S^2_{DD}(F1C)$	$S^2_{DD}(D1C)$
540	14.4%	13.1%	18.5%	17.5%	22.6%	22.0%
1800	10.9%	8.9%	14.5%	12.6%	18.2%	16.6%
14000	6.0%	5.2%	8.6%	8.1%	11.5%	11.2 %

As we noted, the soft input can be obtained from either a model fit to the soft scattering data or directly from the measured values of $\sigma_{tot}, \sigma_{el}$ and R^{H^2}. The first method is denoted F1C and the second is denoted D1C. Note that having no LHC data, $S^2_{DD}(D1C)$, at this energy, is calculated on the basis of model estimates for the total and elastic cross sections. The constant hard radius $R^{H^2} = 7.2$ is deduced from HERA J/Ψ photoproduction forward exponential slope which shows only diminishing shrinkage [57, 58]. This is a conservative choice which may be changed slightly with the improvement of the Tevatron CDF estimates [61] of R^{H^2}. The two sets of results obtained are compatible, even though, $S^2(D1C)$ is consistently lower than $S^2(F1C)$. The S^2 output presented above depends crucially on the quality of the data base from which we obtain the input parameters. The two sets of Tevatron data at $1800\,GeV$ have a severe $10 - 15\%$ difference resulting in a non trivial ambiguity of the S^2 output.

The global GLM two channel fit [38] reproduces the soft scattering data (including SD) remarkably well with $\beta = 0.464$. Its fitted parameters are used for the soft input required for the S^2 calculations. Our cross section predictions for LHC are: $\sigma_{tot} = 103.8\,mb$, $\sigma_{el} = 24.5\,mb$, $\sigma_{sd} = 12\,mb$ and $B_{el} = 20.5\,GeV^{-2}$. The input for the calculation of S^2 requires, in addition to the soft parameters, also the values of $\nu^H_{i,k}$ and ${R^H_{i,k}}^2$. The needed hard radii can be estimated, at present, only for the CD channel, where we associate the hard radii $R^H_{1,1}$ with the hard radius obtained in HERA exclusive J/Ψ photoproduction [57, 58] and $R^H_{1,2}$ with HERA inclusive J/Ψ DIS production [59]. Accordingly, we have ${R^H_{1,1}}^2 = 7.2\,GeV^{-2}$, and ${R^H_{1,2}}^2 = 2.0\,GeV^{-2}$. We do not have experimental input to determine $\nu^H_{i,k}$. We overcome this difficalty by assuming a Regge-like factorization $\sigma^{H0}_{i,k}/\sigma^{S0}_{i,k} = constant$. Our predictions for the CD survival probabilities are: 6.6% at $540\,GeV$, 5.5% at $1800\,GeV$ and 3.6% at $14000\,GeV$.

These results may be compared with a recent, more elaborate, eikonal formulation [60] aiming to calculate the survival probability of a final exclusive $N + LRG + 2J(orH) + LRG + N$ state. These calculations were done in one and two channel models. The one channel S^2_{CD} predicted values are 14.9% at $540\,GeV$, 10.8% at $1800\,GeV$ and 6.0% at $14000\,GeV$. These values are remarkably similar to the GLM one channel output. In the two channel calculations the corresponding predictions are 5.1%, 4.4%

and 2.7%, which are marginally smaller than the GLM two channel output numbers.

In our assessment, the two channel calculations provide a more reliable estimate of S^2 since they reproduce well the soft scattering forward data. Our estimate for the survival probability associated with LHC Higgs production is $2.5\% - 4.0\%$.

4 The KKMR Model

The main part of this section (**4.1-4.3**) was written by V.A. Khoze, A.D. Martin and M. Ryskin (KMR) and is presented here without any editing.

The KKMR model calculation [40, 41, 42, 43, 44] of the survival probabilities is conceptually quite similar to the GLM model, in as much as unitarization is enforced through an eikonal model whose parameters provide a good reproduction of the high energy soft scattering data. However, the GLM model is confined to a geometrical calculation of S^2 for which we need just the value of ${R^H}^2$, without any specification of the hard dynamics. This value is an external input to the model. The KKMR model contains also a detailed pQCD calculation of the hard diffractive proccess, specifically, central diffractive Higgs production. Consequently, it can predict a cross section for the channel under investigation.

4.1 KKMR model for soft diffraction

The KMR description [41] of soft diffraction in high energy pp (or $p\bar{p}$) collisions embodies

(i) *pion-loop* insertions in the bare Pomeron pole, which represent the nearest singularity generated by t-channel unitarity,

(ii) a *two-channel eikonal* which incorporates the Pomeron cuts generated by elastic and quasi-elastic (with N^* intermediate states) s-channel unitarity,

(iii) high-mass *diffractive dissociation*.

The KKMR model gives a good description of the data on the total and differential elastic cross section throughout the ISR-Tevatron energy interval, see [41]. Surprisingly, KMR found the bare Pomeron parameters to be

$$\Delta \;\equiv\; \alpha(0) \,-\, 1 \;\simeq\; 0.10, \qquad \alpha' \;=\; 0. \tag{53}$$

On the other hand it is known that the same data can be described by a simple effective Pomeron pole with [62, 47, 48]

$$\alpha_{I\!P}^{\rm eff}(t) \;=\; 1.08 \,+\, 0.25\, t. \tag{54}$$

In this approach the shrinkage of the diffraction cone comes not from the bare pole ($\alpha' = 0$), but has components from the three ingredients, (i)–(iii), of the model. That is, in the ISR-Tevatron energy range

$$\text{“}\alpha'_{\rm eff}\text{”} \;=\; (0.034 \,+\, 0.15 \,+\, 0.066)\;{\rm GeV}^{-2} \tag{55}$$

from the π-loop, s-channel eikonalisation and diffractive dissociation respectively. Moreover, eikonal rescattering suppresses the growth of the cross section and so $\Delta \simeq 0.10 > \Delta_{\rm eff} \simeq 0.08$.

Since the model [41] embodies all the main features of soft diffraction KMR expect it to give reliable predictions for the *survival probability* S^2 of the rapidity gaps against population by secondary hadrons from the underlying event, that is hadrons originating from soft rescattering. In particular, KMR predict $S^2 = 0.10$ (0.06) for single diffractive events and $S^2 = 0.05$ (0.03) for exclusive Higgs boson production, $pp \to p + H + p$, at Tevatron (LHC) energies.

4.2 Calculation of the exclusive Higgs signal

The basic mechanism for the exclusive process, $pp \to p + H + p$, is shown in Fig. 4. The left-hand gluon Q is needed to screen the colour flow caused by the active gluons q_1 and q_2. Since the dominant

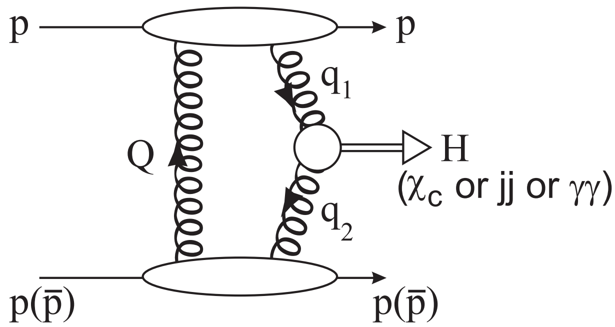


Fig. 4: Schematic diagram for central exclusive production, $pp \to p + X + p$. The presence of Sudakov form factors ensures the infrared stability of the Q_t integral over the gluon loop. It is also necessary to compute the probability, S^2, that the rapidity gaps survive soft rescattering.

contribution comes from the region $\Lambda^2_{\rm QCD} \ll Q_t^2 \ll M_H^2$, the amplitude may be calculated using perturbative QCD techniques [63, 40]

$$\mathcal{M}_H \simeq N \int \frac{dQ_t^2}{Q_t^4}\, f_g(x_1, x_1', Q_t^2, \mu^2) f_g(x_2, x_2', Q_t^2, \mu^2), \tag{56}$$

where the overall normalisation constant N can be written in terms of the $H \to gg$ decay width [64, 40]. The probability amplitudes (f_g) to find the appropriate pairs of t-channel gluons (Q, q_1) and (Q, q_2) are given by the skewed unintegrated gluon densities at the hard scale μ, taken to be $0.62\ M_H$. Since the momentum fraction x' transfered through the screening gluon Q is much smaller than that (x) transfered through the active gluons ($x' \sim Q_t/\sqrt{s} \ll x \sim M_H/\sqrt{s} \ll 1$), it is possible to express $f_g(x, x', Q_t^2, \mu^2)$, to single log accuracy, in terms of the conventional integrated density $g(x)$ [65, 66, 67, 68]. The f_g's embody a Sudakov suppression factor T, which ensures that the gluon does not radiate in the evolution from Q_t up to the hard scale $\mu \sim M_H/2$, and so preserves the rapidity gaps.

It is often convenient to use the simplified form [40]

$$f_g(x, x', Q_t^2, \mu^2) \;=\; R_g\, \frac{\partial}{\partial \ln Q_t^2} \left[\sqrt{T_g(Q_t, \mu)}\; xg(x, Q_t^2) \right], \tag{57}$$

which holds to 10–20% accuracy.[1] The factor R_g accounts for the single $\log Q^2$ skewed effect [67]. It is found to be about 1.4 at the Tevatron energy and about 1.2 at the energy of the LHC.

4.3 The Sudakov factor

The Sudakov factor $T_g(Q_t, \mu)$ reads [65, 66, 69]

$$T_g(Q_t, \mu) = \exp\left(-\int_{Q_t^2}^{\mu^2} \frac{\alpha_S(k_t^2)}{2\pi} \frac{dk_t^2}{k_t^2} \left[\int_{\Delta}^{1-\Delta} z P_{gg}(z) dz \;+\; \int_0^1 \sum_q \; P_{qg}(z) dz \right] \right), \tag{58}$$

with $\Delta = k_t/(\mu + k_t)$. The square root arises in (57) because the (survival) probability not to emit any additional gluons is only relevant to the hard (active) gluon. It is the presence of this Sudakov factor which makes the integration in (56) infrared stable, and perturbative QCD applicable[2].

[1] In the actual computations a more precise form, as given by Eq. (26) of [68], was used.

[2] Note also that the Sudakov factor inside t integration induces an additional strong decrease (roughly as M^{-3} [44]) of the cross section as the mass M of the centrally produced hard system increases. Therefore, the price to pay for neglecting this suppression effect would be to considerably overestimate the central exclusive cross section at large masses.

It should be emphasized that the presence of the double logarithmic T-factors is a purely classical effect, which was first discussed in 1956 by Sudakov in QED. There is strong bremsstrahlung when two colour charged gluons 'annihilate' into a heavy neutral object and the probability not to observe such a bremsstrahlung is given by the Sudakov form factor[3]. Therefore, any model (with perturbative or non-perturbative gluons) must account for the Sudakov suppression when producing exclusively a heavy neutral boson via the fusion of two coloured particles.

More details of the role of the Sudakov suppression can be found in J. Forshaw's review in these proceedings [34]. Here KMR would like to recall that the T-factors in [70, 44] were calculated to *single* log accuracy. The collinear single logarithms were summed up using the DGLAP equation. To account for the 'soft' logarithms (corresponding to the emission of low energy gluons) the one-loop virtual correction to the $gg \to H$ vertex was calculated explicitly, and then the scale $\mu = 0.62\ M_H$ was chosen in such a way that eq.(58) reproduces the result of this explicit calculation. It is sufficient to calculate just the one-loop correction since it is known that the effect of 'soft' gluon emission exponentiates. Thus (58) gives the T-factor to single log accuracy.

In some sense, the T-factor may be considered as a 'survival' probability not to produce any hard gluons during the $gg \to H$ fusion subprocess. However, it is not just a number (i.e. a numerical factor) which may be placed in front of the integral (the 'bare amplitude'). Without the T-factors hidden in the unintegrated gluon densities f_g the integral (56) diverges. From the formal point of view, the suppression of the amplitude provided by T-factors is infinitely strong, and without them the integral depends crucially on an ad hoc infrared cutoff.

4.4 Summary of KKMR S^2 predictions

A compilation of S^2 values obtained in the KKMR model is presented below:

$\sqrt{s}$ (GeV)	$S^2_{2C}(CD)$	$S^2_{2C}(SD_{incl})$	$S^2_{2C}(DD)$
540	6.0%	13.0%	20.0%
1800	4.5%	10.0%	15.0%
14000	2.6%	6.0%	10.0%

A comparison with the corresponding GLM two channel model is possible only for the available GLM CD channel, where, the KKMR output is compatible with GLM. KKMR SD and DD output are compatible with the corresponding GLM single channel numbers. Overall, we consider the two models to be in a reasonable agreement.

A remarkable utilization of the KKMR model is attained when comparing the HERA [18, 19, 20, 21, 22, 23, 24, 25, 26, 27] and CDF [8, 9, 10, 11, 12, 17] di-jets diffractive structure functions derived for the dynamically similar GJJ channels. To this end, the comparison is made between the kinematically compatible HERA $F^D_{jj}(Q^2 = 75\,GeV^2, \beta)$ and the CDF $F^D_{jj}(< E_T^2 >= 75\,GeV^2, \beta)$. The theoretical expectation is that $F^D_{jj}(\beta)$, as measured by the two experiments, should be very similar. As can be seen in Fig.5, the normalizations of the two distributions differ by approximately an order of magnitude and for very small $\beta < 0.15$ there is a suggestive change in the CDF distribution shape. This large discrepancy implies a breaking of QCD and/or Regge factorization. Reconsidering, it is noted, that HERA DIS data is measured at a high Q^2 where the partonic interactions induced by the highly virtual photon are point like and, hence, $S^2 = 1$. On the other hand, CDF GJJ measurement is carried out at $1800\,GeV$ and, as we saw, its survival probability is rather small. The convolution between the HERA determined GJJ $F^D_{jj}(\beta)$ and the β dependent survival probabilities, as calculated by KKMR, provides the $F^D_{jj}(\beta)$ distribution corrected for the soft rescattering of the spectator partons. This is shown in Fig.5

[3] It is worth mentioning that the $H \to gg$ width and the normalization factor N in (56) is an 'inclusive' quantity which includes all possible bremsstrahlung processes. To be precise, it is the sum of the $H \to gg + ng$ widths, with n=0,1,2,... . The probability of a 'purely exclusive' decay into two gluons is nullified by the same Sudakov suppression.

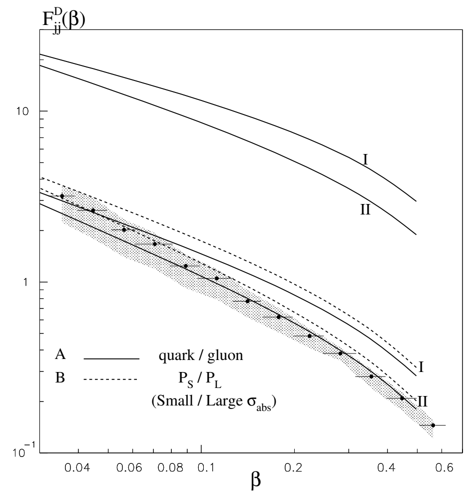


Fig. 5: The predictions for the diffractive di-jets production at the Tevatron (lower lines), obtained from two alternative sets of HERA diffractive parton distributions I and II, compared with the CDF data (shaded area). The upper lines correspond to the Tevatron prediction neglecting the survival probability correction.

and provides an impressive reproduction of the experimental distribution. We were informed [71] that this analysis was successfully redone with an updated H1 produced structure function distribution.

The weak element in the above analysis is that it is crucially dependent on the H1 determined $F^D_{jj}(\beta)$ distribution. ZEUS has constructed a somewhat different structure function. Clearly, a very different experimental determination of $F^D_{jj}(\beta)$, such as been recently suggested by Arneodo [72], will re-open this analysis for further studies, experimental and theoretical.

4.5 A Comparison between KKMR and GLM

The approach of GLM and KKMR to the calculation of forward soft scattering in the ISR-Tevatron range are basically similar. Both models utilize the eikonal model assuming different input soft profiles which have, nevertheless, compatible effective radii. There are, though, a few particular differences between the two sets of calculations:
1) The GLM model, with a Gaussian soft profile, is applicable only in the forward cone ($|t| < 0.3\,GeV^2$), where we have most of the data of interest. KKMR use a multipole power behaviour profile which enables applicability over a, somewhat, wider t range, $|t| < 0.5\,GeV^2$. Note that, the GLM output is not significantly changed with a multipole power behaviour profile provided its radii are compatible with the Gaussian input [60].
2) The GLM input Pomeron trajectory is specified by $\Delta = 0.12$ and $\alpha'_{I\!P} = 0.2$. These evolve due to eikonalisation to an effective output of $\epsilon = 0.08$ and $\alpha'_{I\!P} = 0.25$. Note that, Δ is obtained in GLM as a fitted output parameter. In KKMR, the relatively high input $\Delta \simeq 0.2$ is theoretically tuned by a pion loop renormalization resulting in an input value of $\Delta \simeq 0.1$. KKMR have a more elaborate treatment of $\alpha_{I\!P}(t)$ than GLM, resulting, nevertheless, with forward cone output predictions similar to GLM. However, KKMR accounts for a somewhat wider t range than GLM and reproduces the t dependence of B_{el} well. Similar results are obtained in a GLM version [39, 56] in which the soft profile is given by a dipole distribution. KKMR can predict a few differential properties of S^2, which are beyond the scope of GLM.
3) Both models treat the high mass diffraction with the triple Pomeron formalism [53]. In GLM the final SD cross section is obtained by a convolution of the input $\frac{d\sigma_{sd}}{d^2b}$ with $P^S(s,b)$. In KKMR the treatment

of the SD amplitude is more elaborate, ending, though, with no detailed SD data reconstruction which is presented in GLM.
4) The LHC predictions of the two models for cross sections and slopes are compatible, with the exception of σ_{dd} which is neglected in GLM and acquires a significant KKMR predicted value of $9.5\,mb$.

GLM is a geometrical model where both the input hard LRG non corrected matrix element squared and the soft elastic scattering amplitude, are approximated by central Gaussians in b-space. This property enables us to easily calculate the survival probabilities which depend on ν, R^2 and ${R^H}^2$ in a single channel input, and on $\nu_{i,k}$, $R^2_{i,k}$ and ${R^H_{i,k}}^2$ in a two channel input. As we have noted, the GLM model, on its own, cannot provide a calculation of F_{gap} and f_{gap} as it needs the hard radii as an external input. The KKMR model is more sophisticated. This is attributed to the fact that the hard diffractive LRG process is explicitly calculated in pQCD, hence the non corrected F_{gap} and the corrected f_{gap} and F^D_{jj} are model predictions. As we have just noted, given the hard diffractive matrix element, the actual calculation of the diffractive LRG survival probability damping is almost identical to GLM. Keeping this basic observation in mind, it is constructive to compare the features of the two models with a special interest on the input assumptions and output differences of the two models.

The main difference between the two models is reflected in the level of complexity of their inputs. GLM soft input is obtained from a simple eikonal model for the soft forward scattering, to which we add the hard radii which are derived from the HERA data. KKMR calculations of P^S are equally simple. The calculation of the hard sector matrix elements are, naturally, more cumbersome. Given HERA $F^D_{jj}(Q^2,\beta)$, a Tevatron diffractive F^D_{jj} in which $< E_T >$ and Q^2 are comparable, can be calculated, parameter free, without the need to calculate the hard amplitude. But this is a particular case and, in general, the KKMR calculation depends on an extended parameter base, such as the the input p.d.f. and pQCD cuts. These input parameters are not constrained tightly enough.

The elaborate structure of the KKMR model provides a rich discovery potential which is reflected in the model being able to define and calculate the dependence of S^2 not only on b, but also on other variables, notably β, and experimental cuts such as the recoil proton transverse momentum. GLM depends on the hard radii external information obtained from HERA data. It lacks the potential richness of KKMR. GLM can serve, though, as a standard through which we can compare different unitarized models. Given such a model, we can extract effective values for ν, R^2 and ${R^H}^2$ and proceed to a simple calculation of S^2. We shall return to this proposed procedure in the final discussion.

Even though both GLM and KKMR are two channel models, they are dynamically different. GLM two channel formulation relates to the diversity of the intermediate soft re-scatterings, i.e. elastic and diffractive for which we have different soft amplitudes $a_{i,k}$, each of which is convoluted with a different probability $P^S_{i,k}$ which depends on a different interaction radius $R^2_{i,k}$. In the KKMR model the two channels relate to two different dynamical options of the hard process. In model A the separation is between valence and sea interacting partons. In model B the separation is between small and large dipoles. The two models give compatible results. The key point, though, is that the KKMR opacities $\Omega_{i,k}$, in the definition of $P^S_{i,k}$, differ in their normalization, but have the same b-dependence. Regardless of this difference the output of the GLM and KKMR models is reasonably compatible. The compatibility between GLM and KKMR is not surprising since the explicit KKMR calculation of the hard LRG amplitude is approximated relatively well by the GLM simple Gaussian.

Our final conclusion is that the two model output sets are compatible. The richness of the KKMR model has a significant discovery potential lacking in GLM. On the other hand, the GLM simplicity makes it very suitable as a platform to present different models in a uniform way, which enables a transparent comparison.

5 Discussion

As we shall see, at the end of this section, there is no significant difference between the values of σ_{tot} predicted by DL and GLM up to the top Cosmic Rays energies. This is, even though, DL is a Regge model without unitarity corrections. The explanation for this "paradox" is that the DL amplitude violations of s-unitarity are confined, even at super high energies, to small b which does not contribute significantly to σ_{tot}. Note, though, that $\frac{\sigma_{el}}{\sigma_{tot}}$ grows in DL like s^{ϵ} whereas in GLM its growth is continuosly being moderated with increasing s (see table in **5.3**). The DL model predicts that S^2 is identical to unity or very close to it in the DL high-t model where a weak $I\!\!P I\!\!P$ cut is added. The need for survival probabilities so as to reproduce the the experimental soft SD cross section values and the hard di-jets rates, is the most compelling evidence in support of unitarization at presently available energies. As such, the study of high energy soft and hard diffraction serves as a unique probe substantiating the importance of s-channel unitarity in the analysis of high energy scattering.

5.1 S^2 in unitarized models

Most, but not all, of the unitarized models dealing with LHC S^2 predictions have roughly the same S^2 values. This calls for some clarifications. The first part of our discussion centers on the correlated investigation of two problems:
1) How uniform are the output predictions of different unitarization procedures?
2) How sensitive are the eikonal calculations to the details of the eikonal model they use?

We start with two non eikonal models which have contradictory predictions.

The first is a model suggested by Troshin and Tyurin [52]. In this model the single channel unitarity constraint (Eq.(13)) is enforced with an asymptotic bound where $G_{in} = 0$ and $|a_{el}| = 2$ i.e. asymptotically, $\sigma_{tot} = \sigma_{el}$ and $P^S(s,b) = 1$. The parameters of the model are set so as to obtain a "normal" survival probability monotonically decreasing with energy up to about $2500\,GeV$ where it changes its behavior and rises monotonically to its asymptotic limit of 1. Beside the fact that the model has a legitimate but non appealing asymptotics, its main deficiency is that it suggests a dramatic change in the systematics of S^2 without being able to offer any experimental signature to support this claim. Regardless of this criticism, this is a good example of a proper unitarity model whose results are profoundly different from the eikonal model predictions.

Another non eikonal procedure is Goulianos flux renormalization model [17]. This is a phenomenological model which formally does not enforce unitarity, but rather, a bound of unity on the Pomeron flux in diffractive processes. Note that, the Pomeron flux is not uniquely defined so this should be regarded as an ad hoc parametrization. Nevertheless, it has scored an impressive success in reproducing the soft and hard diffractive data in the ISR-Tevatron range. The implied survival probabilities of this procedure are compatible with GLM and KKMR. However, the model predicts suppression factors for the diffractive channels which are t-independent and, thus, b-independent. The result is that, even though the output diffractive cross section is properly reduced relative to its input, there is no change of the output profile from its input Gaussian form. Consequently, the Pumplin bound is violated. We are informed that Goulianos plans to improve his model by eikonalizing the output of his present model.

As noted, there are a few eikonal models on the market [73, 74, 75, 76, 77, 78, 79, 80], and their predictions are compatible with GLM and KKMR. Reconsidering the procedure of these calculations, their compatibility is not surprising once we translate their input to a GLM format. The GLM eikonal S^2 calculation has two input sectors in either a single or a two channel version. They are the soft ν and R^2, and the hard radius R^{H^2}. Since the soft input is based on a fit of the soft scattering data base, the potential variance in the soft parameters is relatively small. The input hard radius is obtained from either the HERA data or a theoretical calculation, be it a pQCD diagram or a Regge model. All in all, this is a reasonably stable input. In this context, it is interesting to discuss the eikonal model of Block and Halzen [73], where the calculated survival probabilities for Higgs production through W-W fusion are seemingly too high, $S^2(540) = 27\%$, $S^2(1800) = 21\%$ and $S^2(14000) = 13\%$. Even though, Higgs production is a CD process, the above S^2 values are in agreement with the KKMR calculations of S^2_{DD} with a relatively high $R^{H^2} = 11\,GeV^{-2}$. In a proper S^2_{CD} calculation, these high S^2 values correspond to an even higher $R^{H^2} \simeq 20\,GeV^{-2}$, which is far too high as an estimate of the hard radius of $WW \to H$. A possible interpretation of Block-Halzen results is to associate them with a soft, rather than a hard, LRG CD process. This would couple with the non screened interpretation of CD Higgs through the soft CEM model [74, 75], which predicts very high S^2 values. Since the CEM model is not screened we may, as well, assign a survival probability to its output result. This translates into $S^2_{CD} = S^2_{BH} S^2_{CEM}$, providing rather reasonable one channel predictions, $S^2_{CD}(540) = 18.9\%$ and $S^2_{CD}(1800) = 7.2\%$.

Obviously, each of the eikonal models, quoted above has its own particular presentation and emphasis. They do, however, have compatible results reflecting the observation that their input translates into similar values of ν, R^2 and R^{H^2}.

5.2 Compatibility between HERA and the Tevatron di-jets data

Much attention has been given recently to the compatibility between the Tevatron and HERA DIS GJJ data. The starting point made by KKMR and CDF is that rather than depend on a p.d.f. input to calculate F_{gap}, we may use, the GJJ di-jets diffractive structure function, F^D_{jj}, inferred from HERA DIS data [18, 19, 20, 21, 22, 23, 24, 25, 26, 27] and associate it with the F^D_{jj} derived from the Tevatron GJJ data. As it stands, this procedure ignores the role of the survival probability. Consequently, F^D_{jj} obtained from the Tevatron is an order of magnitude smaller than the HERA output [8, 9, 10, 11, 12, 17, 40, 41, 42, 43, 44]. This result led to speculations about a possible breaking of QCD or Regge factorization or both. Once the Tevatron di-jets diffractive structure function is rescaled by the appropriate survival probability, the compatibility between the Tevatron and HERA DIS diffractive data is attained. The conclusion of this analysis is that the breaking of factorization is attributed to the soft re-scatterings of the the colliding projectiles. Additional hard contribution to the factorization breaking due to gluon radiation is suppressed by the Sudakov factor included in the pQCD calculation (see **4.3**).

One should note, though, that the H1 determination [18, 19, 20, 21, 22, 23, 24, 25, 26, 27] of F^D_{jj} is not unique. Arneodo [72] has suggested a different F^D_{jj} output based on HERA di-jets data which has a different normalization and β dependences. Should this be verified, there might well be a need to revise the KKMR calculations.

The evolution of HERA F^D_{jj} from high Q^2 DIS to $Q^2 = 0$ di-jets photoproduction has raised additional concern with regard to the validity of the factorization theorems [28, 29]. This is a complicated analysis since one has to be careful on two critical elements of the calculations:
1) The determination of the ratio between direct and resolved exchanged photon (real or virtual). This is a crucial element of the theoretical calculation since survival probability is applicable only to the resolved photon component. For very high Q^2 data the hard scattering process with the target partons is direct. At $Q^2 = 0$ there is a significant resolved photon contribution.
2) For di-jets production there is a big difference between the LO and the NLO pQCD calculated cross sections [81, 82, 83]. Since the HERA analysis compares the pQCD calculation with the di-jets measured cross section the normalization and shape of the theoretical input is most crucial in the experimental comparison between the high Q^2 and $Q^2 = 0$ data.
On the basis of a NLO calculation, Klasen and Kramer [81, 82] conclude that they can reproduce the photoproduction data with $S^2 = 0.34$, applied to the resolved sector. This survival probability is in agreement with KKMR and GLM calculations.
Regardless of the above, preliminary photoproduction GJJ HERA data [28, 29] suggest that both the direct and resolved photon sectors are suppressed at $Q^2 = 0$. A verification of this observation has severe consequences for our understanding of the evolution of the diffractive structure function from

DIS to photoproduction. It does not directly relate, though, to the issue of soft survival probability which apply, per definition, only to the resolved photon sector. The suggested effect in the direct photon sector should, obviously be subject to a good measure of caution before being substantiated by further independent analysis.

5.3 Diffraction at energies above the LHC

We end with a table showing the GLM two channel predictions for energies including the LHC, and up to the top Cosmic Rays energies.

$\sqrt{s}\,[GeV]$	$\sigma_{tot}^{DL}\,[mb]$	$\sigma_{tot}^{GLM}\,[mb]$	$\sigma_{el}^{GLM}\,[mb]$	$\sigma_{sd}^{GLM}\,[mb]$	$B_{el}^{GLM}\,[GeV^{-2}]$	$S_{CD}^{GLM\,2}$
540	60.1	62.0	12.3	8.7	14.9	0.066
1800	72.9	74.9	15.9	10.0	16.8	0.055
14000	101.5	103.8	24.5	12.0	20.5	0.036
30000	114.8	116.3	28.6	12.7	22.0	0.029
60000	128.4	128.7	32.8	13.2	23.4	0.023
90000	137.2	136.5	35.6	13.5	24.3	0.019
120000	143.6	142.2	37.6	13.7	24.9	0.017

The, somewhat, surprizing observation is that the GLM calculated total cross sections are compatible with the DL simple Regge predictions all over the above energy range. This is a reflection of the fact that even at exceedingly high energies unitarization reduces the elastic amplitude at small enough b values to be relatively insensitive to the calculation of σ_{tot}. On the other hand, we see that σ_{el} becomes more moderate in its energy dependence and σ_{el}/σ_{tot} which is 23.6% at the LHC is no more than 26.4% at the highest Cosmic Rays energy, $120\,TeV$. The implication of this observation is that the nucleon profile becomes darker at a very slow rate and is grey (well below the black disc bound) even at the highest energy at which we can hope for a measurment. A check of our results at the Planck scale shows $\sigma_{tot} = 1010\,mb$ and the profile to be entirely black. i.e., $\frac{\sigma_{el}}{\sigma_{tot}} = \frac{1}{2}$. σ_{sd} is even more moderate in its very slow rise with energy. The diminishing rates for soft and hard diffraction at exceedingly high energies are a consequence of the monotonic reduction in the values of S^2 with a Planck scale limit of $S^2 = 0$. This picture is bound to have its effect on Cosmic Rays studies.

Our LHC predictions are compatible with KMR. Note, though, that: i) σ_{sd}^{GLM} is rising slowly with s gaining 20% from the Tevatron to LHC. KMR has a much faster rise with energy, where, σ_{sd}^{KMR} is gaining $77\% - 92\%$ over the same energy interval. ii) At the LHC $B_{el}^{GLM} = 20.5\,GeV^{-2}$, to be compared with a DL slope of $19\,GeV^{-2}$ and a KMR slope of $22\,GeV^{-2}$. The GLM $30\,TeV$ cross sections are compatible with Block-Halzen.

6 Acknowledgements

We are very thankful to our colleagues Valery Khoze, Alan Martin, Misha Ryskin and Leif Lönnblad, who generously contributed to **Section 4** and the **Appendix**. Needless to say, they bear no responsibility for the rest of this review.

Appendix: Monte Carlo modeling of gap survival

The following was contributed by Leif Lönnblad and is presented without any editing.

An alternative approach to gap survival and factorization breaking is to implement multiple interactions in Monte Carlo event generators. These models are typically based on the eikonalization of the partonic cross section in hadronic collisions and can be combined with any hard sub process to describe the additional production of hadrons due to secondary partonic scatterings. Some of these programs, such as PYTHIA [84, 85] and HERWIG/JIMMY [86, 87, 88], are described in some detail elsewhere in these proceedings [89]. Common for all these models is that they include exact kinematics and flavour conservation, which introduces some non-trivial effects and makes the multiple scatterings process-dependent. Also, the predictions of the models are very sensitive to the cutoff used to regularize the partonic cross section and to the assumptions made about the distribution of partons in impact parameter space. Nevertheless, the models are quite successful in describing sensitive final-state observables such as multiplicity distributions and jet-pedestal effects [89]. In particular this is true for the model in PYTHIA which has been successfully tuned to Tevatron data[4] by Rick Field [90], the so-called *CDF tune A*.

The PYTHIA model does not make any prediction for the energy dependence of the total cross section - rather this is an input to the model used to obtain the distribution in the number of multiple interactions. PYTHIA can, however, make predictions for gap survival probabilities. This was first done for Higgs production via W-fusion [2], and amounts to simply counting the fraction of events which do not have any additional scatterings besides the W-fusion process. The basic assumption is that any additional partonic scattering would involve a colour exchange which would destroy any rapidity gap introduced by W-fusion process. Since PYTHIA produces complete events, these can also be directly analyzed with the proper experimental cuts. A similar estimate was obtained for the gaps between jets process, both for the Tevatron and HERA case [91].

Recently, PYTHIA was used to estimate gap survival probabilities also for the case of central exclusive Higgs production [92]. As in the case of gaps between jets, the actual signal process is not implemented in PYTHIA, so direct analysis with proper experimental cuts was not possible. Instead a similar hard sub process was used (standard inclusive Higgs production via gluon fusion in this case) and the fraction of events without additional secondary partonic scatterings was identified with the gap survival probability. Using the *CDF tune A* the gap survival probability was estimated to be 0.040 for the Tevatron and 0.026 for the LHC. This is remarkably close both to the values used in [64] obtained in the KKMR model [43], and to the GLM values presented in section 3.4 especially the two-channel ones obtained in [60].

References

[1] Y. L. Dokshitzer, S. I. Troian, and V. A. Khoze, Sov. J. Nucl. Phys. **46**, 712 (1987).

[2] Y. L. Dokshitzer, V. A. Khoze, and T. Sjöstrand, Phys. Lett. **B274**, 116 (1992).

[3] J. D. Bjorken, Int. J. Mod. Phys. **A7**, 4189 (1992).

[4] J. D. Bjorken, Phys. Rev. **D47**, 101 (1993).

[5] D0 Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **72**, 2332 (1994).

[6] D0 Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **76**, 734 (1996). `hep-ex/9509013`.

[7] D0 Collaboration, B. Abbott *et al.*, Phys. Lett. **B440**, 189 (1998). `hep-ex/9809016`.

[8] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 855 (1995).

[9] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **80**, 1156 (1998).

[10] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **81**, 5278 (1998).

[11] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **84**, 5043 (2000).

[12] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **85**, 4215 (2000).

[4]Note that the model in PYTHIA has recently been revised [89]. However, the reproduction of Tevatron data is not as good for the revised model.

[13] K. Goulianos (2003). hep-ph/0306085.

[14] K. Goulianos, J. Phys. **G26**, 716 (2000). hep-ph/0001092.

[15] CDF Collaboration, K. Goulianos, Nucl. Phys. Proc. Suppl. **99A**, 37 (2001). hep-ex/0011059.

[16] K. Goulianos (2004). hep-ph/0407035.

[17] K. Goulianos. Private communication to UM.

[18] ZEUS Collaboration, M. Derrick *et al.*, Phys. Lett. **B315**, 481 (1993).

[19] ZEUS Collaboration, M. Derrick *et al.*, Z. Phys. **C68**, 569 (1995). hep-ex/9505010.

[20] ZEUS Collaboration, M. Derrick *et al.*, Phys. Lett. **B356**, 129 (1995). hep-ex/9506009.

[21] ZEUS Collaboration, M. Derrick *et al.*, Phys. Lett. **B369**, 55 (1996). hep-ex/9510012.

[22] ZEUS Collaboration, J. Breitweg *et al.*, Eur. Phys. J. **C6**, 43 (1999). hep-ex/9807010.

[23] H1 Collaboration, T. Ahmed *et al.*, Nucl. Phys. **B429**, 477 (1994).

[24] H1 Collaboration, T. Ahmed *et al.*, Phys. Lett. **B348**, 681 (1995). hep-ex/9503005.

[25] H1 Collaboration, C. Adloff *et al.*, Z. Phys. **C76**, 613 (1997). hep-ex/9708016.

[26] F.-P. Schilling, *Measurement and nlo dglap qcd interpretation of diffractive deep-inelastic scattering at hera*. Paper 089 submitted to EPS 2003 Conf. Aachen.

[27] A. A. Savin. Prepared for NATO Advanced Research Workshop on Diffraction 2002, Alushta, Ukraine, 31 Aug - 6 Sep 2002.

[28] H. Abramowicz, ECONF **C0406271**, MONT04 (2004). hep-ex/0410002.

[29] O. Gutsche, *Dijets in diffractive photoproduction and deep-inelastic scattering at hera*. Contribution No. 6-0177 submitted to ICHEP04, Beijing Aug 2004.

[30] E. Gotsman, E. M. Levin, and U. Maor, Z. Phys. **C57**, 677 (1993). hep-ph/9209218.

[31] E. Gotsman, E. M. Levin, and U. Maor, Phys. Rev. **D49**, 4321 (1994). hep-ph/9310257.

[32] E. Gotsman, E. M. Levin, and U. Maor, Phys. Lett. **B353**, 526 (1995). hep-ph/9503394.

[33] E. Gotsman, E. M. Levin, and U. Maor, Phys. Lett. **B347**, 424 (1995). hep-ph/9407227.

[34] J. Forshaw, *Diffractive Higgs production: theory*. These Proceedings.

[35] E. Gotsman, E. M. Levin, and U. Maor, Phys. Lett. **B309**, 199 (1993). hep-ph/9302248.

[36] E. Gotsman, E. Levin, and U. Maor, Nucl. Phys. **B493**, 354 (1997). hep-ph/9606280.

[37] E. Gotsman, E. Levin, and U. Maor, Phys. Lett. **B438**, 229 (1998). hep-ph/9804404.

[38] E. Gotsman, E. Levin, and U. Maor, Phys. Lett. **B452**, 387 (1999). hep-ph/9901416.

[39] E. Gotsman, E. Levin, and U. Maor, Phys. Rev. **D60**, 094011 (1999). hep-ph/9902294.

[40] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C14**, 525 (2000). hep-ph/0002072.

[41] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C18**, 167 (2000). `hep-ph/0007359`.

[42] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C19**, 477 (2001). `hep-ph/0011393`.

[43] A. B. Kaidalov, V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C21**, 521 (2001). `hep-ph/0105145`.

[44] A. B. Kaidalov, V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C33**, 261 (2004). `hep-ph/0311023`.

[45] P. D. B. Collins, *An Introduction to Regge Theory and High-Energy Physics*. Cambridge 1977, 445p.

[46] L. e. Caneschi, *Current physics sources and comments vol. 3: Regge theory of low p_t hadronic interactions*. North Holland Pub (1989).

[47] A. Donnachie and P. V. Landshoff, Phys. Lett. **B296**, 227 (1992). `hep-ph/9209205`.

[48] A. Donnachie and P. V. Landshoff, Z. Phys. **C61**, 139 (1994). `hep-ph/9305319`.

[49] M. M. Block, K. Kang, and A. R. White, Int. J. Mod. Phys. **A7**, 4449 (1992).

[50] CDF Collaboration, F. Abe *et al.*, Phys. Rev. **D50**, 5535 (1994).

[51] T. T. Chou and C.-N. Yang, Phys. Rev. **170**, 1591 (1968).

[52] S. M. Troshin and N. E. Tyurin, Eur. Phys. J. **C39**, 435 (2005). `hep-ph/0403021`.

[53] A. H. Mueller, Phys. Rev. **D2**, 2963 (1970).

[54] J. Pumplin, Phys. Rev. **D8**, 2899 (1973).

[55] J. Pumplin, Phys. Scripta **25**, 191 (1982).

[56] T. Gitman. TAU M.Sc. Thesis, unpublished (2003).

[57] H1 Collaboration, C. Adloff *et al.*, Phys. Lett. **B483**, 23 (2000). `hep-ex/0003020`.

[58] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C24**, 345 (2002). `hep-ex/0201043`.

[59] H. Kowalski and D. Teaney, Phys. Rev. **D68**, 114005 (2003). `hep-ph/0304189`.

[60] E. Gotsman, H. Kowalski, E. Levin, U. Maor, and A. Prygarin. Preprint in preparation (2005).

[61] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **79**, 584 (1997).

[62] A. Donnachie and P. V. Landshoff, Nucl. Phys. **B231**, 189 (1984).

[63] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Phys. Lett. **B401**, 330 (1997). `hep-ph/9701419`.

[64] V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C23**, 311 (2002). `hep-ph/0111078`.

[65] M. A. Kimber, A. D. Martin, and M. G. Ryskin, Phys. Rev. **D63**, 114027 (2001). `hep-ph/0101348`.

[66] M. A. Kimber, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C12**, 655 (2000). hep-ph/9911379.

[67] A. G. Shuvaev, K. J. Golec-Biernat, A. D. Martin, and M. G. Ryskin, Phys. Rev. **D60**, 014015 (1999). hep-ph/9902410.

[68] A. D. Martin and M. G. Ryskin, Phys. Rev. **D64**, 094017 (2001). hep-ph/0107149.

[69] G. Watt, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C31**, 73 (2003). hep-ph/0306169.

[70] A. B. Kaidalov, V. A. Khoze, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C31**, 387 (2003). hep-ph/0307064.

[71] V. Khoze. Private communication to UM, July 2005.

[72] M. Arneodo. Talk at HERA/LHC CERN Meeting and private communication to UM, Oct. 2004.

[73] M. M. Block and F. Halzen, Phys. Rev. **D63**, 114004 (2001). hep-ph/0101022.

[74] O. J. P. Eboli, E. M. Gregores, and F. Halzen, Phys. Rev. **D61**, 034003 (2000). hep-ph/9908374.

[75] O. J. P. Eboli, E. M. Gregores, and F. Halzen, Nucl. Phys. Proc. Suppl. **99A**, 257 (2001).

[76] L. Frankfurt, M. Strikman, and C. Weiss, Annalen Phys. **13**, 665 (2004). hep-ph/0410307.

[77] L. Frankfurt, M. Strikman, C. Weiss, and M. Zhalov (2004). hep-ph/0412260.

[78] V. A. Petrov and R. A. Ryutin, Eur. Phys. J. **C36**, 509 (2004). hep-ph/0311024.

[79] A. Bialas, Acta Phys. Polon. **B33**, 2635 (2002). hep-ph/0205059.

[80] A. Bialas and R. Peschanski, Phys. Lett. **B575**, 30 (2003). hep-ph/0306133.

[81] M. Klasen and G. Kramer, Eur. Phys. J. **C38**, 93 (2004). hep-ph/0408203.

[82] M. Klasen and G. Kramer, Phys. Rev. Lett. **93**, 232002 (2004). hep-ph/0410105.

[83] A. B. Kaidalov, V. A. Khoze, A. D. Martin, and M. G. Ryskin, Phys. Lett. **B567**, 61 (2003). hep-ph/0306134.

[84] T. Sjöstrand and M. van Zijl, Phys. Rev. **D36**, 2019 (1987).

[85] T. Sjöstrand, and others, Comput. Phys. Commun. **135**, 238 (2001). hep-ph/0010017.

[86] G. Corcella *et al.*, JHEP **01**, 010 (2001). hep-ph/0011363.

[87] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, Z. Phys. **C72**, 637 (1996). hep-ph/9601371.

[88] J. Butterworth *et al.* http://jetweb.hep.ucl.ac.uk/JIMMY.

[89] C. Buttar *et al.*, *Underlying events*. These Proceedings.

[90] R. Field, *Min-bias and the underlying event at the tevatron and the lhc*. http://www.phys.ufl.edu/~rfield/cdf/FNAL_Workshop_10-4-02.pdf. Talk presented at the Fermilab ME/MC Tuning Workshop, October 4, 2002.

[91] B. Cox, J. Forshaw, and L. Lönnblad, JHEP **10**, 023 (1999). hep-ph/9908464.

[92] L. Lönnblad and M. Sjödahl, JHEP **05**, 038 (2005). hep-ph/0412111.

Multi-Jet Production and Multi-Scale QCD

Z. Czyczula[1,2], *G. Davatz*[3], *A. Nikitenko*[4], *E. Richter-Was*[1,5,*], *E. Rodrigues*[6], *N. Tuning*[6]
[1] Institute of Physics, Jagiellonian University, Krakow, Poland
[2] Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
[3] Institute for Particle Physics, ETH Zürich, Switzerland
[4] Imperial College, London, UK
[5] Institute of Nuclear Physics PAN, Krakow, Poland
[6] NIKHEF, Amsterdam, The Netherlands

Abstract
We summarize the contributions in Working Group II on "Multi-jet final states and energy flows" related to the topic of jet production, multi-jet topologies and multi-scale QCD. Different parton shower models will lead to systematic differences in the event topology. This may have a significant impact on predictions for the LHC. Here we will look at a few examples, such as the acceptance of $H \to \tau\tau$ events and in applying a jet veto in the non-hadronic $H \to WW \to l\nu l\nu$ decay channel. We also study the effect of CCFM evolution on the jet veto and on the event topology at the LHC in the forward region. Finally, we show that the choice of the QCD scale leads to large uncertainties in e.g. the $H \to \tau\tau$ analysis.

1 Introduction

In simulating high-energy interactions, the sequence of branchings such as $q \to qg$, can be modelled by calculating the exact amplitude of the Feynman diagrams, known as the matrix-element method, or, alternatively, can be modelled using the parton-shower approach. Matrix elements are in principle the exact approach but lead to increasingly complicated calculations in higher orders, and are therefore only used for specific exclusive physics applications, such as background estimates with multiple hard jets (see also [1]).

Since no exact way of treating partonic cascades exist, various Monte Carlo programs model the parton showers in different ways. In HERWIG [2] the parton showers are performed in the soft or collinear approximation, treating the soft gluon emission correctly. The shower is strictly angular ordered, where the angle between emitted partons is smaller at each branching. The hardest gluon emission is then matched to the first order matrix-element. This matrix-element correction has recently been implemented for $gg \to H$, leading to harder jets, and thus a more stringent jet veto in e.g. the non-hadronic decay $H \to WW \to l\nu l\nu$, where the jet veto is crucial to reduce the top background. PYTHIA [3] applies the collinear algorithm with the cascade ordered according to the virtuality Q^2. Corrections to the leading-log picture using an angular veto, lead to an angular ordering of subsequent emissions. The initial parton branchings are weighted to agree with matrix-elements. ARIADNE [4] on the other hand, does not emit gluons from single partons, but rather from the colour dipoles between two dipoles, thus automatically including the coherence effects approximated by angular ordering in HERWIG. From the resulting two dipoles softer emission occurs, resulting in a p_T ordering of subsequent emissions. ARIADNE has proven to predict the event shapes at HERA accurately [5], and could be explored more widely for simulation studies for the LHC.

* Supported in part by the Polish Government grant KBN 1 P03 091 27 (years 2004-2006) and by the EU grant MTKD-CT-2004-510126, in partnership with the CERN Physics Department.

The way parton showers are implemented affects the emission of soft gluons, and therefore affect both the transverse momentum of the produced Higgs, as well as the p_T of the balancing jets. In the following we will discuss the effect of the different parton showers on the selection of $H \to \tau\tau$ by applying angular cuts on the jets and on the selection of $H \to WW \to l\nu l\nu$ by rejecting events with jets with large p_T.

Both PYTHIA and HERWIG are general purpose leading order (LO) parton shower Monte Carlo programs, based on LO matrix elements. MC@NLO [6] on the other hand, uses exact next-to-leading order (NLO) calculations and is matched to the HERWIG parton shower Monte Carlo. Its total cross section is normalized to NLO predictions. The different predictions of these programs for the high part of the transverse momentum spectrum of the Higgs will be described in detail.

In the parton cascade as implemented in e.g. PYTHIA, the parton emissions are calculated using the DGLAP approach [7], with the partons ordered in virtuality. DGLAP accurately describes high-energy collisions of particles at moderate values of the Bjorken-x by resummation of the leading log terms of transverse momenta ($(\alpha_s \ln Q^2)^n$). However, to fixed order, the QCD scale used in the ladder is not uniquely defined. Different choices of the scale lead to large differences in the average transverse momentum of the Higgs in e.g. the processes $gb \to bH$ and $gg \to bbH$.

In the CCFM formalism [8] there is no strict ordering along the parton ladder in transverse energy, contrary to the DGLAP formalism. The CASCADE Monte Carlo program [9] has implemented the CCFM formalism, inspired by the low-x F_2 data and forward jet data from HERA, and became recently available for pp scattering processes. Until now, CASCADE only includes gluon chains in the initial state cascade. Different sets of unintegrated gluon densities are available, which all describe HERA data equally well [9]. Note, however, that it is questionable if these densities are constrained enough for Higgs production, as discussed elsewhere in these proceedings [10].

CCFM is expected to provide a better description of the gluon evolution at very low values of x compared to DGLAP, as it also takes leading-logs of longitudinal momenta ($(\alpha_s \ln x)^n$) into account. Since the partons at the bottom of the ladder (furthest away from the hard scatter) are closest in rapidity to the outgoing proton, effects might be expected in the forward region. The event topology in terms of jets and charged multiplicity is investigated at rapidities $2<\eta<5$, corresponding to the acceptance of the LHCb detector.

2 MSSM Higgs production with the Yukawa bbH coupling induced mechanisms

In the MSSM, the Yukawa coupling of the heavy neutral Higgs bosons to the bottom quarks is strongly enhanced for large tan(β) with respect its SM value, which makes the Higgs boson production in association with bottom quarks the primary production mechanism in LHC pp collisions. Currently, the inclusive cross section for this process is under good control up to NNLO, both in the so called fixed-flavour-scheme (FFS) and varying-flavour-scheme (VFS). The impressive level of theoretical uncertainty in the order of 15% is achieved on the predictions for the total cross-section for m_H=120 GeV [11, 12].

The observability potential for the $H \to \tau\tau$ channel [13] is, however, very sensitive to the topology of the events, due to the reconstruction of the invariant mass of the tau-pair, using the collinear approximation of τ-leptons decay, in order to account for the neutrino momenta. The impact of the event topology on the final acceptance of the signal has been discussed elsewhere [14]. Here, we pursue the subject further and we study more quantitatively the systematic effects from the parton shower model and the choice of the QCD scale selected in the event generation.

Currently available Monte Carlo generators for the Higgs boson production are based on the LO matrix elements, with the QCD part of physics event simulated with a parton shower approach. Clearly, the kinematics of the Higgs boson (and therefore the final acceptance for the signal) depends strongly on the algorithm used to simulate the QCD cascade. At tree level, the following exclusive processes have been studied, combining the observability of events with and without spectator b-tagged jets accompa-

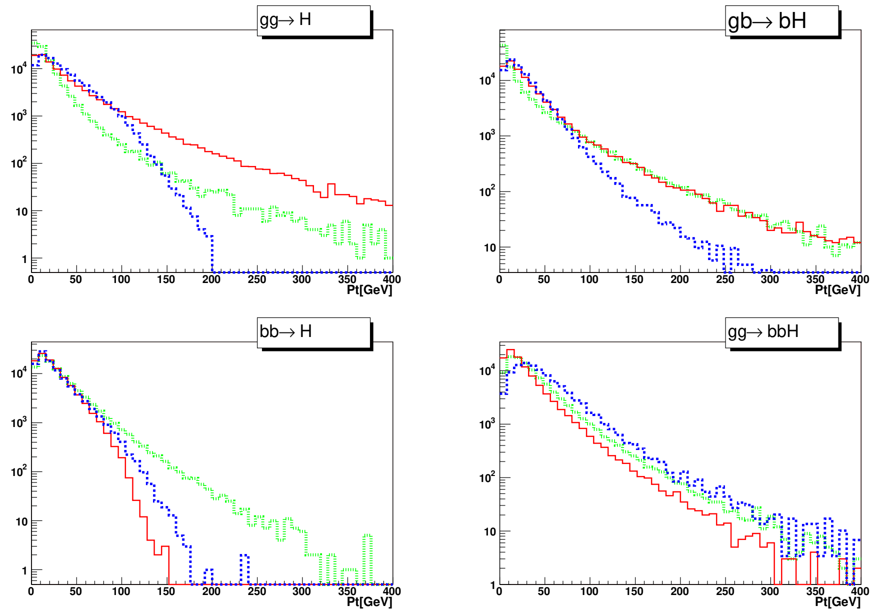


Fig. 1: The transverse momenta of the Higgs boson, p_T^{Higgs} for 3 different shower models for each production mechanism. The red solid line represents PYTHIA, the dashed green line ARIADNE and the dotted blue line HERWIG events. The vertical scale gives the number of events per bin, and a total of 10^5 events have been generated with each program.

nying the reconstructed tau-pair: $gb \to bH$ (VFS), $gg \to b\bar{b}H$ (FFS), $b\bar{b} \to H$ (VFS) and $gg \to H$.

For the purpose of the discussion presented here we have studied the SM Higgs boson production with a mass of 120 GeV, decaying into a tau pair, where one tau decays hadronically and one leptonically. The reconstruction of the Higgs boson mass and the selection criteria were performed at the level of generated particles (leptons, hadrons) or, where necessary (missing energy, b-jets), on objects reconstructed from simplified simulation of the detector response [15].

2.1 Systematics from the choice of parton shower model

As discussed in the introduction, the various parton shower models predict different spectra of the transverse momentum, p_T^{Higgs}, of the produced Higgs boson. This leads to a large variation in the prediction for the fraction of accepted events. The obvious starting point for the discussion is the Higgs boson transverse momentum spectra in complete physics events [1]. In case of the 2→2 and 2→3 processes, the p_T of the Higgs boson arrises predominantly from matrix elements, whereas in the 2→1 events p_T^{Higgs} purely comes from the parton shower. Therefore, the Higgs transverse momentum spectra differ significantly for different models of the QCD cascade. Figure 1 shows these spectra for each production mechanism [2].

Clearly, the spectra of the Higgs boson transverse momenta show substantial dependence not only on the topology of the hard process, but also on the shower model used in the simulation of the event. The shower model as implemented in PYTHIA includes hard matrix element corrections for inclusive gluon-gluon fusion, $gg \to H$, hence leading to a harder spectrum compared to the one obtained from the standard HERWIG shower. In this production mode the shower model from ARIADNE fails because of the missing splitting kernel for $g \to q\bar{q}$. On the other hand, the ARIADNE model predicts the hardest

[1] The AcerMC 2.4 framework [16] with interfaces to PYTHIA 6.2, ARIADNE 4.12 and HERWIG 6.5 was used to generate events and AcerDET [15] was used to simulate the detector performance.

[2] The CTEQ5L parton density functions were used in all simulations. It has been checked that both final acceptance of the signal and the mean Higgs boson transverse momentum is almost independent of the pdf parametrization. Uncertainties below 10% are observed by using CTEQ5L, CTEQ6L, MRST2001 interfaced with LHAPDF [17]).

spectra for the process $b\bar{b} \to H$. In this production channel, predictions from PYTHIA and HERWIG are in quite good agreement. However, almost a factor of two difference for the prediction of the mean transverse momenta can be reported between PYTHIA and HERWIG in $gg \to b\bar{b}H$ process.

Numerical values for the average Higgs boson transverse momentum in different production processes and parton shower models are given in Table 1. It is important to stress that these results were obtained with default settings of the parameters for each parton shower model.

The steps of the analysis that lead to the reconstruction of the tau-pair invariant mass are indicated in Table 1, including the acceptances for all the discussed production processes and parton shower models. They consist of the basic selection (including the trigger and p_T and $|\eta|$ cuts on the lepton and jet), and the additional selection that is needed to improve the mass resolution of the accepted tau-pair. The acceptance of the signal after the basic selection is rather stable, at the level of 12%-14% depending on the production mechanism. The significant differences start to appear when a cut on the angle between the lepton and hadron is applied. A difference of almost a factor two is observed for the $b\bar{b} \to H$ production process with the parton shower from the HERWIG or ARIADNE model, respectively.

For the final acceptance values, the uncertainty from the parton shower model varies between 85% for inclusive gluon fusion to 135% for $gg \to bbH$ (between HERWIG and PYTHIA models). In the case of the Higgs production through $bb \to H$, predictions from HERWIG and PYTHIA models are in excellent agreement. However, the prediction of the acceptance in this production channel differs by 115% if the parton shower from ARIADNE is used. For the $gb \to bH$ production mechanism, the uncertainty due to the shower model from either PYTHIA or HERWIG is about 90%.

The systematic theoretical uncertainty on the predictions for the final acceptance ranges from 85% to 135% for the three different shower models studied here. The uncertainty is even larger, when the requirement of an additional tagging b-jet is introduced, up to 170% for $bb \to H$ [3]. Figure 2 shows the

[3]It should be stressed, that the problem of the efficiency of b-jet tagging was not touched upon, nor was the problem of

Hard process	$\mathbf{gg \to H}$			$\mathbf{b\bar{b} \to H}$		
Shower model	PY	AR	HW	PY	AR	HW
$< p_T^{Higgs}$ (generated)$>$ (GeV)	37.2	X	32.2	23.1	29.9	24.6
$< p_T^{Higgs}$ (accepted)$>$ (GeV)	129.4	X	75.27	58.6	91.64	68.4
basic selection	14.2%	X	12.7%	12.8%	13.8%	11.8%
+($cos(\phi)$ >-0.9 , $\|sin(\phi)\|$ >0.2)	5.5%	X	4.5%	2.9%	4.3%	2.7%
+($p_T^{miss} >$ 30 GeV, $m_T^{lep-miss}$ <50GeV)	3.8%	X	2.3%	1.4%	2.3%	1.5%
+(mass window: 120 $\pm$ 20 GeV)	2.4%	X	1.3%	0.6%	1.3%	0.6%
+(1 tagged b-jet)				0.4%	1.0%	0.4%
Hard process	$\mathbf{gb \to bH}$			$\mathbf{gg \to b\bar{b}H}$		
Shower model	PY	AR	HW	PY	AR	HW
$< p_T^{Higgs}$ (generated)$>$ [GeV]	32.5	26.0	26.9	27.2	35.8	47.4
$< p_T^{Higgs}$ (accepted)$>$ [GeV]	125.1	133.9	82.1	95.0	99.6	105.3
basic selection	13.3%	12.6%	11.7%	13.0%	13.6%	12.1%
+($cos(\phi) > -0.9, \|sin(\phi)\|$ >0.2)	4.4%	3.4%	3.2%	3.5%	5.1%	6.7%
+($p_T^{miss} >$ 30 GeV,$m_T^{lep-miss}$ <50GeV)	2.7%	2.4%	1.7%	2.0%	2.9%	3.8%
+(mass window: 120 $\pm$ 20 GeV)	1.7%	1.5%	0.9%	1.1%	1.8%	2.6%
+(1 tagged b-jet)	1.3%	1.4%	0.6%	0.9%	1.2%	2.1%

Table 1: The average transverse momenta of the Higgs boson and acceptance of selection criteria for different hard processes and parton shower models. Events were generated with default initialization of these generators. Columns marked PY, AR and HW denote results from PYTHIA, ARIADNE and HERWIG shower model respectively.

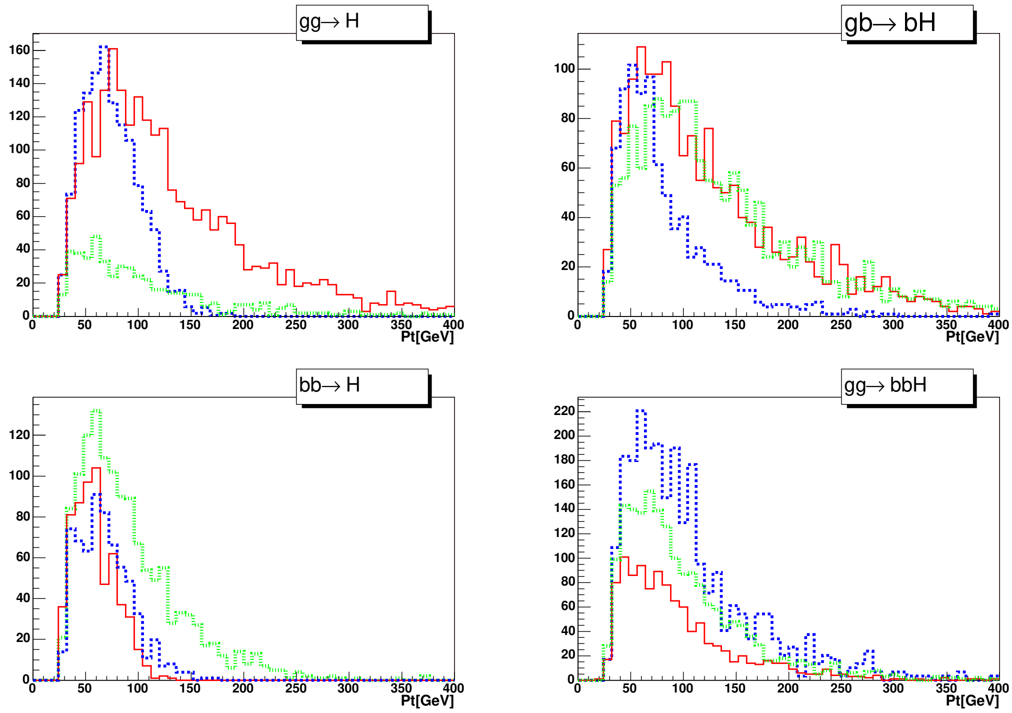


Fig. 2: Same as Fig. 1 but after selection presented in Table 1. The vertical scale in in number of events entering given bin after selection procedure, in each case 10^5 events were initially generated.

Higgs boson transverse momentum for those events that passed all selection criteria. As can be observed, the selection criteria rejected most of events with $p_T^{Higgs} < 40$ GeV.

2.2 Systematics from the choice of QCD scale

Having considered here the available Monte Carlo generators with the overall precision of the leading order only, large uncertainties are expected for the predictions coming from different scale choices. Here we concentrate only on the effects on the event topology, neglecting the effects from the choice of the QCD scale on the total cross-section. Table 2 shows the Higgs boson mean transverse momentum and final acceptance of the signal for 2→2 and 2→3 processes for some possible choices in PYTHIA and ARIADNE. The Q^2 value sets the scale not only for the hard scattering process, but also for the initial state parton shower. For the 2→1 production, the Q^2 scale is naturally set to be the mass of the Higgs boson mass. The uncertainty in the acceptance due to scale choice for the $gg \to b\bar{b}H$ production mechanism is about 60% in the case of PYTHIA and 25% in the case of ARIADNE parton shower model. For the exclusive process $gb \to bH$, the uncertainties are 75% and 100%, respectively.

3 gg → H at the LHC: Uncertainty due to a Jet Veto

In the Higgs mass range between 155 and 180 GeV, $H \to W^+W^- \to \ell\nu\ell\nu$ is considered to be the main Higgs discovery channel [18, 19]. The signal consists of two isolated leptons with large missing E_T and with a small opening angle in the plane transverse to the beam, due to spin correlations of the W-pair. In order to reduce the top background, a jet veto has to be applied. The signal over background ratio is found to be around 2:1 for Higgs masses around 165 GeV. For lower and higher Higgs masses, the signal over background ratio decreases slightly [19]. The experimental cross section σ_{meas} of the Higgs signal

the efficiency for the reconstruction of the τ-jet. Discussing these effects, very important for complete experimental analysis, would complicate the problem and dilute the aim of the phenomenological studies presented here.

Hard process	$\mathbf{gb \to bH}$			$\mathbf{gg \to b\bar{b}H}$			
Q^2 scale	`default`	$\hat{s}$	$\frac{2\hat{s}\hat{t}\hat{u}}{\hat{s}^2+\hat{t}^2+\hat{u}^2}$	`default`	m_b^2	m_b^2	$\hat{s}$
<Q> (GeV)	94	257	49	27	4.8	120	255
$< p_T^{Higgs}$ (generated)$>$ (GeV)[`PY`]	32.5	42.7	43.2	27.2	29.8	32.1	36.2
Acceptance (%) [`PY`]	1.7	2.6	2.96	1.1	1.3	1.4	1.8
$< p_T^{Higgs}$ (generated)$>$ (GeV)[`AR`]	26.0	25.5	44.9	35.8	38.	35.3	34.5
Acceptance (%) [`AR`]	1.5	1.6	3.1	1.8	2.1	1.7	1.7

Table 2: The average transverse momenta of the Higgs boson and acceptance of selection criteria for different scale choices. Events were generated with default initialization of these generators. Events marked `PY` and `AR` denote results from PYTHIA and ARIADNE shower model respectively.

and other final states is given by:

$$\sigma_{meas} = N_s/(\epsilon_{sel} \times L_{pp}), \tag{1}$$

with N_s being the number of signal events, ϵ_{sel} the efficiency after all signal selection cuts are applied and L_{pp} the proton-proton luminosity. In order to get an estimate of the cross section uncertainty, the statistical and systematic uncertainties have to be determined. The systematic uncertainties come from the experimental selection, background and luminosity uncertainties. As the signal over background ratio is small in the channel under study, the systematic uncertainties should be known precisely. This study concentrates on the uncertainty of the signal efficiency due to the jet veto, by studying the systematics using different Monte Carlo simulations. To do so, four different parton-shower Monte Carlo programs were used, as described in the introduction. The effect of different parton shower models are discussed by comparing PYTHIA 6.225 [3] and HERWIG 6.505 [2], whereas the comparison to MC@NLO 2.31 [6] leads to an uncertainty estimate of higher-order effects [4]. Then, also CASCADE 1.2009 [9] is studied to compare the DGLAP approach to the CCFM formalism.

Jets are reconstructed using an iterative cone algorithm with a cone size of 0.5. The leading particle (seed) of the jet is required to have a p_T larger than 1 GeV. The pseudo-rapidity $|\eta|$ of the jet should be smaller than 4.5, corresponding to the CMS detector acceptance [20]. The event is rejected if it contains a jet with a p_T higher than 30 GeV. The Higgs mass for this study was chosen to be 165 GeV, corresponding to the region of phase space with the highest signal over background ratio. First, all events are studied without considering the underlying event. Finally, PYTHIA is also studied including different underlying event schemes.

3.1 Matrix Element Corrections

At leading order, the transverse momentum of the Higgs boson, p_T^{Higgs}, is zero. However, parton shower Monte Carlos emit soft gluons which balance the Higgs and introduce a transverse momentum in LO parton shower Monte Carlos. As the Higgs is balanced by jets, the transverse momentum is very sensitive to the jet veto and therefore also the efficiency of a jet veto dependends stongly on p_T^{Higgs}.

In Fig. 3, the normalized p_T^{Higgs} spectra are shown for PYTHIA, HERWIG and MC@NLO. HERWIG and MC@NLO are very similar at low p_T, as can be seen on the linear scale, which is to be expected as the soft and collinear emissions of MC@NLO are treated by HERWIG. Figure 4 shows that PYTHIA predicts a softer leading jet spectrum than HERWIG and therefore also a softer p_T^{Higgs} spectrum. HERWIG implements angular ordering exactly and thus correctly sums the LL (Leading Log) and part of the N^kLL (Next-to..Leading Log) contributions. However, the current version of HERWIG available does not treat hard radiations in a consistent way. Hence the spectrum drops quickly at high p_T, see Fig. 3b). PYTHIA on the other hand does not treat angular ordering in an exact way, but includes hard matrix ele-

[4] In the following, HERWIG and PYTHIA use the pdf-set CTEQ5L, whereas MC@NLO uses CTEQ5M.

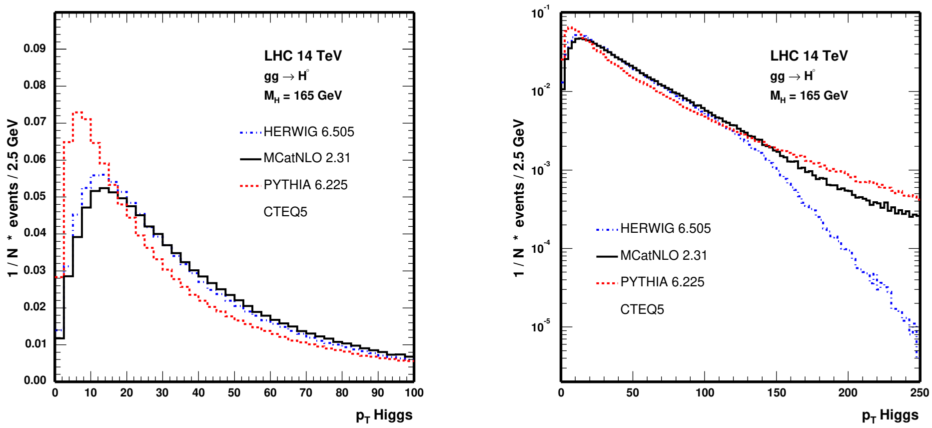


Fig. 3: p_T^{Higgs} spectra for PYTHIA, HERWIG and MC@NLO in linear and logarithmic scale.

ment corrections. Therefore PYTHIA looks more similar to MC@NLO at high p_T. MC@NLO correctly treats the hard radiation up to NLO, combining the high p_T spectrum with the soft radiation of HERWIG.

In Fig. 5, the efficiency of the jet veto is shown for the three different Monte Carlos as a function of p_T^{Higgs}. One observes a strong dependency of p_T^{Higgs} on the jet veto. Once a jet veto is defined, the efficiency starts to drop quickly as soon as p_T^{Higgs} is close to the p_T used to define a jet veto. However, as the transverse momentum of the Higgs can be balanced by more than one jet, the efficiency is not zero above this value.

G. Corcella provided a preliminary version of HERWIG including hard matrix element corrections for $gg \to H$ [21]. The hard matrix element corrections lead to harder jets, see Fig. 6, and therefore the jet veto is more effective. At high p_T, PYTHIA and HERWIG now show very similar predictions. Table 3 shows the efficiencies of the jet veto of 30 GeV for MC@NLO, PYTHIA and HERWIG with and without matrix element corrections. In addition, the numbers for CASCADE are shown, which will be discussed

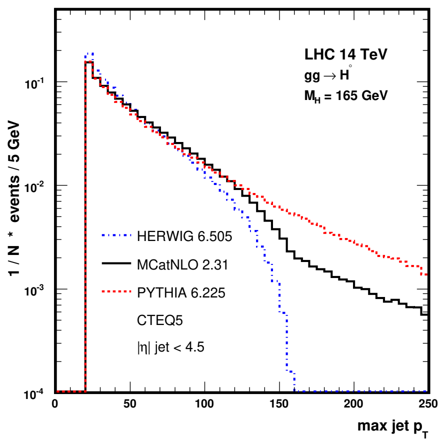


Fig. 4: p_T of the leading jet for PYTHIA, HERWIG and MC@NLO

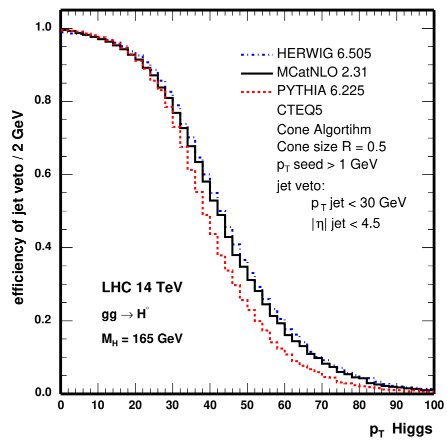


Fig. 5: Efficiency of the jet veto of 30 GeV as a function of p_T^{Higgs}.

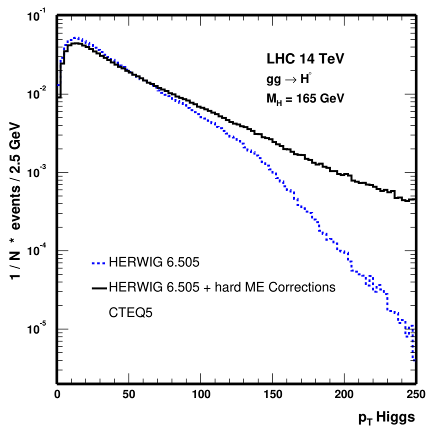


Fig. 6: HERWIG with and without hard Matrix Element Corrections, logarithmic scale.

in more detail later. In the first row, the number of the efficiency for p_T^{Higgs} between 0 and 80 GeV is shown. The second column shows the inclusive efficiency for all events. One has to keep in mind that after all selection cuts, only the low p_T region is important [19].

	Efficiency for events with a p_T Higgs between 0 and 80 GeV	Inclusive efficiency (all events)
MC@NLO 2.31	0.69	0.58
PYTHIA 6.225	0.73	0.62
HERWIG 6.505	0.70	0.63
HERWIG 6.505 + ME Corrections	0.68	0.54
CASCADE 1.2009	0.65	0.55

Table 3: Efficiency of the jet veto for MC@NLO, PYTHIA, HERWIG, HERWIG + ME Corrections and CASCADE.

In order to estimate the effect from the detector resolution on the jet veto, the E_T of the jet is smeared with the jet resolution of e.g. CMS, as given by [20]:

$$\Delta E_T/E_T = 118\%/\sqrt{E_T} + 7\%. \tag{2}$$

More jets at initially low p_T are shifted to higher p_T than vice versa, as the jets are generally soft. However, the effect of the smearing is limited and the difference between the smeared and unsmeared case is smaller than 1%.

In the last years, a lot of progress has been made in understanding the Higgs boson production and decays on a theoretical basis. The gluon fusion cross section has been calculated up to NNLO [22]. Such corrections are known to increase the LO cross section by a factor of more than two. In order to include these higher order corrections in a parton shower Monte Carlo, each event is reweighted with its corresponding p_T-dependent effective K-factor (which includes all selection cuts) [19]. This technique can be applied to other processes which are sensitive to jet activity, e.g. the WW background for this channel. The result is an overall effective K-factor of 2.04 for a Higgs mass of 165 GeV, which is only about 15% lower than the inclusive K-factor (without any cuts) for the same mass. This reweighting method allows to optimize the selection cuts and thus also helps to improve the discovery potential. We observe that the uncertaintiy of the jet veto efficiency does not change significantly including those higher order corrections.

3.2 Underlying Event

So far all events were generated without considering the underlying event. However, to study a jet veto, it is important to consider also the effect of the underlying event. Therefore, PYTHIA was studied with different underlying event tuning schemes, which are the ATLAS Tune [23], CDF Tune A [24] and PYTHIA default (MSTP(81)=1, MSTP(82)=3 [3]). The different tunings lead to approximately the same efficiency, and also the difference in the efficiency with and without underlying event is smaller than 1%, see Table 4.

	Efficiency for events with a p_T Higgs between 0 and 80 GeV	Inclusive efficiency (all events)
PYTHIA no UE	0.730	0.620
PYTHIA default	0.723	0.613
ATLAS tune	0.706	0.600
CDF tune	0.709	0.596

Table 4: Efficiency numbers for different underlying event tunings in PYTHIA.

3.3 Comparing to CCFM evolution

Finally, we compared the PYTHIA, HERWIG and MC@NLO predictions with the ones obtained using CASCADE. One has to keep in mind that this Monte Carlo is dedicated to low-x physics, and is about to be released for LHC physics applications. There were many improvements implemented during this workshop. In Fig. 7, the p_T^{Higgs} spectra for PYTHIA, HERWIG+ME Corrections, MC@NLO and CASCADE are shown. The prediction from CASCADE lies within the ones from PYTHIA and HERWIG. When looking at different p_T regions, one generally observes that CASCADE produces more jets compared to the other Monte Carlos, and the jets are harder. The jet veto efficiency as a function of the p_T of the Higgs is shown in Fig. 8, indicating that the main differences are in the low p_T range and that the efficiency for CASCADE is slightly smaller than unity at a p_T^{Higgs} of zero. A reason for this is that the Higgs boson is balanced by more than one jet, with at least one of the jets with a p_T higher than 30 GeV and thus vetoed. For the same reason, the efficiency in general is lower than for the other Monte Carlo programs at low p_T^{Higgs}. Results in the high p_T region have to be studied carefully.

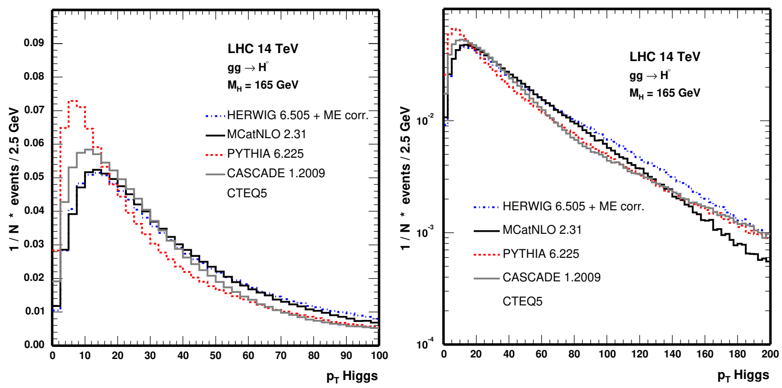


Fig. 7: p_T^{Higgs} Higgs of PYTHIA, HERWIG + ME Corrections, MC@NLO and CASCADE, linear and logarithmic scale.

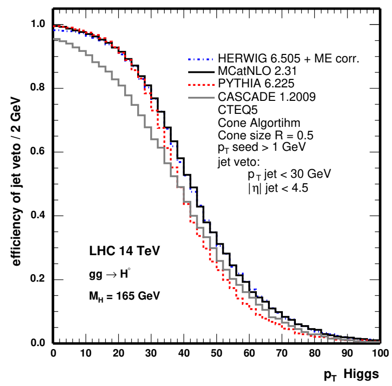


Fig. 8: Efficiency as a function of p_T for PYTHIA, HERWIG+ME Corrections, MC@NLO and CASCADE.

4 Forward Studies with CASCADE at LHC Energies

The applicability of DGLAP evolution [7] is known to be limited in the very forward region, that is at small values of Bjorken-x, where $ln(x)$ terms are expected to become large [25]. Since the partons at the bottom of the ladder (furthest away from the hard scatter) are closest in rapidity to the outgoing proton, effects might be expected in the forward region. The CCFM evolution [8] takes these BFKL-like terms into account, and is implemented in the CASCADE Monte Carlo program [9].

We have studied the topology of forward particle and jet production in the LHCb detector at the LHC. LHCb is a forward spectrometer covering roughly the forward region 1.8$<\eta<$4.9 [26]. Its main goal is the study of CP violation in the B-meson sector and the measurement of rare B-decays. But its very nature makes LHCb a suitable environment for QCD forward studies.

The usage of another Monte Carlo program in LHCb is important in order to estimate the uncertainty on the PYTHIA [3] predictions. In particular, the track multiplicity seen in the detector is an important factor to take into account, as it affects the performance of the trigger, the tracking and the B-tagging. But here we will concentrate on another aspect: the study of the QCD evolution itself, proving that LHCb has the potential to be a natural test bed of QCD in the forward region, complementing the studies done at present at the Tevatron and the future studies to be made with the central detectors – ATLAS and CMS – at the LHC. The predictions in the forward region as given by CASCADE are here compared with that of PYTHIA, the default Monte Carlo generator used in LHCb. This is a "natural" way to test CCFM versus DGLAP QCD evolution in the region of the phase space where differences are most likely to show.

In what follows we will compare both predictions for the event kinematics and topology, and the particle and jet production. We used CASCADE version 1.2009 "out of the box" and PYTHIA 6.227 with the LHCb tune. We used for the comparisons a sub-sample of the QCD processes of PYTHIA, as CASCADE only includes (unintegrated) gluons. PYTHIA was run with the only sub-processes $fg \to fg$, $gg \to f\bar{f}$ and $gg \to gg$, and multiple interactions (MI) were also switched off, since they are as yet not implemented in CASCADE; this version is denoted *"PYTHIA gluon"* in the plots. Another configuration named *"PYTHIA gluon incl MI"* has the multiple interactions switched on, for a cross-check of the influence of such inclusion. All the plots refer to minimum bias events.

4.1 Event Kinematics

Figure 9 shows the kinematic variables Q^2 and Bjorken-x variables x_1 and x_2 (referring to both LHC proton beams of energy E_p), using the definitions given below. For PYTHIA the standard definitions

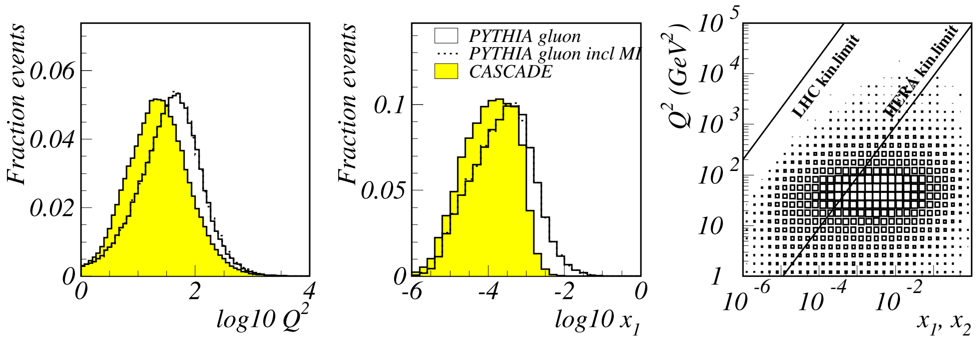


Fig. 9: Comparison between CASCADE and PYTHIA for the general event kinematics variables (refer to the text for the definitions). Note that $x_1 < x_2$ by construction.

from the PYPARS common block were used:

$$x_1 = \text{PARI(33)} \quad x_2 = \text{PARI(34)};$$
$$Q^2 = \text{PARI(18)},$$

whereas for CASCADE we set [5]:

$$x_{1,2} = \frac{(E + |p_z|)_{in.\ parton\ 1,2}}{2E_p};$$
$$Q^2 = p^2_{T\ out.\ parton}.$$

There is a reasonable agreement between both Monte Carlo programs, although a direct comparison seems difficult and unnatural given the definitions above. The phase space spanned by the kinematic variables $x_{1,2}$ and Q^2 is shown also in Fig. 9 for PYTHIA.

4.2 Forward Particle Production

Some general event variables are compared in Fig. 10 in the region of the LHCb acceptance, 1.8<η<4.9, including the charged track multiplicity, the acoplanarity ($\Delta\phi$) of the outgoing partons, the average track transverse momentum in the event <p_T> and the maximum track transverse momentum $p_{T,max}$. The predictions from both Monte Carlo programs agree well – neglecting the multiple interactions in PYTHIA – likely because the same final state parton showering is performed. The effect of including the multiple interactions is seen mainly in the event multiplicity, as expected. Interesting is the distribution of the acoplanarity of the two outgoing partons: PYTHIA predicts a strong (anti-)correlation whereas CASCADE exibits a distribution that is nearly flat.

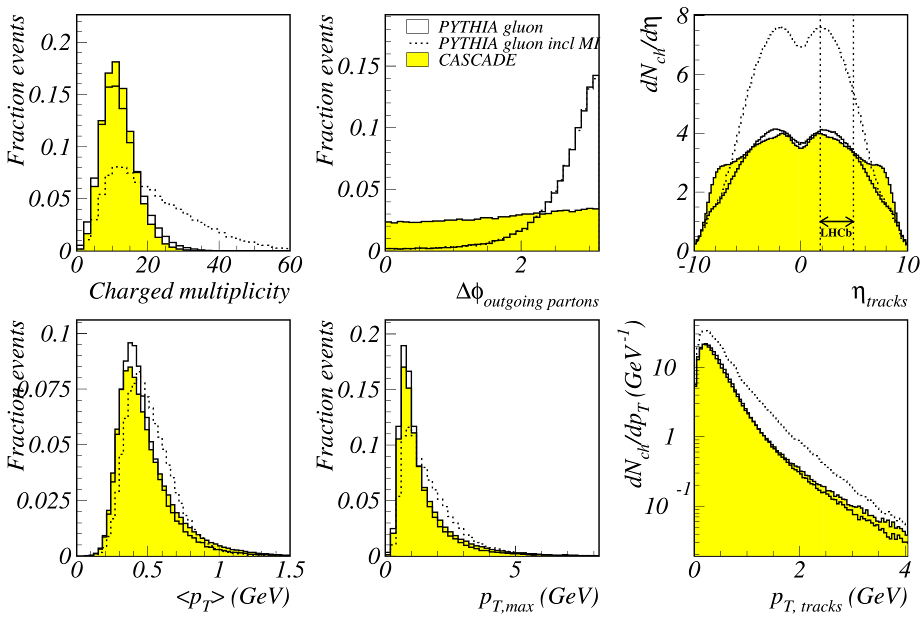


Fig. 10: Comparison between CASCADE and PYTHIA for general event variables, and between charged tracks variables in the region of the LHCb acceptance defined as $1.8 < \eta < 4.9$. No acceptance cuts are applied on the $\Delta\phi$ and η_{tracks} distributions.

The number of charged tracks per unit rapidity, $dN/d\eta_{tracks}$, and the differential distribution of the number of charged tracks (in the LHCb acceptance) as a function of the transverse momentum $p_{T,tracks}$ are also included in Fig. 10. Note that these 2 distributions were normalized to the mean

[5]The two incoming partons in the hard interaction are obtained from the variables NIA1 and NIA2, corresponding to the positions 4 and 6 in the CASCADE event record, whereas the outgoing partons are at positions 7 and 8.

track multiplicity in the full and LHCb acceptance, respectively. The p_T distributions compare very well, leading us to conclude that the general hard dynamics of the event is predicted in a rather similar way by both programs. CASCADE however, produces more forward tracks than PYTHIA, as the η-distribution is clearly flatter than the rather steep distribution of PYTHIA. This is particularly true in the region $5<\eta<8$, just beyond the acceptance of the LHCb spectrometer – shown between the 2 vertical dashed lines –, but could make LHCb a candidate environment to discriminate between the two predicted forward behaviours.

4.3 Forward Jet Production

We have also looked at jet production. Jets were found in the laboratory frame with the KTCLUS algorithm on all stable hadrons, in the longitudinally invariant inclusive mode. We looked at the jet production in the LHCb acceptance with a rather loose selection of $E_{T,jets} > 1$ GeV. The number of jets found in PYTHIA or CASCADE is shown in Fig. 11. The number of events with no jets satisfying $E_{T,jet} > 1$ GeV inside $1.8<\eta<4.9$ is much larger for PYTHIA. In other words, CASCADE predicts a jet cross-section larger than PYTHIA, a fact already shown by the HERA experiments in low-x jet analyses. This difference leads us to believe that strong angular ordering in CASCADE favours a "clustered production" of particles and therefore the production of jets, whereas PYTHIA tends to give a more spreaded transverse energy flow. Furthermore, though the effect is small, we already saw from Fig. 10 that the highest-p_T track is somewhat softer in PYTHIA compared to CASCADE.

The rapidity distribution and the transverse energy distribution of the jets is also shown in Fig. 11; they have been normalized to the average number of jets per event in the full acceptance and LHCb acceptance, respectively. PYTHIA and CASCADE predict similar jets in the LHCb acceptance, but the inclusion of multiple interactions gives a harder spectrum.

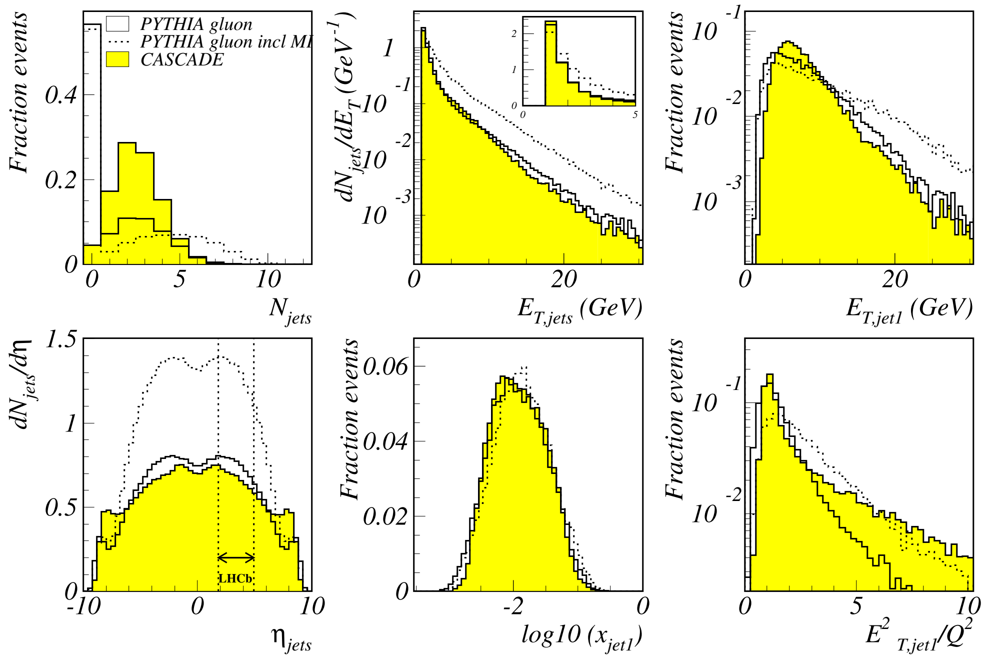


Fig. 11: Distributions of the number of jets N_{jets}, the jets transverse energy $E_{T,jets}$, the E_T and x_{jet} of the highest-E_T jet, $jet1$, all in the LHCb acceptance. Also the number of jets per unit pseudorapidity is shown. The distribution of the ratio of $E^2_{T,jet}/Q^2$ in the LHCb acceptance shows a comparison of the two scales. Jets were selected with $E_{T,jet} > 1$ GeV.

Also shown are the event distributions in the LHCb acceptance of the highest-E_T jet in the event, $E_{T,jet1}$, and the energy fraction of the proton carried by the highest-E_T jet, $x_{jet1} = E_{jet1}/E_p$. The hardest jet in the event is on average harder in CASCADE compared to PYTHIA. The distributions of x_{jet} and $E^2_{T,jet}/Q^2$ are interesting in that they correspond to variables now in standard use within the HERA experiments as a means of selecting samples where forward effects are expected. Indeed both experiments have published a series of "forward QCD" analyses [25] applying cuts of the kind $E^2_{T,jet} \sim Q^2$ and $x_{jet} \gg x_{Bjorken}$. The phase space is selected such that it suppresses jet production via DGLAP evolution and enhances production from BFKL dynamics:

- DGLAP evolution is suppressed in the small phase space for Q^2 evolution requiring $E^2_{T,jet} \sim Q^2$;
- CCFM evolution enhanced when large phase space for x evolution requiring $x_{jet} \gg x_{Bjorken}$.

At the LHC such a selection becomes rather delicate, since there are two proton beams and the comparison of x_{jet} with $x_{Bjorken}$ gets an ambiguity between the choice of x_1 or x_2. A way out – though it lowers significantly the statistics – would be to make the selection based on $x_{jet} \gg max(x_1, x_2)$. From the distributions presented in this paper we are lead to believe that such a forward selection is indeed possible. But we leave this issue open for further investigation.

5 Summary

Various ways of treating parton showers have been compared, as implemented by the HERWIG, PYTHIA and ARIADNE Monte Carlo programs. We have studied the uncertainties that arrise from these different models to the p_T-spectrum of the jets, and the p_T-spectrum of the Higgs boson.

The theoretical systematic uncertainty on predictions for inclusive cross section at NNLO for Higgs production with bbH Yukawa coupling is under good theoretical control with an uncertainty of about 15% for a Higgs mass around 120 GeV. However, the predictions for the exclusive cross section determined by the event selection of a simplified experimental analysis indicates at present an order by magnitude larger uncertainty in e.g. $H \to \tau\tau$ events. Uncertainties due to the shower model can reach 170% and depend strongly on the production mechanism. Another factor of two arises from the choice of the QCD scale. Higher order Monte Carlo generators will therefore be mandatory to achieve better precision on the theoretical predictions.

On the other hand, the uncertainty of the jet veto efficiency in the $H \to WW \to l\nu l\nu$ decay channel by using different Monte Carlo generators in the $gg \to H$ process is estimated to be around 10%. Including higher order QCD corrections does not enhance the uncertainty significantly. Also the effect of including a realistic jet-E_T resolution is very small. The effect of including an underlying event in the simulation is smaller than 1%, and does not vary significantly for various tuning models.

Furthermore we have studied the predictions at the LHC using the CCFM formalism as implemented in the full hadron level Monte Carlo generator. We conclude that CASCADE produces more and harder jets compared to the other Monte Carlo programs, leading to a bigger uncertainty of the jet veto efficiency in the small p_T^{Higgs} range. In the forward region larger differences are expected between the DGLAP and CCFM approach, but in the moderate forward rapidity range $2<\eta<5$, as covered by the LHCb detector, a fairly good agreement between CASCADE and PYTHIA is observed for most of the distributions looked at, and despite their different philosophies. However, this result has to be treated with care, as the program is only recently developped for proton physics at such high energies as produced in the future LHC. It also comes out of this simple study that CASCADE is indeed a potential Monte Carlo tool to use for QCD studies at the LHC in the forward region. In the future one should further investigate regions of phase space where large differences in behaviour are expected at the LHC from DGLAP and BFKL dynamics. LHCb seems a natural experimental environment in which to study such differences.

Finally, we would like to encourage the community by stating that it is very interesting and instructive to study the predictions at the LHC by using tools developed and tuned at HERA, such as the CCFM Monte Carlo CASCADE, and by using parton shower models such as ARIADNE, that have proven their validity at HERA.

References

[1] S. Höche *et al.*, *Matching parton showers and matrix elements*. These proceedings.

[2] G. Corcella *et al.*, JHEP **0101**, 010 (2001).

[3] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[4] L. Lönnblad, Comput. Phys. Commun. **71**, 15 (1992).

[5] ZEUZ Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C27**, 531 (2003).

[6] S. Frixione and B. Webber, JHEP **0206**, 029 (2002).

[7] V. Gribow and L. Lipatov, Sov. J. Nucl. Phys. **15**, 438 and 675 (1971);
G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977);
Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[8] M. Ciafaloni, Nucl. Phys. **B296**, 49 (1988);
S. Catani, F. Fiorani, and G. Marchesini, Nucl. Phys. **B336**, 18 (1990).

[9] H. Jung, Comput. Phys. Commun. **143**, 100 (2002).

[10] J. Collins *et al.*, *Unintegrated parton density functions*. These proceedings.

[11] R. Harlander and W. Kilgore, Phys. Rev. **D68**, 013001 (2003).

[12] J. Campbell, R. K. Ellis, F. Maltoni, and S. Willenbrock, Phys. Rev. **D67**, 095002 (2003). `hep-ph/0204093`.

[13] ATLAS Collaboration, *Detector performance and physics potential TDR*, 1999. CERN-LHCC-99-15, vol.II, ch.19.

[14] E. Richter-Was, T. Szymocha, and Z. Was, Phys. Lett. **B589**, 125 (2004). `hep-ph/0402159`.

[15] E. Richter-Was, *AcerDET: a particle level fast simulation and reconstruction package for phenomenological studies on high p_t physics at LHC*. Preprint hep-ph/0207355, 2002.

[16] B. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*. Preprint hep-ph/0405247, 2004.

[17] *LHAPDF online manual*, available on `http://durpdg.dur.ac.uk/lhapdf`.

[18] M. Dittmar and H. Dreiner, Phys. Rev. **D55**, 167 (1997).

[19] G. Davatz, G. Dissertori, M. Dittmar, M. Grazzini, and F. Pauss, JHEP **05**, 009 (2004).

[20] CMS Collaboration, *CMS TDR 6.2*, 2002. CERN-LHCC-2002-26, ch.15, p.317.

[21] G. Corcella and S. Moretti, Phys. Lett. **B590**, 249 (2004).

[22] S. Catani, D. de Florian, and M. Grazzini, JHEP **0105**, 025 (2001);
R. V. Harlander and W. B. Kilgore, Phys. Rev. **D64**, 013015 (2001);
R. V. Harlander and W. B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002);
C. Anastasiou and K. Melnikov, Nucl. Phys. **B646**, 220 (2002);
C. Anastasiou and K. Melnikov, Phys. Rev. Lett. **93**, 262002 (2004);
J. S. V. Ravindran and W. L. van Neerven, Nucl. Phys. **B665**, 325 (2003).

[23] I. D. A. Moraes, C. Buttar and P. Hodgson, available on http://amoraes.home.cern.ch/amoraes/. ATLAS internal notes.

[24] R. Field, available on http://www.phys.ufl.edu/rfield/cdf/. CDF Note 6403.

[25] ZEUS Collaboration, S. Chekanov *et al.* (2005). hep-ex/0502029;
H1 Collaboration, C. Adloff *et al.*, Nucl. Phys. **B38**, 3 (1999).

[26] LHCb Collaboration, *LHCb reoptimized detector design and performance TDR*, 2003. CERN-LHCC-2003-030.

Unintegrated parton density functions

John Collins[1]*, Markus Diehl*[2]*, Hannes Jung*[2]*, Leif Lönnblad*[3]*, Michael Lublinsky*[4]*, Thomas Teubner*[5]
[1] Physics Department,Penn State University, U.S.A.
[2] Deutsches Elektronen-Synchroton Hamburg, FRG
[3] Department of Theoretical Physics, Lund University, Sweden
[4] University of Connecticut, U.S.A.
[5] University of Liverpool, U.K

Abstract
An overview on activities to determine unintegrated parton density functions is given and the concept and need for unintegrated PDFs is discussed. It is also argued that it is important to reformulate perturbative QCD results in terms of fully unintegrated parton densities, differential in all components of the parton momentum. Also the need for non-linear BFKL evolution is discussed and results using the BK equation supplemented by DGLAP corrections at short distances is reviewed. Finally the use unintegrated generalized parton distributions for hard diffractive processes is discussed.

1 Unintegrated parton density functions[1]

The parton distributions of hadrons still cannot be calculated from first principles, but have to be determined experimentally. However, once the initial distributions $f_i^0(x, \mu_0^2)$ at the hadronic scale ($\mu^2 \sim 1$ GeV^2) are determined, different approximations allow to calculate the parton density functions (PDFs) for different kinematic regions:

- DGLAP [1, 2, 3, 4] describes the evolution with the scale μ^2
- BFKL [5, 6, 7] describes the evolution in the longitudinal momenta x
- CCFM [8, 9, 10, 11] describes the evolution in an angular ordered region of phase space while reproducing DGLAP and BFKL in the appropriate asymptotic limits

The different evolution equations attempt to describe different regions of phase space on the basis of in perturbative QCD (pQCD).

1.1 Introduction to uPDFs and $k_\perp$ factorization

In the collinear factorization ansatz the cross sections are described by x-dependent density functions $f_i(x, \mu^2)$ of parton i at a given factorization scale μ convoluted with an (on-shell) coefficient function (matrix element):

$$\sigma(a + b \to X) = \int dx_1 dx_2 f_i(x_1, \mu^2) f_j(x_2, \mu^2) \hat{\sigma}_{ij}(x_1, x_2, \mu^2) \tag{1}$$

with $\hat{\sigma}_{ij}(x_1, x_2, \mu^2)$ being the hard scattering process for the partons $i+j \to X$. In this equation we have left implicit all external kinematic variables, keeping only the variables used in the parton densities. This ansatz is very successful in describing inclusive cross sections, such as the structure function $F_2(x, Q^2)$ at HERA or the inclusive production of vector bosons or Drell-Yan in proton proton collisions. The free parameters of the starting distributions $f_i^0(x, \mu_0^2)$ are determined such that after a DGLAP evolution to the scale $\mu^2 = Q^2$ and convolution with the coefficient functions the measured structure function $F_2(x, Q^2)$ at HERA (and, usually, some other cross sections, e.g., in hadron-hadron and neutrino-hadron scattering) are best described.

[1] authors: Hannes Jung and Leif Lönnblad.

However, as soon as, for example, final-state processes are considered, the collinear factorization ansatz becomes more and more unreliable, because neglecting the transverse momenta of the partons during the (DGLAP) evolution leads to inconsistencies, as will be discussed in more detail in section 2. Collinear factorization is only appropriate when (a) the transverse momentum (and virtuality) of the struck parton(s) can be neglected with respect to Q, and (b) the integrals over these variables can be treated as independent and unrestricted up to the scale Q. (Certain complications concerning high transverse momentum partons are correctly treated by NLO and higher corrections to the hard scattering.) When these requirements are not met, a more general treatment using unintegrated parton densities (uPDFs) is better.

For example, in the small x regime, when the transverse momenta of the partons are of the same order as their longitudinal momenta, the collinear approximation is no longer appropriate and high energy or $k_\perp$ - factorization has to applied, with the appropriate BFKL or CCFM evolution equations. Cross sections are then $k_\perp$- factorized [12, 13, 14, 15] into an off-mass-shell ($k_\perp$- dependent) partonic cross section $\hat{\sigma}(x_1, x_2, k_{\perp 1}, k_{\perp 2})$ and a $k_\perp$- unintegrated parton density function (uPDF) $\mathcal{F}(z, k_\perp)$:

$$\sigma = \int dx_1 dx_2 d^2k_{\perp 1} d^2k_{\perp 2} \hat{\sigma}_{ij}(x_1, x_2, k_{\perp 1}, k_{\perp 2}) \mathcal{F}(x_1, k_{\perp 1}) \mathcal{F}(x_2, k_{\perp 2}) \tag{2}$$

The unintegrated gluon density $\mathcal{F}(z, k_\perp)$ is described by the BFKL evolution equation in the region of asymptotically large energies (small x). It is important to note that only when the $k_\perp$dependence of the hard scattering process $\hat{\sigma}$ can be neglected, i.e. if $\hat{\sigma}(x_1, x_2, k_{\perp 1}, k_{\perp 2}) \sim \hat{\sigma}(x_1, x_2, 0, 0)$, then the $k_\perp$integration can be factorized and an expression formally similar to eq.(1) is obtained.

An appropriate description valid for both small and large x, is given by the CCFM evolution equation, resulting in an unintegrated gluon density $\mathcal{A}(x, k_\perp, \mu)$, which is a function also of the additional evolution scale μ. This scale is connected to the factorization scale in the collinear approach.

Further examples where uPDFs are needed are the Drell-Yan and related processes at low transverse momentum, as in the CSS formalism [16]. However, the relation between CSS method (which does not need small x) and $k_\perp$-factorization of the BFKL/CCFM kind (for small x) has not yet been properly worked out.

1.2 Extraction and determination of uPDFs

In this section we will review how measurements of uPDFs have been extracted from DIS data at small x, mostly from the inclusive structure function F_2. For measurements of the uPDFs in Drell-Yan processes using the CSS formalism, see [17].

From the DIS data, the uPDF can be obtained by adjusting the non-perturbative input distribution $f_i^0(x, \mu_0^2)$ and the free parameters of the perturbative evolution such that after convolution with the appropriate off-shell matrix element (according to eq.(2)) a measured cross section is best described.

Applying $k_\perp$-factorization to determine the uPDF from DIS data until now mainly the inclusive structure function measurements of $F_2(x, Q^2)$ at HERA have been used. The exceptions are those which are simply derivatives of integrated PDFs, which then neglects fully the transverse momentum dependence of the matrix element. Extracting a uPDF from the integrated PDF is appropriate only if the $k_\perp$-dependence of the hard scattering process $\hat{\sigma}$ in eq.(2) can be neglected. In addition, contributions from $k_\perp > \mu$, which are present in a full calculation, are entirely neglected. It thus can only provide an estimate of the proper kinematics in the collinear approach, which is otherwise fully neglected when using integrated PDFs.

Here we will compare some of these parameterizations which have been obtained in different ways

- **CCFM set A0** was obtained using CCFM evolution in [18, 19].

- **LDC standard** was similarly obtained in [20] using LDC evolution [21], which is a reformulation and generalization of CCFM.
- **KKSL** [22] was obtained from a combined BFKL and DGLAP evolution following [23]
- **GLLM** [24] was obtained applying the BK equation to HERA F_2 measurements, as described in section 3
- **KMR** is one of the more advanced derivatives of integrated PDFs, using Sudakov form factors [25].

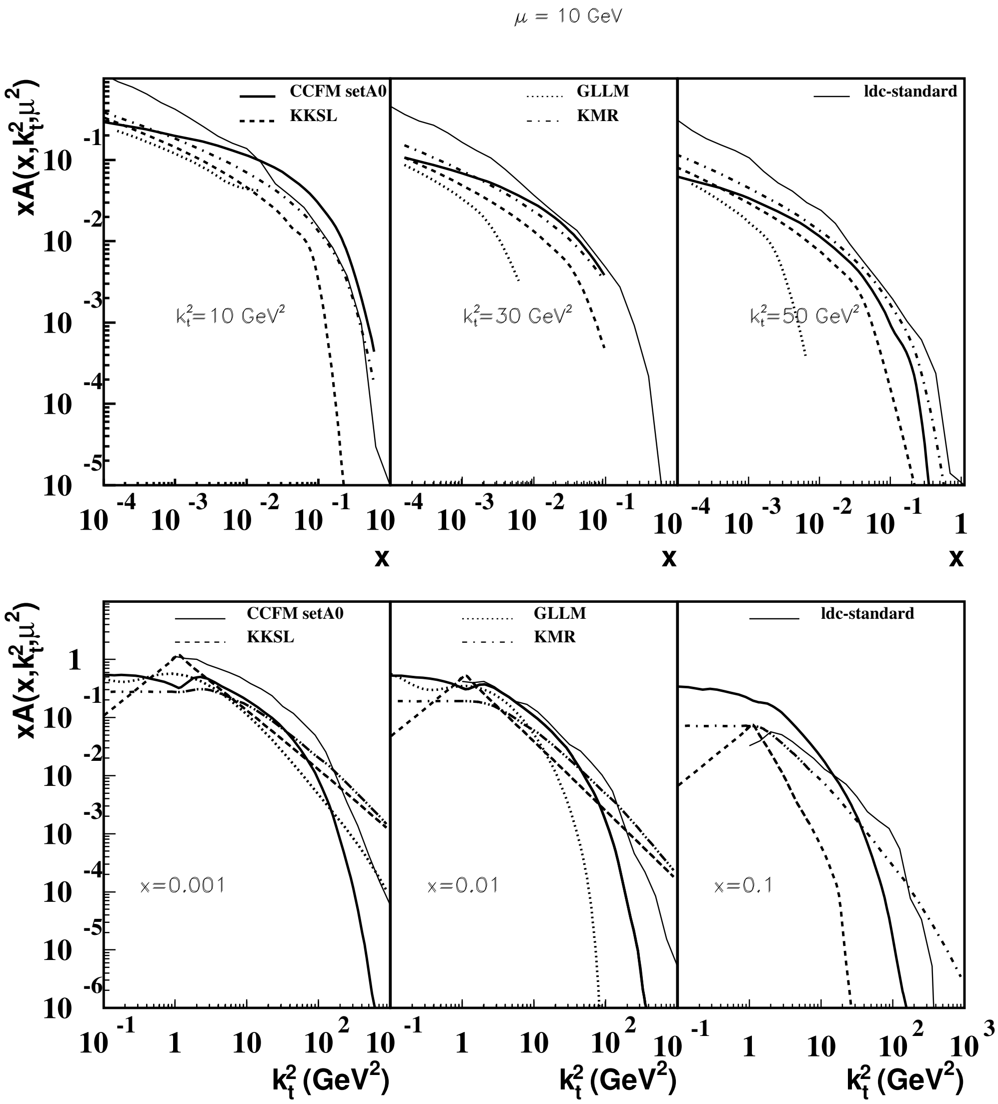


Fig. 1: Comparison of different uPDFs at $\mu = 10$ GeV.

In Fig. 1 we show a comparison of the different uPDFs as a function of x and $k_\perp$ at a factorization scale $\mu = 10$ GeV. All the parameterizations are able to describe the measured $F_2(x, Q^2)$ in the small x range reasonably well, with a $\chi^2/ndf \sim 1$. In Fig. 2 the same uPDFs are compared at a factorization scale which is relevant at LHC energies, e.g. for inclusive Higgs production ($\mu = 120$ GeV). One should note that the uPDFs from KKSL and GLLM have no explicit factorization scale dependence, therefore they are the same as in Fig 1.

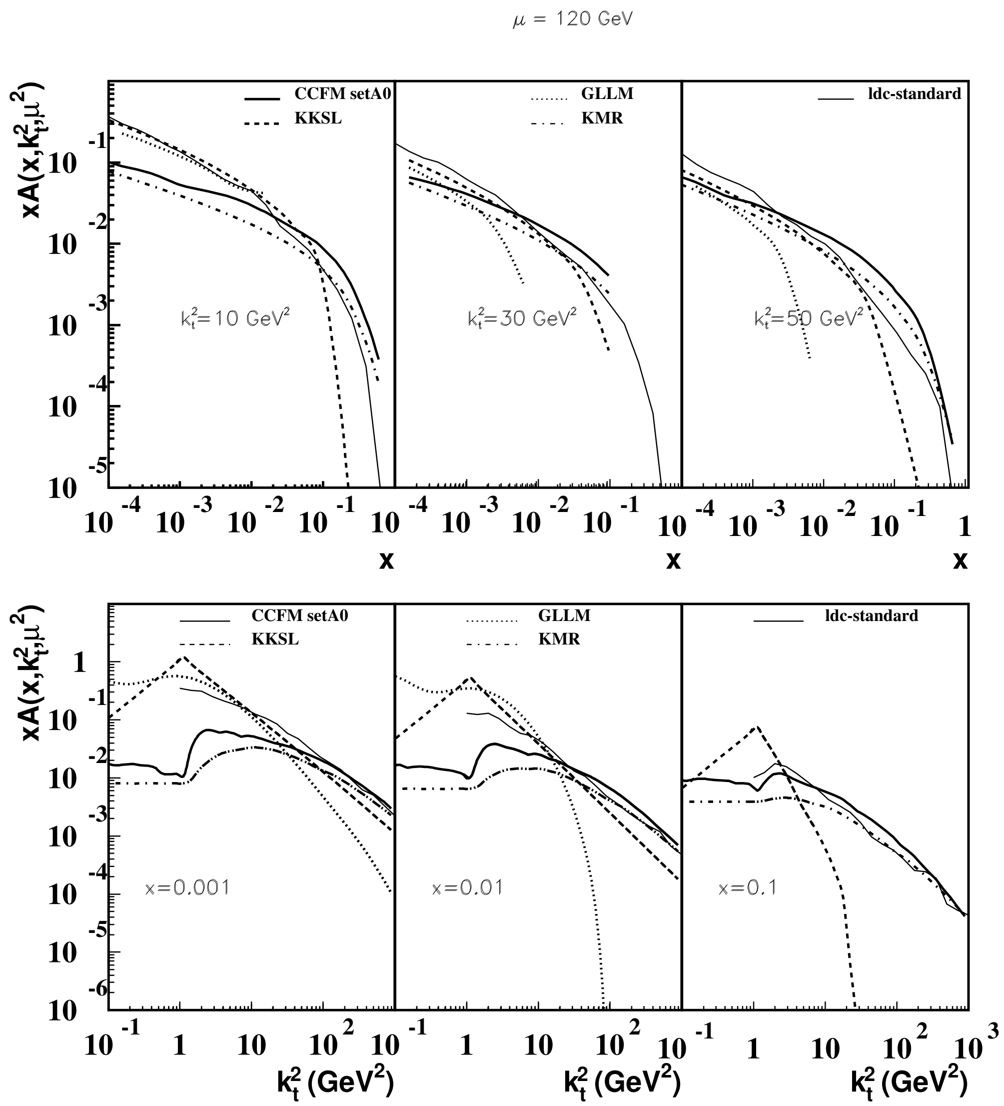


Fig. 2: Comparison of different uPDFs at $\mu = 120$ GeV.

1.3 Extrapolation to LHC energies

All the parameterizations of uPDFs considered in this report give a fairly good fit to HERA F_2 data. This means that they are well constrained mainly in the region of small x and relatively small Q^2, where the bulk of the HERA data is concentrated. For higher x and Q^2, a fit to HERA data is less constraining, and indeed some of the parameterizations based on the CCFM and LDC evolution of the gluon alone are only fitted in the small-x region (typically $x < 0.01$, $Q^2 < 100\,\mathrm{GeV}^2$).

When evolving the uPDFs to apply them to the processes of main interest at the LHC, such as Higgs production, this is a serious concern. Although the x-values in such processes are typically below 0.01, the scales involved are of the order of 10^4 GeV^2 or more. Through the evolution one then becomes sensitive to larger x-values at lower scales where the current parameterizations are less constrained.

A notable exception is the KMR [25] densities which are obtained from a global fit of integrated PDFs, which should give reliable prediction at LHC at least for integrated observables such as the inclusive Higgs cross section. In contrast, it was shown in [20] that the CCFM [8, 9, 10, 11] and LDC [21] evolved uPDFs have unreasonably large uncertainties for such cross sections. On the other hand it was also shown in [20] that there are some questions about the constraint of the actual $k_\perp$ distribution of the KMR uPDFs resulting eg. in a too soft $p_\perp$ spectrum of W or Z production at the Tevatron for small transverse momenta. Hence, although the KMR prediction for inclusive quantities may be reliable at the LHC, the predictions of eg. the detailed distribution of low-$p_\perp$ Higgs may be questionable.

What is needed is clearly to obtain fits of the uPDFs, not only to HERA F_2 data, but also to observables more sensitive to higher x and Q^2 values, as well as to observables directly sensitive to the $k_\perp$ distribution. To obtain such global fits there is a need for both theoretical and phenomenological developments. Examples of the former is the inclusion of quarks in the CCFM evolution, while the latter involves the development of $k_\perp$-sensitive observables, where HERA data at small x, such as forward jet or heavy quark production, will play an important role, as discussed in the following.

1.4 Global uPDF fits

Until now the uPDFs obtained from DIS were only determined and constrained by the inclusive structure function $F_2(x, Q^2)$. It is clear that the inclusive measurements are not very sensitive to the details of the $k_\perp$dependence. In Fig. 3 we show the $k_\perp$distribution of the gluon in $\gamma^* g^* \rightarrow q\bar{q}$ which is the relevant process for F_2 at small x. The $k_\perp$-distributions in Fig. 3 are obtained with CASCADE [26, 27] using the CCFM uPDFs. The bins in Q^2 are typical for HERA F_2 measurements. It is interesting to observe that even at large Q^2 essentially only the small $k_\perp$region is probed by F_2.

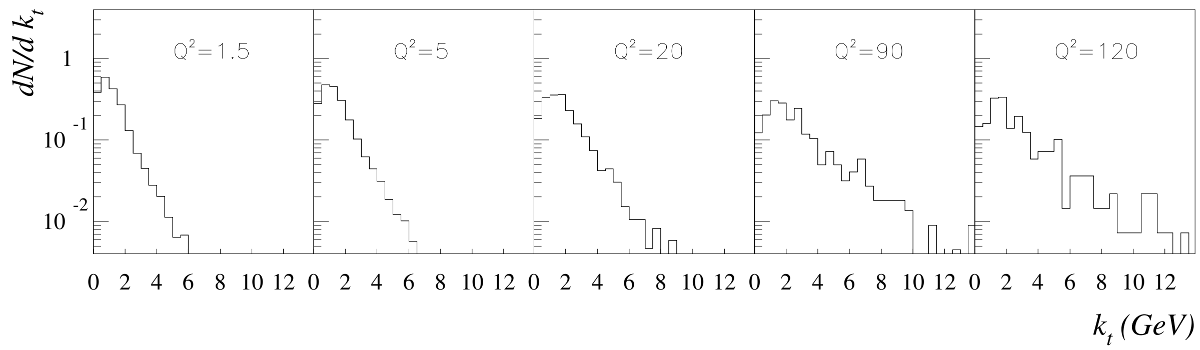


Fig. 3: $k_\perp$distribution in different Q^2 bins used in $F_2(x, Q^2)$ at HERA.

A larger lever arm for the $k_\perp$distribution can be obtained with photoproduction of D^* + jet events at HERA. In Fig. 4 the relevant diagram is shown. The quantity x_γ, normally designed to separate direct from resolved photon processes, can be also used to distinguish small and large $k_\perp$- regions. The region of large x_γ corresponds to measuring jets coming from the quark-box. The region of small x_γ corresponds to the situation where one of the jets originates from a gluon, as indicated in Fig. 4. Thus, the transverse momentum of the gluon i can be probed, as shown in Fig. 5 for two different regions of x_γ using CASCADE . It is interesting to note that the average $k_\perp$distribution for bottom production at the Tevatron is similar to what it shown in Fig. 5.

To further constrain the uPDF it would be desirable to perform a common fit to inclusive measurements like F_2 and simultaneously to final state measurements.

Once the data sets and the sensitivity to the uPDFs have been identified, a systematic error treatment of the data used in the uPDF fits can be performed. Until now, the uPDFs are not really the result of a fit but rather a proof that the uPDF is consistent with various measurements.

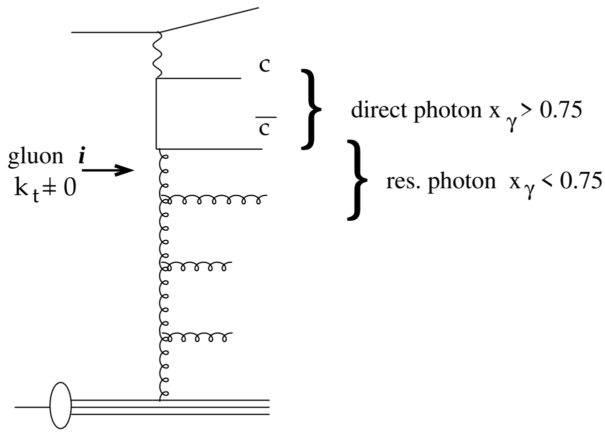


Fig. 4: Diagram of charm photoproduction, showing the sensitivity to the gluon transverse momentum

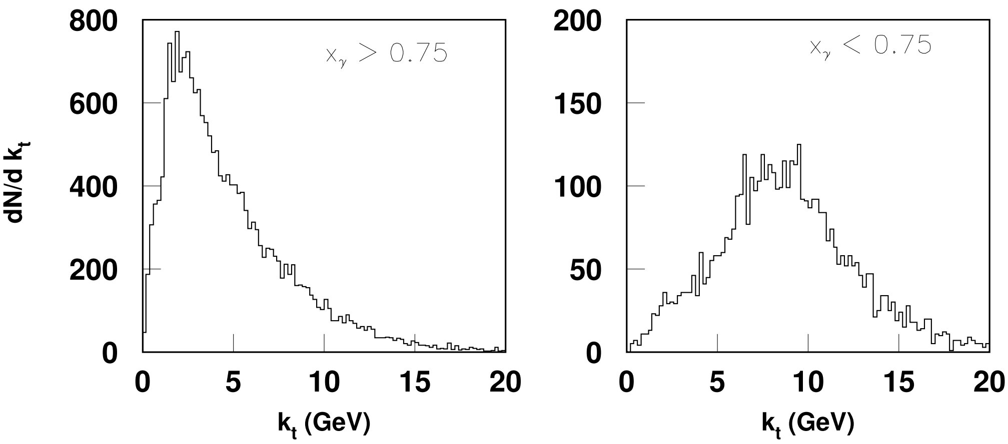


Fig. 5: $k_\perp$distribution in different x_γ bins obtained from D^*+jet photo-production at HERA.

A uPDF fit would require a systematic variation of the parameters used to specify the non-perturbative input gluon distribution as well as a systematic treatment of the experimental systematic uncertainties. Only then an uncertainty band of the uPDFs can be given. To consider the uncertainty of the uPDF given from the spread of different available parameterizations is a very rough estimate.

1.5 Outlook and Summary

Clearly, the extraction of uPDFs from data is still in its infancy, especially if compared to the well developed industry of fitting integrated PDFs. The uPDFs are only leading order parameterizations, they have mainly been fitted to F_2 data at small x, and besides the KMR and LDC parameterizations, no attempts have been made to obtain unintegrated quark densities. Taken together, this means that the applicability to LHC processes are uncertain. However, the field is maturing and we hope to soon be able to do more global uPDF fits which will greatly enhance the reliability of the predictions for the LHC. In doing so the small-x data from HERA will be very important, but also eg. Tevatron data will be able to provide important constraints.

2 Need for fully unintegrated parton densities[2]

2.1 Introduction

Conventional parton densities are defined in terms of an integral over all transverse momentum and virtuality for a parton that initiates a hard scattering. While such a definition of an integrated parton density is appropriate for very inclusive quantities, such as the ordinary structure functions F_1 and F_2 in DIS, the definition becomes increasingly unsuitable as one studies less inclusive cross sections. Associated with the use of integrated parton densities are approximations on parton kinematics that can readily lead to unphysical cross sections when enough details of the final state are investigated.

We propose that it is important to the future use of pQCD that a systematic program be undertaken to reformulate factorization results in terms of fully unintegrated densities, which are differential in both transverse momentum and virtuality. These densities are called "doubly unintegrated parton densities" by Watt, Martin and Ryskin [28, 29], and "parton correlation functions" by Collins and Zu [30]; these authors have presented the reasoning for the inadequacy, in different contexts, of the more conventional approach. The new methods have their motivation in contexts such as Monte-Carlo event generators where final-state kinematics are studied in detail. Even so, a systematic reformulation for other processes to use unintegrated densities would present a unified methodology.

These methods form an extension of $k_\perp$-factorization. See Sec. 1 for a review of $k_\perp$-factorization, which currently involves two different formalisms, the BFKL/CCFM methods [5, 6, 7, 8, 9, 10, 11] and the CSS method [16].

2.2 Inadequacy of conventional PDFs

The problem that is addressed is nicely illustrated by considering photoproduction of $c\bar{c}$ pairs. In Figs. 6, we compare three methods of calculation carried out within the CASCADE event generator [26, 27]:

- Use of a conventional gluon density that is a function of parton x alone.
- Use of a $k_\perp$ density that is a function of parton x and $k_\perp$. These are the "unintegrated parton densities" (uPDFs) that are discussed in Sec. 1
- Use of a doubly unintegrated density that is a function of parton x, $k_\perp$ and virtuality, that is, of the complete parton 4-momentum.

The partonic subprocess in all cases is the lowest order photon-gluon-fusion process $\gamma + g \longrightarrow c + \bar{c}$ (Fig. 7). Two differential cross sections are plotted: one as a function of the transverse momentum of the $c\bar{c}$ pair, and the other as a function of the x_γ of the pair. By x_γ is meant the fractional momentum of the photon carried by the $c\bar{c}$ pair, calculated in the light-front sense as

$$x_\gamma = \frac{\sum_{i=c,\bar{c}}(E_i - p_{z\,i})}{2yE_e} = \frac{p^-_{c\bar{c}}}{q^-}.$$

Here E_e is the electron beam energy and the coordinates are oriented so that the electron and proton beams are in the $-z$ and $+z$ directions respectively.

In the normal parton model approximation for the hard scattering, the gluon is assigned zero transverse momentum and virtuality, so that the cross section is restricted to $p_{Tc\bar{c}} = 0$ and $x_\gamma = 1$, as shown by the solid lines in Fig. 6(a,b). When a $k_\perp$ dependent gluon density is used, quite large gluonic $k_\perp$ can be generated, so that the $p_{Tc\bar{c}}$ distribution is spread out in a much more physical way, as given by the dashed line in Fig. 6(a). But as shown in plot (b), x_γ stays close to unity. Neglecting the full recoil mass $m_{\rm rem}$ (produced in the shaded subgraph in Fig 7) is equivalent of taking $k^2 = -k_\perp^2/(1-x)$ with k^2 being the virtuality of the gluon in Fig. 7, $k_\perp$ its transverse momentum and x its light cone energy fraction. This gives a particular value to the gluon's k^-. When we also take into account the correct virtuality of gluon, there is no noticeable change in the $p_{Tc\bar{c}}$ distribution — see Fig. 6(c) (dashed line) —

[2]authors: John Collins and Hannes Jung.

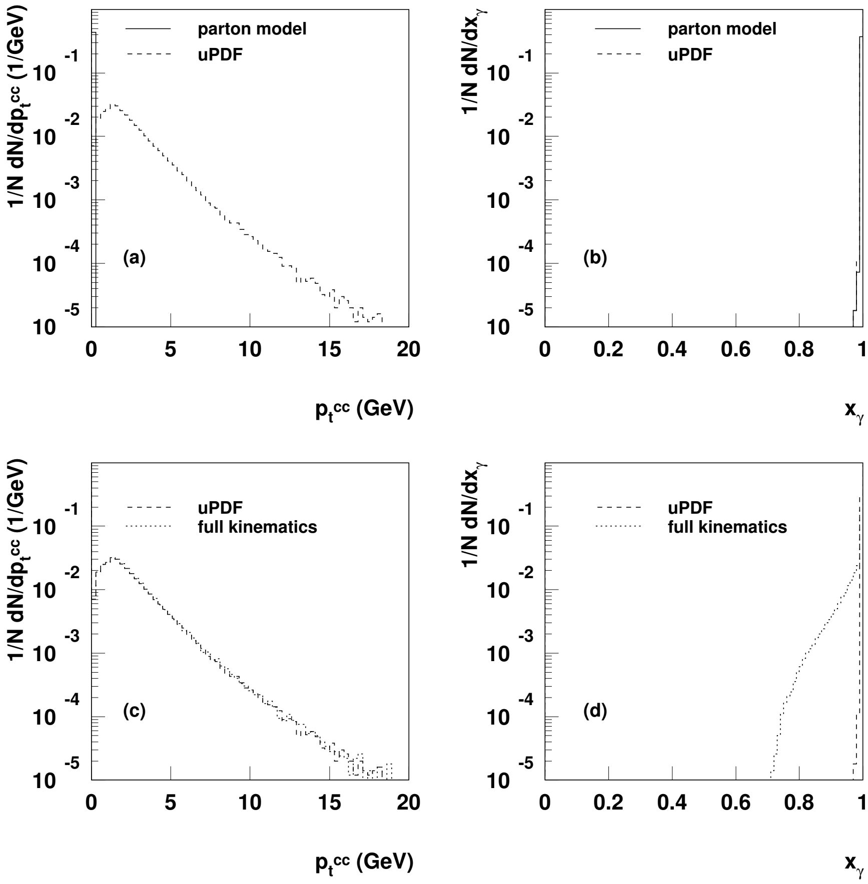


Fig. 6: (a) and (b): Comparison between use of simple LO parton model approximation and of the use of $k_\perp$ densities for the p_T of $c\bar{c}$ pairs in photoproduction, and for the x_γ. (c) and (d): Comparison of use of $k_\perp$ densities and full simulation.

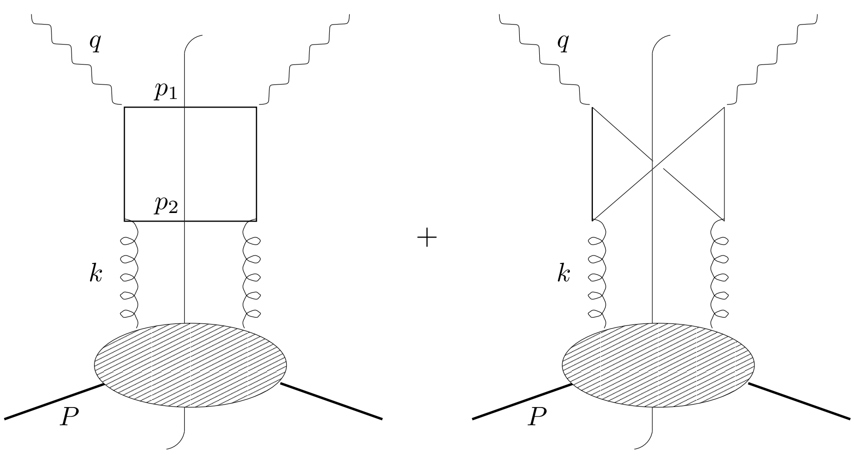


Fig. 7: Photon-gluon fusion.

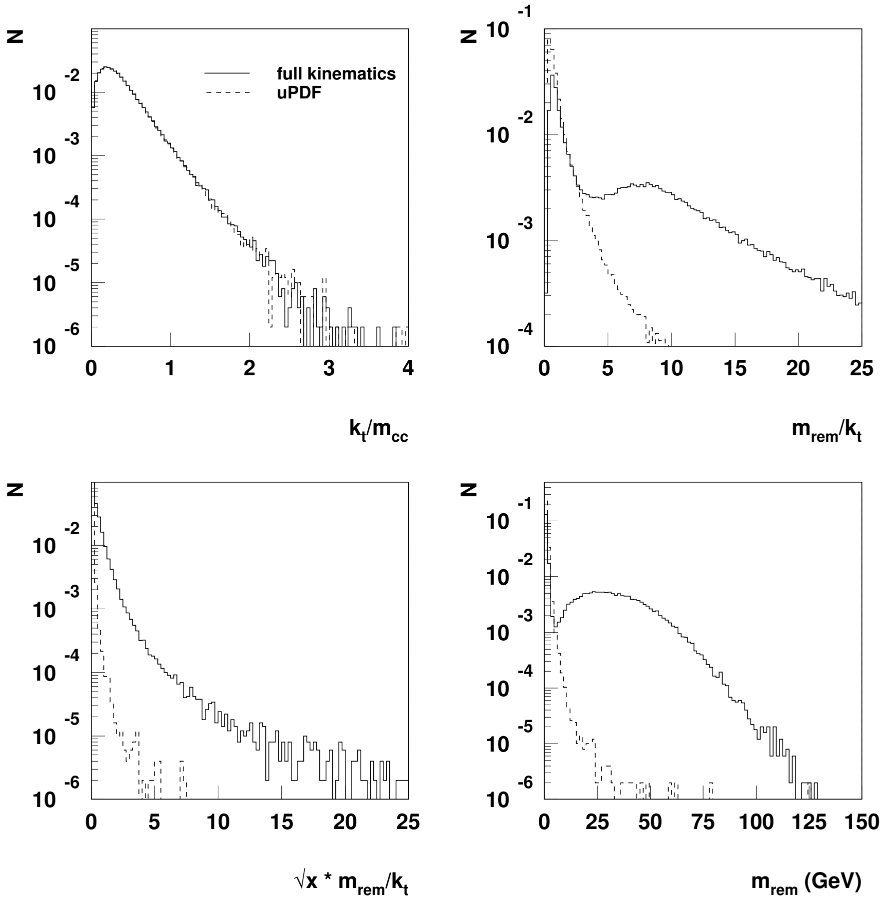


Fig. 8: Comparison of distributions in partonic variables between calculations with full parton kinematics and with ordinary unintegrated PDFs.

since that is already made broad by the transverse momentum of the gluon. But the gluon's k^- is able to spread out the x_γ distribution, as in Fig. 6(d) with the dashed line. This is equivalent with a proper treatment of the kinematics and results in $k^2 = -(k_\perp^2 + x m_{\rm rem}^2)/(1-x)$, where $m_{\rm rem}$ is the invariant mass of the beam remnant, the part of the final state in the shaded blob in Fig. 7. This change can be particularly significant if x is not very small.

Note that if partons are assigned approximated 4-momenta during generation of an event in a MC event generator, the momenta need to be reassigned later, to produce an event that conserves total 4-momentum. The prescription for the reassignment is somewhat arbitrary, and it is far from obvious what constitutes a correct prescription, especially when the partons are far from a collinear limit. A treatment with fully unintegrated PDFs should solve these problems.

If, as we claim, an incorrect treatment of parton kinematics changes certain measurable cross sections by large amounts, then we should verify directly that there are large discrepancies in the distributions in partonic variables themselves. We see this in Fig. 8. Graph (a) plots the gluonic transverse momentum divided by the charm-pair mass. As is to be expected, the typical values are less than one, but there is a long tail to high values. But the use of full parton kinematics does not have much of an effect, the unintegrated parton distributions already providing realistic distributions in transverse momentum.

On the other hand, a simple collinear approximation for showering sets the remnant mass, $m_{\rm rem}$, to zero. As can be seen from the formulae for the gluon virtuality, this only provides a good approximation to the gluon kinematics if $m_{\rm rem}$ is much less than $k_\perp$. In reality, as we see from graph (b), there is a long tail to large values of $m_{\rm rem}/k_\perp$, and the tail is much bigger when correct kinematics are used. A more correct comparison uses $xm_{\rm rem}^2$, with an extra factor of x. Even then, there is a large effect, shown in graph (c). The vertical scale is logarithmic, so the absolute numbers of events are relatively small, but the tail is broad. Finally, graph (d) shows that the distribution in $m_{\rm rem}$ itself is very broad, extending to many tens of GeV. This again supports the argument that unless a correct treatment of parton kinematics is made, very incorrect results are easily obtained.

It is important to note that, for the cross sections themselves, the kinematic variables used in Fig. 6 are normal ones that are in common use. Many other examples are easily constructed. Clearly, the use of the simple parton-model kinematic approximation gives unphysically narrow distributions. The correct physical situation is that the gluon surely has a distribution in transverse momentum and virtuality, and for the considered cross sections neglect of parton transverse momentum and virtuality leads to wrong results. It is clearly better to have a correct starting point even at LO, for differential cross sections such as we have plotted.

2.3 Kinematic approximations

The standard treatment of parton kinematics involves replacing the incoming parton momentum k by its plus component only: $k^\mu \mapsto \hat{k}^\mu \equiv (k^+, 0, 0_T)$. There are actually two parts to this. The first is to neglect the $^-$ and transverse components of k with respect to the large transverse momenta in the calculation of the numerical value of the hard-scattering amplitude; this is a legitimate approximation, readily corrected by higher order terms in the hard scattering. The second part is to change the kinematics of the final-state particles, p_1 and p_2, so that their sum is q plus the approximated gluon momentum. It is this second part that is problematic, for it amounts to the replacement of the momentum conservation delta function $\delta^{(4)}(k + q - p_1 - p_2)$ by $\delta^{(4)}(\hat{k} + q - p_1 - p_2)$. These delta-functions are infinitely different, point-by-point. Only when integrated with a sufficiently smooth test function can they be regarded as being approximately the same, as in a fully inclusive cross section.

In an event generator, the effect is to break momentum conservation, which is restored by an ad hoc correction of the parton kinematics. Note that the change of parton kinematics is only in the hard scattering, i.e., in the upper parts of the graphs. Parton kinematics are left unaltered within the parton density part, and the integrals over $k_\perp$ and virtuality are part of the standard definition of integrated PDFs.

The situation is ameliorated by inclusion of NLO terms, and perhaps also by some kind of resummation. But these do not correct the initial errors in the approximation, and lead to a very restricted sense in which the derivation of the cross section can be regarded as valid. Furthermore, when much of the effect of NLO terms is to correct the kinematic approximations made in LO, this is an inefficient use of the enormous time and effort going into NLO calculations. A case in point is the BFKL equation, where 70% of the (large) NLO corrections are accounted for [31] by the correction of kinematic constraints in the LO calculation.

2.4 Conclusions

The physical reasoning for the absolute necessity of fully unintegrated densities is, we believe, unquestionable. Therefore it is highly desirable to reformulate perturbative QCD methods in terms of doubly unintegrated parton densities from the beginning. A full implementation will be able to use the full power of calculations at NLO and beyond.

Among other things, a full implementation, as in [30], will provide extra factorization formulae for obtaining the values of the unintegrated densities at large parton transverse momentum and virtuality.

This will incorporate all possible perturbatively calculable information, so that the irreducible nonperturbative information, that must be obtained from data, will be at low transverse momentum and virtuality. In addition, the implementation will quantify the relations to conventional parton densities. With the most obvious definitions, the integrated PDFs are simple integrals of the unintegrated densities. However, in full QCD a number of modifications are required [30, 32], so that the relations between integrated and unintegrated PDFs are distorted.

The fact that we propose new and improved methods does not invalidate old results in their domain of applicability. The work of Watt, Martin and Ryskin, and of Collins and Zu provides a start on this project; but much remains to be done to provide a complete implementation in QCD; for example, there is as yet no precise, valid, and complete gauge-invariant operator definition of the doubly unintegrated densities in a gauge theory.

The outcome of such a program should have the following results:

1. Lowest order calculations will give a kinematically much more realistic description of cross sections. This may well lead to NLO and higher corrections being much smaller numerically than they typically are at present, since the LO description will be better.
2. It will also obviate the need for separate methods (resummation or the CSS technique), which are currently applied to certain individual cross sections like the transverse-momentum distribution for the Drell-Yan process. All these and others will be subsumed and be given a unified treatment.
3. A unified treatment will be possible for both inclusive cross sections using fixed order matrix element calculations and for Monte-Carlo event generators.
4. For a long-term theoretical perspective, the doubly unintegrated distributions will interface to methods of conventional quantum many-body physics much more easily than regular parton densities, whose definitions are tuned to their use in ultra-relativistic situations.

This program is, of course, technically highly nontrivial if it is to be used in place of conventional methods with no loss of predictive power. A start is made in the cited work.

Among the main symptoms of the difficulties are that the most obvious definition of a fully unintegrated density is a matrix element of two parton fields at different space-time points, which is not gauge-invariant. It is often said that the solution is to use a light-like axial gauge $A^+ = 0$. However, in unintegrated densities, this leads to divergences — see [32] for a review — and the definitions need important modification, in such a way that a valid factorization theorem can be derived.

We also have to ask to what extent factorization can remain true in a generalized sense. Hadron-hadron collisions pose a particular problem here, because factorization needs a quite nontrivial cancellation arising from a sum over final-state interactions. This is not compatible with simple factorization for the exclusive components of the cross section, and makes a distinction between these processes and exclusive components of DIS, for example.

3 PDF extrapolation to LHC energies based on combined BK/DGLAP equations [3]

3.1 Introduction

In recent years it became clear that the DGLAP evolution is likely to fail in certain kinematics associated with the *low* x domain. This might be a dangerous problem for certain DGLAP based predictions made for the LHC. The reasons for the failure are well known.

- DGLAP predicts a very steep rise of gluon densities with energy. If not suppressed this rise will eventually violate unitarity.
- The leading twist evolution breaks down when higher twists become of the same order as the leading one. We have to recall here that higher twists are estimated to rise with energy much faster than the leading one [33].

[3] author: Michael Lublinsky.

- The DGLAP evolution is totally unable to describe physics of low photon virtualities.

It is most important to stress that NLO corrections are in principal unable to solve any of the above problems, though they can potentially help to delay their onset.

Fortunately, a solution to the low x problem does exist. We have to rely on a nonlinear evolution based on the BFKL dynamics. So far the best candidate on the market is the Balitsky-Kovchegov (BK) equation [34, 35], which is a nonlinear version of the LO BFKL equation. Compared to the DGLAP equation it has the following advantages:

- it accounts for saturation effects due to high parton densities.
- it sums higher twist contributions.
- it allows an extrapolation to large distances.

Though the BK evolution takes care of the low x domain, it misses the essential part of the short distance physics correctly accounted for by the DGLAP evolution. The reason is that the BFKL kernel involves the $1/z$ part only of the full gluon-gluon splitting function $P_{gg}(z)$. Thus we have to develop a scheme which in a consistent manner would use elements of both the equations. Such scheme was proposed in Ref. [36] and realized in a successful fit to F_2 data in Ref. [37].

One of the main problems of the DGLAP evolution is a necessity to specify the x dependence of the distributions in the initial conditions of the evolution. The scheme which we propose generally avoids this problem and thus can be used for future more elaborated analysis including NLO corrections and the quark sector.

At low x it is very convenient to use the dipole picture. In this approach the structure function F_2 can be expressed through the universal dipole cross section σ^{dipole}:

$$F_2(y,Q^2) \;=\; \frac{Q^2}{4\pi^2}\int d^2\,r\int d\,z\; P^{\gamma^*}(Q^2;r,z)\;\sigma^{dipole}(r,y)\;. \tag{3}$$

with the probability to find a dipole of the transverse size r in the photon‘s wavefunction given by

$$P^{\gamma^*}(Q^2;r,z)^2 \;=\; \frac{N_c}{2\,\pi^2}\sum_{f=1}^{3} Z_f^2\,\left\{(z^2+(1-z)^2)\,a^2\,K_1^2(a\,r)\;+\;4\,Q^2\,z^2\,(1-z)^2\,K_0^2(a\,r)\right\}\,,$$

where $a^2 = Q^2 z(1-z)$, Z_f are the quark charges, and K_i the standard modified Bessel functions.

The dipole cross section is determined through the evolution of the imaginary part of the dipole target elastic amplitude N subsequently integrated over the impact parameter b (in the analysis of Ref. [37] the dependence on b was modeled):

$$\sigma^{dipole}(r,y)\;=\;2\int d^2\,b\;N(r,y;b)\,.$$

In our approach, the amplitude $\tilde{N}$ is given by a sum of two terms

$$N \;=\; \tilde{N} \,+\, \Delta N$$

The first term $\tilde{N}$ follows from the solution of the BK equation whereas ΔN is a DGLAP correction to it (Fig. 9). The strategy of the fit is the following. We trust the DGLAP evolution for x above $x_0 = 10^{-2}$. The gluon density obtained as a result of this evolution is then used as a initial condition for the low x evolution based on the BK equation. In practice the CTEQ6 gluon was used as an input. The large distance behavior was extrapolated using the method proposed in Ref. [38]. The extrapolation is based on the geometrical scaling [39], a phenomenon experimentally observed by HERA. The BK evolved function N is fitted to the low Q^2 data, with the effective proton size being the only fitting parameter entering the b dependence ansatz. As the last step, the DGLAP correction ΔN is switched on and computed by solving a DGLAP-type equation. An inhomogeneous N-dependent term in the equation acts as a source term for ΔN. This allows to have zero initial condition for the DGLAP correction. [4]

[4]The initial condition for the BK equation is CTEQ gluon distribution. In the DGLAP-type equation for ΔN an initial condition at $r = r_0$ is required, which is set to zero and no modelling of the small x behavior is needed.

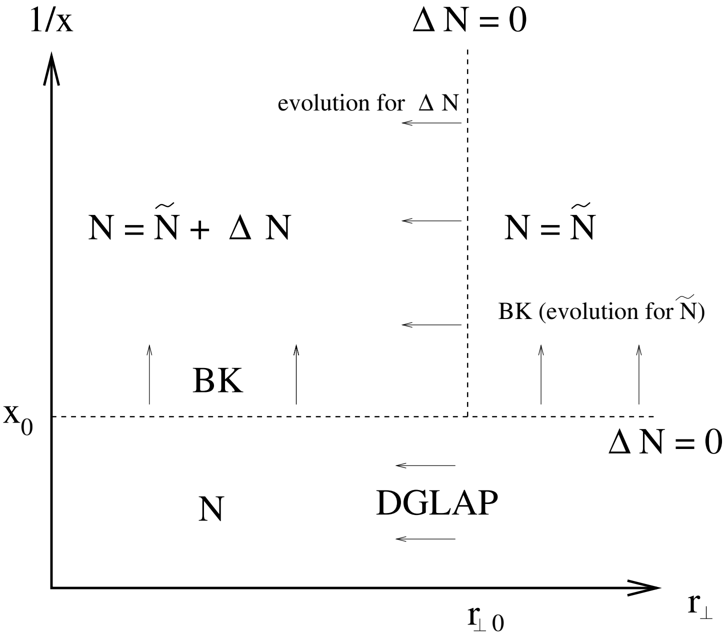


Fig. 9: *The kinematic map for the solutions.*

3.2 Results

We skip most of the technical details reported in Ref. [37] and present a result of the fit with $\chi^2/d.o.f. \simeq 1$. Fig. 10 displays the results vs. a combined set of experimental data for x below 10^{-2}. The solid line is the final parameterization. The dashed line on plot (b) is the result without DGLAP corrections added. Fig. 11, a presents our results for the logarithmic derivative of F_2 with respect to lnx. This graph

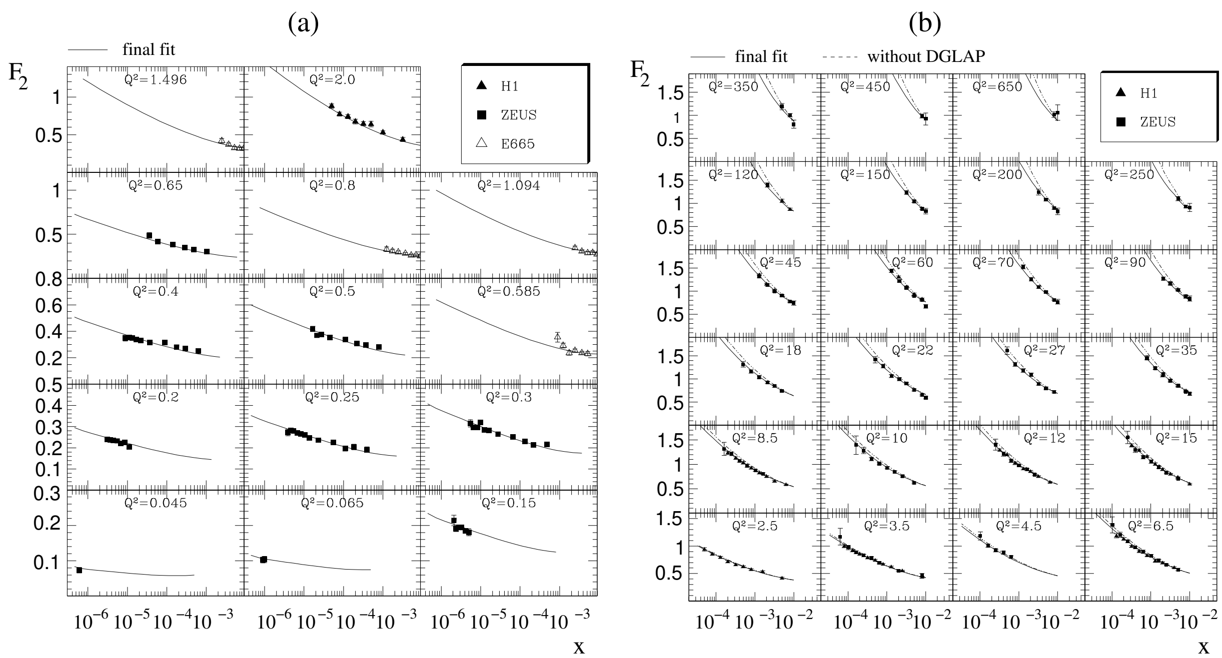


Fig. 10: *Fit to the F_2 structure function.*

illustrates the hard-soft pomeron transition as a result of multiple rescattering of the BFKL pomeron. The intercept decreases from the LO BFKL intercept of the order 0.3 to the hadronic value of the order 0.1. As clearly observed from Fig. 11a, the intercept depends strongly on the photon virtuality Q^2 and decreases towards hadronic value when the virtuality decreases. If we further increase the energy, the

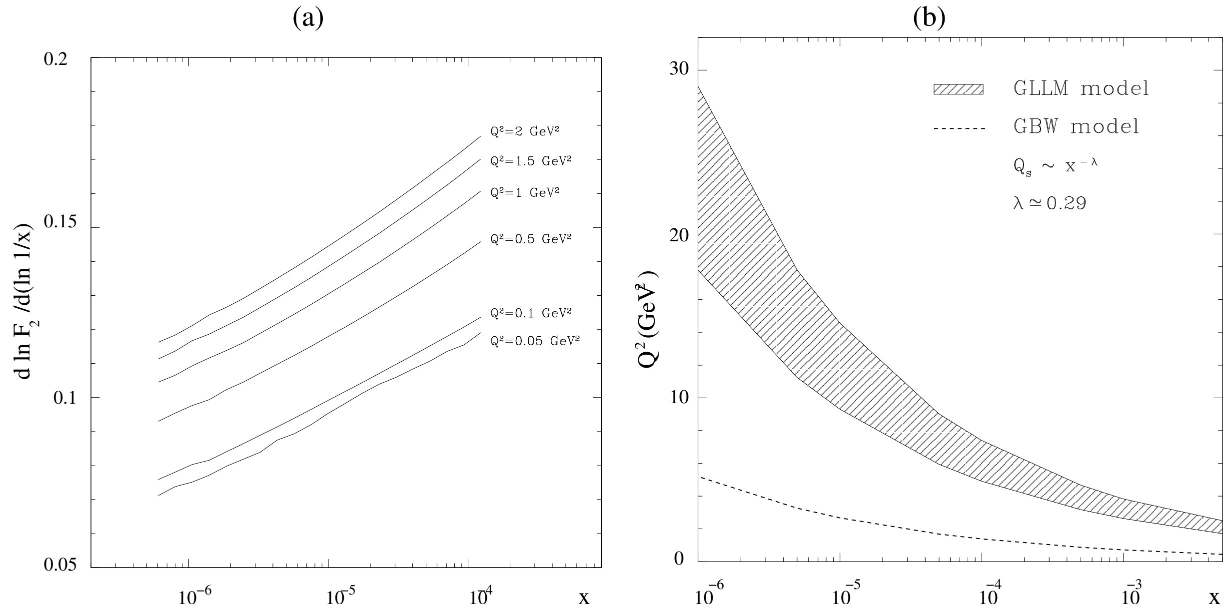


Fig. 11: *(a) The logarithmic derivative $\lambda = \partial \ln F_2 / \partial \ln 1/x$ plotted at low Q^2 and very low x. (b) Saturation scale. the hatched area defines a prediction band of Ref. [37]; dashed line is the GBW model.*

intercept would eventually vanish in accord with the unitarity requirements. The band of our estimates for the value of saturation scale at LHC is displayed on Fig. 11b together with the most popular Golec-Biernat Wüsthoff saturation model [40]. Based on our analysis we predict much stronger saturation effects compared to the ones which could be anticipated from the GBW model. Though the power growth of the saturation scale in both cases is given by the very same exponent of the order $\lambda \simeq 0.3$, we had to take a much stronger saturation input at the beginning of the evolution.

Model predictions for F_L at HERA and F_2 at LHC can be found in Ref. [37]. Having determined the dipole cross section we can relate it to the unintegrated gluon distribution $f(k, y)$:

$$\sigma^{dipole}(r, y) \;=\; \frac{4\,\pi^2}{N_c} \int \frac{d\,k^2}{k^4}\,[1\,-\,J_0(k\,r)]\;\alpha_s(k^2)\,f(k,y)\,. \tag{4}$$

The relation (4) can be inverted for f which can be then used as an input for any computation based on the k_t factorization scheme. The data set for the dipole cross section σ^{dipole} as well as for the unintegrated gluon f can be found in [24]. The uPDF is compared to other parameterizations in Fig.1.

3.3 Outlook

We have reported on, so far, the most advanced analysis of the F_2 data based on combined BK/DGLAP evolution equations. Though our approach incorporates most of the knowledge accumulated in saturation physics, it is not yet fully developed. The next essential steps would be to include NLO corrections both to BFKL and DGLAP. The quark sector should be also added into a unique scheme.

4 Generalized parton distributions[5]

The theoretical description of hard diffractive processes involves the gluon distribution in the proton. Such processes have a proton in the final state which carries almost the same momentum as the incident proton. Due to the small but finite momentum transfer, it is not the usual gluon distribution which appears, but its generalization to nonforward kinematics. Prominent example processes are the exclusive

[5] authors: Markus Diehl and Thomas Teubner.

production of mesons from real or virtual photons (Figure 12a) when either the photon virtuality or the meson mass provides a hard scale, virtual Compton scattering $\gamma^* p \to \gamma p$, and the diffractive production of a quark-antiquark pair (Figure 12b) in suitable kinematics. The generalized gluon distribution depends on the longitudinal momentum fractions x and x' of the emitted and reabsorbed gluon (which differ because of the longitudinal momentum transfer to the proton) and on the invariant momentum transfer $t = -(p-p')^2$. In its "unintegrated" form it depends in addition on the transverse momentum k_t of the emitted gluon. Another important process involving this distribution is exclusive diffractive production of a Higgs in pp scattering (Figure 13), discussed in detail in [41]. Note that the description of this process requires the gluon distribution to be unintegrated with respect to k_t, whereas the processes in $\gamma^{(*)}p$ collisions mentioned above can be treated either in k_t-factorization or in the collinear factorization framework, where k_t-integrated generalized parton distributions occur. Note also that Figures 12 and 13 show graphs for the process *amplitudes*: the cross section depends hence on the square of the gluon distribution for Figure 12, and on its fourth power for Figure 13.

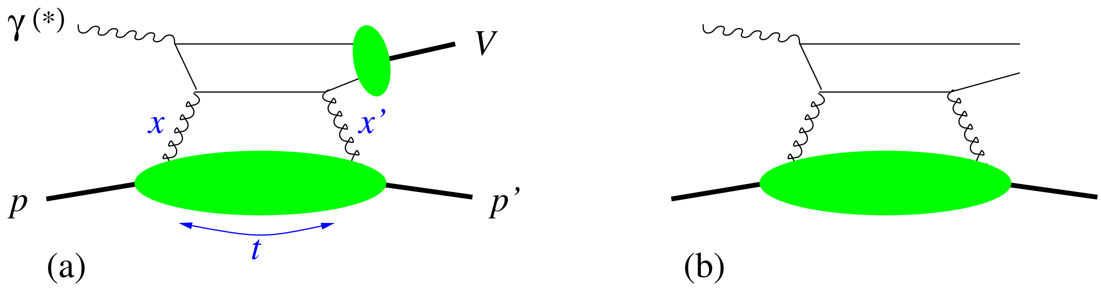


Fig. 12: Example graphs for the diffractive production of (a) a vector meson V or (b) a quark-antiquark pair. The large blob denotes the generalized gluon distribution of the proton and the small one the vector meson wave function.

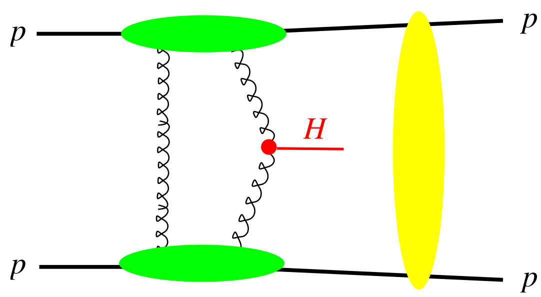


Fig. 13: Graph for the exclusive diffractive production of a Higgs boson, $p + p \to p + H + p$. The horizontal blobs indicate generalized gluon distributions, and the vertical blob represents secondary interactions between the projectiles.

To extract the generalized gluon distribution from vector meson production data requires knowledge of the meson wave function, which is an important source of uncertainty for the ρ^0 and ϕ and, to a lesser extent, for the J/Ψ. In this respect Υ production is by far the cleanest channel but experimentally challenging because of its relatively low production rate. An approach due to Martin, Ryskin and Teubner (MRT) [42] circumvents the use of the meson wave function by appealing to local parton-hadron duality, where the meson production cross section is obtained from the one for open quark-antiquark production, integrated over an interval of the invariant $q\bar{q}$ mass around the meson mass. The choice of that interval is then mainly reflected in an uncertainty in the overall normalization of the cross section. Virtual Compton scattering $\gamma^* p \to \gamma p$ does not involve any meson wave function and for sufficiently large Q^2 is again theoretically very clean.

By a series of steps one can relate the generalized gluon distribution to the usual gluon density, obtained for instance in global parton distribution fits.

1. The t dependence is typically parameterized by multiplying the distribution at $t = 0$ with an exponential $\exp(-b|t|)$, whose slope b has to be determined from measurement. In more refined models this slope parameter may be taken to depend on the other kinematic variables of the process.
2. To leading logarithmic accuracy in $\log(1/x)$ one can neglect the difference between the longitudinal momentum fractions of the two gluons. The amplitude for meson production is then proportional to the usual gluon density evaluated at $x_g = (M_V^2 + Q^2)/W^2$, where M_V is the meson mass, Q^2 the photon virtuality, and W the $\gamma^* p$ c.m. energy. For phenomenology this leading logarithmic approximation is however insufficient. A weaker approximation allows one to express the amplitude in terms of the gluon density at x_g times a correction factor for the kinematic asymmetry ("skewing") between the two momentum fractions [43].
3. The problem to relate the k_t unintegrated gluon distribution to the k_t integrated one is quite analogous to the case of the usual forward gluon density (see Sect. 1.1), with some specifics concerning Sudakov form factors in the nonforward case [44].

An overview and discussion of theoretical aspects and uncertainties in describing vector meson production in this framework can be found in [45].

To illustrate the sensitivity of such processes to the gluon distribution we show in Figure 14 data for photo- and electroproduction of J/Ψ compared to calculations in the MRT approach [46], with different gluon densities taken as input to construct the generalized gluon distribution as just described. The potential of such processes to constrain the gluon distribution is evident from this plot.

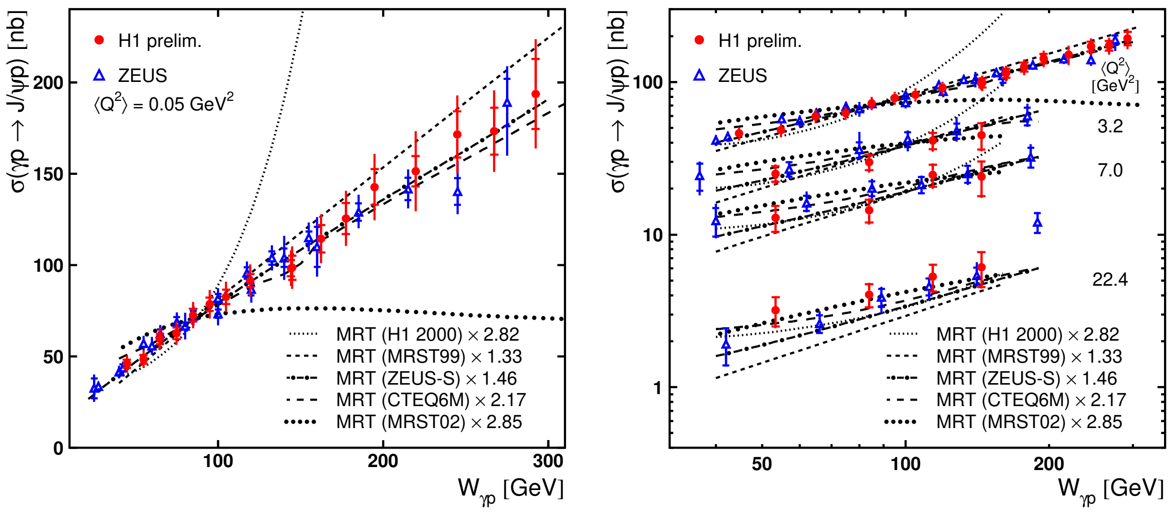


Fig. 14: Data for the $\gamma^* p \to J/\Psi\, p$ cross section from H1 [47] and ZEUS [48, 49] compared to calculations in the MRT approach [42, 46] with different gluon densities. The upper data points in the right panel correspond to those in the left one. The ZEUS data has been shifted to the Q^2 values of the H1 analysis using the Q^2 dependence measured by ZEUS, as described in [47]. Figure courtesy of Philipp Fleischmann (H1 Collaboration).

We finally note that the theoretical description of diffractive Higgs production in pp collisions is very similar to the description of diffractive processes in ep scattering using k_t factorization (much more than to the description of, say, inclusive DIS in collinear factorization, which provides the main input to the determination of conventional gluon densities at small x), see [41, 50] for further discussion. The analysis of diffractive ep scattering is hence well suited to provide input to estimate the diffractive Higgs cross section at the LHC.

Acknowledgments
This work is supported in part (JC) by the U.S. DOE.

References

[1] V. Gribov and L. Lipatov, Sov. J. Nucl. Phys. **15**, 438 and 675 (1972).

[2] L. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975).

[3] G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977).

[4] Y. Dokshitser, Sov. Phys. JETP **46**, 641 (1977).

[5] E. Kuraev, L. Lipatov, and V. Fadin, Sov. Phys. JETP **44**, 443 (1976).

[6] E. Kuraev, L. Lipatov, and V. Fadin, Sov. Phys. JETP **45**, 199 (1977).

[7] Y. Balitskii and L. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[8] M. Ciafaloni, Nucl. Phys. **B 296**, 49 (1988).

[9] S. Catani, F. Fiorani, and G. Marchesini, Phys. Lett. **B 234**, 339 (1990).

[10] S. Catani, F. Fiorani, and G. Marchesini, Nucl. Phys. **B 336**, 18 (1990).

[11] G. Marchesini, Nucl. Phys. **B 445**, 49 (1995).

[12] L. Gribov, E. Levin, and M. Ryskin, Phys. Rep. **100**, 1 (1983).

[13] E. M. Levin, M. G. Ryskin, Y. M. Shabelski, and A. G. Shuvaev, Sov. J. Nucl. Phys. **53**, 657 (1991).

[14] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B 366**, 135 (1991).

[15] J. Collins and R. Ellis, Nucl. Phys. **B 360**, 3 (1991).

[16] J. C. Collins, D. E. Soper, and G. Sterman, Nucl. Phys. **B250**, 199 (1985).

[17] F. Landry, R. Brock, P. M. Nadolsky, and C. P. Yuan, Phys. Rev. **D67**, 073016 (2003). `hep-ph/0212159`.

[18] H. Jung and G. Salam, Eur. Phys. J. **C 19**, 351 (2001). hep-ph/0012143.

[19] H. Jung, *Un-integrated updfs in ccfm*, 2004. hep-ph/0411287.

[20] L. Lonnblad and M. Sjodahl, JHEP **05**, 038 (2005). `hep-ph/0412111`.

[21] B. Andersson, G. Gustafson, and J. Samuelsson, Nucl. Phys. **B467**, 443 (1996); B. Andersson, G. Gustafson, and H. Kharraziha, Phys. Rev. **D57**, 5543 (1998). `hep-ph/9711403`.

[22] H. Jung, K. Kutak, K. Peters, and L. Motyka, *Nonlinear gluon evolution and heavy quark production at the lhc*. These proceedings, 2005.

[23] K. Kutak and A. M. Stasto, Eur. Phys. J. **C41**, 343 (2005). `hep-ph/0408117`.

[24] M. Lublinsky, *Parameterization of the dipole cross section and updf.* `http://www.desy.de/~lublinm/`.

[25] M. A. Kimber, A. D. Martin, and M. G. Ryskin, Phys. Rev. **D63**, 114027 (2001). hep-ph/0101348.

[26] H. Jung, Comput. Phys. Commun. **143**, 100 (2002). hep-ph/0109102.

[27] H. Jung and G. P. Salam, Eur. Phys. J. **C19**, 351 (2001). hep-ph/0012143.

[28] G. Watt, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C31**, 73 (2003). hep-ph/0306169.

[29] G. Watt, A. D. Martin, and M. G. Ryskin, Phys. Rev. **D70**, 014012 (2004). hep-ph/0309096.

[30] J. C. Collins and X. Zu, JHEP **03**, 059 (2005). hep-ph/0411332.

[31] J. Kwiecinski, A. D. Martin, and J. J. Outhwaite, Eur. Phys. J. **C9**, 611 (1999). hep-ph/9903439.

[32] J. C. Collins, Acta Phys. Polon. **B34**, 3103 (2003). hep-ph/0304122.

[33] J. Bartels, Z. Phys. **C60**, 471 (1993);
J. Bartels, Phys. Lett. **B298**, 204 (1993);
E. M. Levin, M. G. Ryskin, and A. G. Shuvaev, Nucl. Phys. **B387**, 589 (1992).

[34] I. Balitsky, Nucl. Phys. **B 463**, 99 (1996).

[35] Y. V. Kovchegov, Phys. Rev. **D 60**, 034008 (1999).

[36] M. Lublinsky, E. Gotsman, E. Levin, and U. Maor, Nucl. Phys **A 696**, 851 (2001).

[37] E. Gotsman, E. Levin, M. Lublinsky, and U. Maor, Eur. Phys. J. **C 27**, 411 (2003).

[38] M. Lublinsky, Eur. Phys. J. **C 21**, 513 (2001).

[39] A. M. Stasto, K. Golec-Biernat, and J. Kwiecinski, Phys. Rev. Lett. **86**, 596 (2001).

[40] K. Golec-Biernat and M. Wusthoff, Phys. Rev. **D 59**, 014017 (1999).

[41] J. Forshaw, *Diffractive higgs production: theory*. These proceedings.

[42] A. D. Martin, M. G. Ryskin, and T. Teubner, Phys. Rev. **D55**, 4329 (1997). hep-ph/9609448;
A. D. Martin, M. G. Ryskin, and T. Teubner, Phys. Lett. **B454**, 339 (1999). hep-ph/9901420;
A. D. Martin, M. G. Ryskin, and T. Teubner, Phys. Rev. **D62**, 014022 (2000). hep-ph/9912551.

[43] A. G. Shuvaev, K. J. Golec-Biernat, A. D. Martin, and M. G. Ryskin, Phys. Rev. **D60**, 014015 (1999). hep-ph/9902410.

[44] A. D. Martin and M. G. Ryskin, Phys. Rev. **D64**, 094017 (2001). hep-ph/0107149.

[45] M. Diehl, Phys. Rept. **388**, 41 (2003). hep-ph/0307382.

[46] T. Teubner. To appear in Proceedings of DIS05.

[47] H. C. A. Aktas et al., *Elastic J/Ψ production at HERA*. hep-ex/0510016, Submitted to Eur Phys J.

[48] ZEUS Collaboration, S. Chekanov *et al.*, Nucl. Phys. **B695**, 3 (2004). hep-ex/0404008.

[49] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C24**, 345 (2002). `hep-ex/0201043`; ZEUS Collaboration, J. Breitweg *et al.*, Eur. Phys. J. **C6**, 603 (1999). `hep-ex/9808020`; ZEUS Collaboration, J. Breitweg *et al.*, Z. Phys. **C75**, 215 (1997). `hep-ex/9704013`.

[50] L. Motyka, *Talk at the workshop meeting 7–21 January 2005, CERN.* http://agenda.cern.ch/fullAgenda.php?ida=a045699; M. Diehl, *Talk at the workshop meeting 21–24 March 2005, DESY*. http://agenda.cern.ch/fullAgenda.php?ida=a05611.

Resummation

A. Banfi[1]*, G. Corcella*[2]*, M. Dasgupta*[3]*, Y. Delenda*[3]*, G.P. Salam*[4] *and G. Zanderighi*[5].
[1] University of Cambridge, Madingley Road, Cambridge CB3 0HE ,U.K
[2] Department of Physics, Theory Division, CH-1211 Geneva 23, Switzerland.
[3] University of Manchester, Oxford Road, Manchester, M13 9PL, U.K
[4] LPTHE, Universities of Paris VI and VII and CNRS, 75005, Paris, France.
[5] Fermilab, Batavia, IL 60510, USA.

Abstract
We review the work discussed and developed under the topic "Resummation" at Working Group 2 "Multijet final states and energy flow" , of the HERA-LHC Workshop. We emphasise the role played by HERA observables in the development of resummation tools via, for instance, the discovery and resummation of non-global logarithms. We describe the event-shapes subsequently developed for hadron colliders and present resummed predictions for the same using the automated resummation program CAESAR. We also point to on-going studies at HERA which can be of benefit for future measurements at hadron colliders such as the LHC, specifically dijet E_t and angular spectra and the transverse momentum of the Breit current hemisphere.

1 Introduction

Resummed calculations are an invaluable tool, both for the understanding of perturbative QCD dynamics at all orders as well as for extracting, as accurately as possible, QCD parameters such as the strong coupling, quark masses and parton distribution functions. These parameters, which cannot be directly computed from QCD perturbation theory itself, will be vital inputs in new physics searches at the LHC. Moreover, resummed expressions are also an important stepping stone to probing observable distributions in regions where non-perturbative power corrections make a significant contribution. In this region one may expect a smearing of the resummed perturbative result with a non-perturbative function (for which one can adopt, for example, a renormalon-inspired model), and the resulting spectrum can be confronted with data to test our understanding of non-perturbative dynamics. In all these aspects, HERA data and observables have played an important role (sometimes significantly underrated in the literature) in furthering our knowledge, without which accurate studies of several observables at the LHC would simply not be possible.

A concrete example of HERA's important role in this regard is the case of event shape distributions [1, 2], theoretical studies of which led to the finding of non-global single-logarithmic [3] effects (discussed in more detail below). Prior to these studies it was widely believed that the HERA distributions, measured in the current hemisphere Breit frame, were trivially related to their e^+e^- counterparts. Had such ideas, based on independent soft gluon emission by the hard partons, been applied directly to similar variables at the LHC, such as energy flows away from jets, the accuracy of theoretical predictions would have been severely compromised leading almost certainly to erroneous claims and conclusions.

Another area where HERA has played a vital role is in the testing of renormalon inspired models for power corrections, most significantly the dispersive approach [4] to $1/Q$ power corrections, tested against HERA event-shape distributions and mean-values [5]. The fact that HERA data seem to confirm such models , where one can think of the power corrections as arising from the emission of a gluon with transverse momentum $\mathcal{O}(\Lambda_{\rm QCD})$, is significant for the LHC. This is because the agreement of the renormalon model with data demonstrates that the presence of initial state protons does not affect significantly the form of $1/Q$ corrections. It thus sets limits on the additional non-perturbative contribution that may potentially be generated by the flight of struck partons through the proton cloud, which therefore does not appear to be significant. Once again it is accurate resummed predictions [6] which have allowed us access to the non-perturbative domain hence strengthening our understanding of power corrections.

One important aspect of resummed studies, till date, is that stringent comparisons of next-to–leading logarithmic resummed predictions with data have only been carried out in cases involving observables that vanish in the limit of two hard partons. Prominent examples reflecting the success of this program are provided by $e^+e^- \to 2$ jet event shapes and DIS (1+1) jet event shapes as well as Drell-Yan vector boson transverse momentum spectra at hadron colliders. At the LHC (and hadron colliders in general) one already has two hard incoming partons and any observable dealing with final state jet production would take us beyond the tested two hard parton situation. Thus dijet event shapes at hadron colliders (discussed in detail later), which involve much more complicated considerations as far as the resummation goes, represent a situation where NLL resummations and power corrections are as yet untested. Bearing in mind the hadronic activity due to the underlying event at hadron colliders, it is important to test the picture of resummations and power corrections for these multiparton event shapes in cleaner environments. Thus LEP three-jet event shapes and similar $2 + 1$ jet event shapes at HERA become important to study in conjunction with looking at resummation of event shapes at hadron colliders.

Predictions for several LEP and HERA three-jet event shapes already exist (see e.g [7] and for a full list of variables studied Ref. [8]) and at this workshop a prominent development presented was the proposal of several dijet event-shapes in hadron-hadron collisions and the resummed predictions for their distributions [9].

Existing HERA data can also be usefully employed to study soft gluon radiation dynamics from multi-hard–parton ensembles, in the study of dijet E_t and angular spectra. These quantities are somewhat different from event shapes since one defines observables based on aggregate jet-momenta and angles rather than directly constructing them from final-state hadron momenta. Examples are the transverse energy, E_t, mismatch between the leading E_t jets in dijet production and the azimuthal correlation between jets ϕ_{jj}, once again refering to the highest E_t jets in dijet production. For the former quantity there are no direct experimental data as yet, but it is simply related to the dijet total rate in the region of symmetric E_t cuts for which data does exist . For the latter quantity similarly there are direct experimental data [10]. These observables have smaller hadronisation corrections scaling as $1/Q^2$ rather than $1/Q$ as for most event shapes. They thus offer a good opportunity to test the NLL perturbative predictions alone without necessarily probing non-perturbative effects at the same time [1].

At this workshop developments were reported on extending existing calculations [11] for cone dijets, to different jet algorithms, such as the k_t algorithm, comparing to fixed order estimates and performing the leading order matching. Once the HERA data has been well described similar studies can be carried out for hadron–hadron dijets. In fact predictions already exist for hadron-hadron dijet masses near threshold [12] but are not in a form conducive to direct comparisons with data containing neither the jet algorithms in the form actually employed in experiment, nor the matching to fixed order. However these calculations provided a useful starting point for the calculations presented here, which should eventually lead to direct comparisons with data.

Another area where HERA may play an important role is to establish whether unaccounted for small x effects may be significant in comparing theoretical resummations for e.g. vector boson p_t spectra with experimental data. It has been suggested that a non-perturbative intrinsic k_t, growing steeply with x, is required to accomodate HERA data for semi-inclusive DIS processes [13]. When this observation is extrapolated to the LHC kinematical region there is apparently significant small x broadening in the vector boson p_t distribution. Similar effects may well arise in the case of the Higgs boson too. However DIS event shape studies in the Breit current hemisphere [6] apparently do not acquire such corrections since they are well described by conventional NLL resummations supported by dispersive

[1]Although effects to do with intrinsic k_t will eventually have to be accounted for similar to the case of Drell-Yan vector boson p_t spectra.

power corrections [5], which are x independent [2]. However there are some important caveats:

- Unlike vector boson p_t spectra, event shapes receive $1/Q$ hadronisation corrections unrelated to intrinsic k_t. These could mask $1/Q^2$ terms originating from intrinsic k_t which may yet contain the x dependence in question.
- It has already been observed that including H1 data for $Q < 30$ GeV does spoil somewhat the agreement with the dispersive prediction of universal power corrections to event shapes [6]. The origin of this effect could well be extra non-perturbative k_t broadening related to the effects described above for vector boson p_t.

To get to the heart of this matter a useful variable that has been suggested (see plenary talk by G. Salam at the first meeting of this workshop) is the modulus of the vector transverse momentum $\sum_{i\in H_c} \vec{k}_{t,i}$ of the current hemisphere in the DIS Breit frame. This quantity is simply related to the Drell-Yan p_t spectra and comparing theoretical predictions, presented here, with data from HERA should help to finalise whether additional small-x enhanced non-perturbative terms are needed to accomodate the data. We begin by first describing the results for hadron-hadron event shape variables, discussed by G. Salam at this workshop. Then we describe the progress in studying dijet E_t and angular spectra (presented by M. Dasgupta and G. Corcella at the working group meetings). Finally we mention the results obtained thus far, for the Q_t distribution of the current hemisphere and end with a look at prospects for continuing phenomenology at HERA, that would be of direct relevance to the LHC.

2 Event shapes for hadron colliders

Event shape distributions at hadron colliders, as has been the case at LEP and HERA, are important collinear and infrared safe quantities, that can be used as tools for the extraction of QCD parameters, for instance α_s, by comparing theory and data. In contrast however to more inclusive sources of the same information (e.g the ratio of 3 jet to 2 jet rates), event shape distributions provide a wealth of other information, some of which ought to be crucial in disentangling and further understanding the different physics effects, relevant at hadron colliders. These range from fixed-order predictions to resummations, hadronisation corrections and, in conjunction with more detailed studies assesing the structure of, and role played by, the underlying event (beam fragmentation).

Until recently there have only been limited experimental studies of jet-shapes at hadron colliders [15] and no resummed theoretical predictions for dijet shape variables at hadron colliders. Rapid recent developments (see Ref. [9] and references therin) in the field of perturbative resummations have now made theoretical estimates possible for a number of such distributions, introduced in [9] which we report on below.

The three main theoretical developments that have led to the studies of Ref. [9] are:

- Resummation for hadron-hadron dijet observables depends on describing multiple soft gluon emission from a system of four hard partons. The colour structure of the resulting soft anomalous dimensions is highly non-trivial and was explicitly computed by the Stony Brook group in a series of papers (see e.g [12] and references therin).
- The discovery of non-global observables [3]. The realisation that standard resummation techniques based on angular ordering/independent-emission of soft gluons by the hard-parton ensemble, are not valid for observables that are sensitive to emissions in a limited angular range, has led to the introduction of observables that are made global by construction. This means that one can apply the technology developed by the Stony-Brook group to obtain accurate NLL predictions for these observables, without having to resort to large N_c approximations.
- The advent of automated resummation [16]. The development of generalised resummation formulae and powerful numerical methods to determine the parameters and compute the functions

[2] An exception is the jet broadening [14] but the x dependence there is of an entirely different origin and nature.

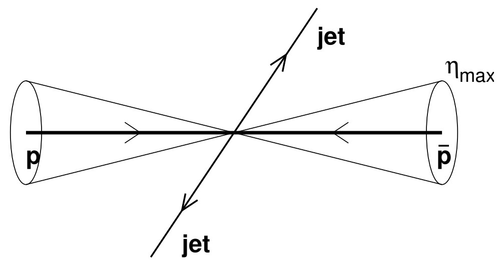


Fig. 1: Cut around the beam direction beyond rapidity $\eta_{\max}$ corresponding to the maximum rapidity reach of the detectors.

thereof, has made it possible to study several variables at once rather than having to perform copious, and in some cases previously unfeasible, calculations for each separate observable.

We now discuss the different types of variables defined and resummed in [9]. The first issue one has to deal with is the fact that experimental detectors have a limited rapidity range, which can be modeled by a cut around the beam direction.

This cut would then correspond to a position in rapidity of the edge of the most forward detector with momentum or energy resolution and the relevant values of the maximum rapidity for measurements is 3.5 units at the Tevatron and 5 units at the LHC. One may then worry about gluon emissions beyond this rapidity (i.e. inside the beam cut, see Fig. 1) that emit softer gluons into the allowed rapidity range, outside the cones depicted in Fig. 1. Such a configuration would of course render the observable non-global.

To get around this potential problem, one can employ an idea suggested for 3-jet observables such as out-of–plane momentum flows in hadron-hadron collisions [17], which helps side-step the issue of non-globalness. We note that all the observables studied here have the following functional dependence on a soft emission, k, collinear to a given hard leg [3] (common to all event shapes studied here and in other processes)

$$V(\tilde{p}, k) = d\left(\frac{k_t}{Q}\right)^a e^{-b\eta} g(\phi), \tag{1}$$

where k_t , η and ϕ are measured wrt a given hard leg and $\tilde{p}$ represent the set of hard parton momenta including recoil against k while Q is the hard-scale of the process. We are particularly interested in emissions soft and collinear to the beam (incoming) partons. Then an emission beyond the maximum detector rapidity $\eta \geq \eta_{\max}$ corresponds to at most a contribution to the observable $V \sim e^{-(a+b_{\min})\eta_{\max}}$ with $b_{\min} = \min(b_1, b_2)$ and b_1 and b_2 are the values of b associated with collinear emission near beam-partons 1 and 2.

If one then choses to study the observable over a range of values such that

$$L \leq (a + b_{\min})\eta_{\max}, \; L \equiv \ln 1/V, \tag{2}$$

then emissions more forward than $\eta_{\max}$ do not affect the observable in the measured range of values. One can thus include the negligible contribution from this region and do the calculation *as if the observable were global*, ignoring the cut around the beam. Including the region beyond $\eta_{\max}$ does not alter the NLL resummed result in the suitably selected range Eq. 2.

The price one has to pay is to limit the range of the study of the observable V, such that emissions beyond $\eta_{\max}$ make a negligible contribution. As we will mention later this is a more significant restriction for some variables compared to others (depending on the parameters a and b) but a range of study can always be found over which the observable can be treated as global.

[3] In general the values of parameters d, a, b and the function g depend on the observable considered. For more details and constraints on the various parameters that ensure globalness and infrared and collinear safety etc., see Ref. [16].

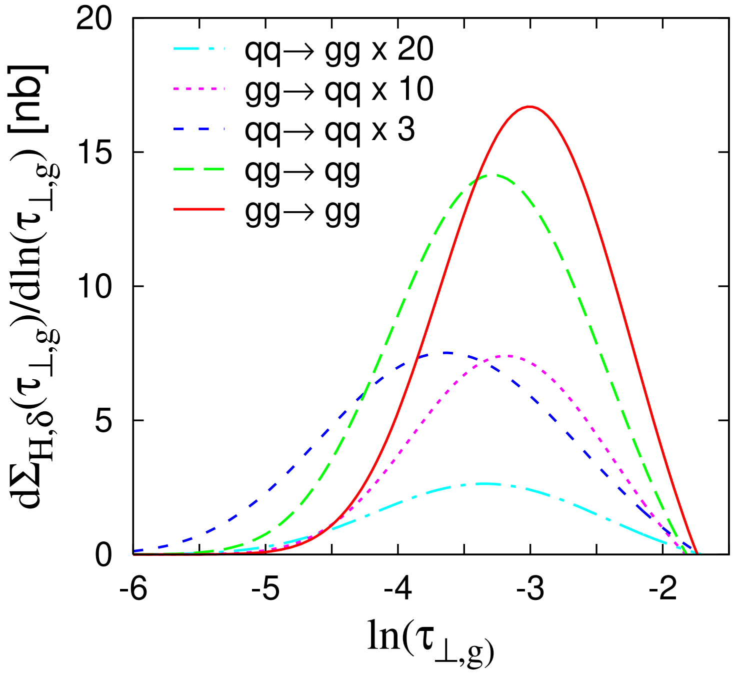


Fig. 2: The global transverse thrust distribution with the contribution from different partonic channels explicitly displayed.

2.1 Global event shapes

With the above caveat in place several variables can be safely studied (treated as global) over a wide range of values. An explicit example is the *global transverse thrust* defined as:

$$T_{\perp,g} \equiv \max_{\vec{n}_T} \frac{\sum_i |\vec{q}_{\perp i} \cdot \vec{n}_T|}{\sum_i q_{\perp i}}\,, \qquad \tau_{\perp,g} = 1 - T_{\perp,g}\,, \tag{3}$$

where the thrust axis $\vec{n}_T$ is defined in the plane transverse to the beam axis. The probability $P(v)$, that the event shape is smaller than some value v behaves as:

$$P(v) = \exp\left[-G_{12}\frac{\alpha_s}{2\pi}L^2 + \cdots\right], \; L = \ln 1/v, \tag{4}$$

with $G_{12} = 2C_B + C_J$, where C_B and C_J represent the total colour charges of the beam and jet (outgoing) partons. The above represents just the double-logarithmic contribution. The full result with control of up to next-to–leading single-logarithms in the exponent is considerably more complicated. It contains both the Stony-Brook colour evolution matrices as well as multiple emission effects (generated by phase-space factorisation). The automated resummation program CAESAR [16] is used to generate the NLL resummed result shown in Fig. 2. In this particular case the effect of the cut around the beam direction can be ignored for values $\tau_{\perp,g} \geq 0.15e^{-\eta_{\max}}$. We note that it is advisable to leave a safety margin between this value and the values included in measurement.

Other global variables studied include the *global thrust minor* and the three jet-resolution threshold parameter y_{23}. For detailed definitions and studies of these variables, the reader is refered to [9].

We shall now proceed to look at two different ways of defining event shapes in a given central region, which on its own would lead to non-globalness, and then adding terms that render them global.

2.2 Forward suppressed observables

Here we shall examine event shapes defined in a chosen central region $\mathcal{C}$ well away from the forward detector edges.

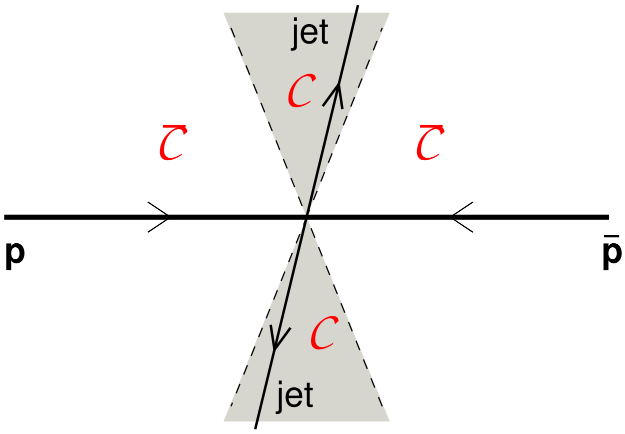


Fig. 3: Figure depicting the central region marked $\mathcal{C}$, containing the two hard jets.

First we define central $\perp$ momentum, and rapidity:

$$Q_{\perp,\mathcal{C}} = \sum_{i\in\mathcal{C}} q_{\perp i}\,, \quad \eta_{\mathcal{C}} = \frac{1}{Q_{\perp,\mathcal{C}}} \sum_{i\in\mathcal{C}} \eta_i\, q_{\perp i} \tag{5}$$

and an *exponentially suppressed forward term,*

$$\mathcal{E}_{\bar{\mathcal{C}}} = \frac{1}{Q_{\perp,\mathcal{C}}} \sum_{i\notin\mathcal{C}} q_{\perp i}\, e^{-|\eta_i-\eta_{\mathcal{C}}|}\,. \tag{6}$$

Then we can define an event shape in the central region $\mathcal{C}$[4] which on its own would be non-global since we measure emissions just in $\mathcal{C}$. The addition of $\mathcal{E}_{\bar{\mathcal{C}}}$ to the event-shape renders the observable global as this term includes suitably the effect of emissions in the remaining region $\bar{\mathcal{C}}$. The exponential suppression of the added term reduces sensitivity to emissions in the forward region which in turn reduces the effect of the beam cut $\eta_{\max}$ considerably, pushing its impact to values of the observable where the shape cross-section is highly suppressed and thus too small to be of interest.

The event shapes are constructed as described stepwise below:

- Split $\mathcal{C}$ into two pieces: *Up, Down*
- Define *jet masses* for each

$$\rho_{X,\mathcal{C}} \equiv \frac{1}{Q^2_{\perp,\mathcal{C}}} \Big(\sum_{i\in\mathcal{C}_X} q_i\Big)^2\,, \qquad X = U, D\,. \tag{7}$$

Define sum and heavy-jet masses

$$\rho_{S,\mathcal{C}} \equiv \rho_{U,\mathcal{C}} + \rho_{D,\mathcal{C}}\,, \qquad \rho_{H,\mathcal{C}} \equiv \max\{\rho_{U,\mathcal{C}}, \rho_{D,\mathcal{C}}\}\,. \tag{8}$$

Define global extension, with extra forward-suppressed term

$$\rho_{S,\mathcal{E}} \equiv \rho_{S,\mathcal{C}} + \mathcal{E}_{\bar{\mathcal{C}}}\,, \qquad \rho_{H,\mathcal{E}} \equiv \rho_{H,\mathcal{C}} + \mathcal{E}_{\bar{\mathcal{C}}}\,. \tag{9}$$

- Similarly: *total and wide jet-broadenings*

$$B_{T,\mathcal{E}} \equiv B_{T,\mathcal{C}} + \mathcal{E}_{\bar{\mathcal{C}}}\,, \qquad B_{W,\mathcal{E}} \equiv B_{W,\mathcal{C}} + \mathcal{E}_{\bar{\mathcal{C}}}\,. \tag{10}$$

At the double-log level the results assume an identical form to Eq. 4 with G_{12} representing a combination of total incoming (beam) and outgoing (jet) parton colour charges [9]. The full NLL resummed results have a substantially more complex form and results from CAESAR [16] are plotted in Fig. 4.

[4]There is considerable freedom on the choice of the central region. For instance this could be a region explicitly delimited in rapidity or the two hard jets themselves.

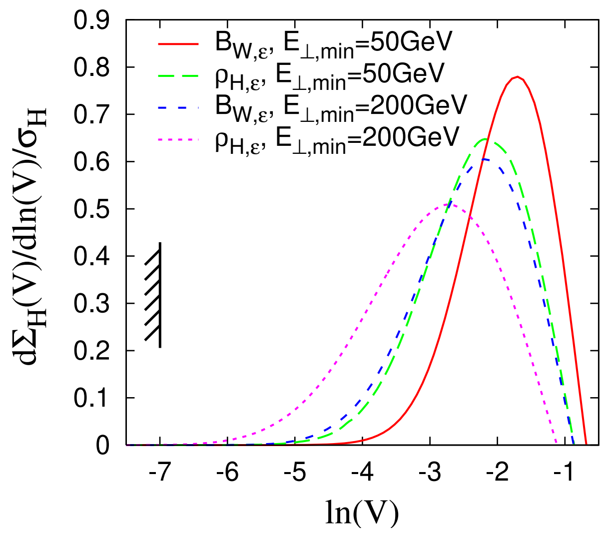


Fig. 4: NLL resummed predictions from CAESAR for the heavy jet-mass and the wide jet-broadening with the minimum jet transverse energy $E_{\perp,\min}$ values of 50 and 200 GeV as shown.

2.3 Indirectly global recoil observables

Here we study observables that are defined exclusively in terms of particles in the central region but are global. Such observables are already familiar from HERA studies. As an example, although the current-jet broadening wrt the photon axis of the DIS Breit frame involves only particles that enter the current hemisphere, the current quark acquires transverse momentum by *recoil* against remnant hemisphere particles. This recoil means that the observable is indirectly sensitive to emissions in the remnant hemisphere which makes the observables global.

To construct similar observables in the hadron-hadron case we observe that by momentum conservation, the following relation holds :

$$\sum_{i\in\mathcal{C}} \vec{q}_{\perp i} = -\sum_{i\notin\mathcal{C}} \vec{q}_{\perp i} \tag{11}$$

which relates the sum of transverse momenta in $\mathcal{C}$ to that in the complementary region. Then the central particles can be used to define a recoil term:

$$\mathcal{R}_{\perp,\mathcal{C}} \equiv \frac{1}{Q_{\perp,\mathcal{C}}} \left| \sum_{i\in\mathcal{C}} \vec{q}_{\perp i} \right| , \tag{12}$$

which contains an indirect dependence on non-central emissions.

Now we can define event shapes explicitly in terms of central particle momenta in $\mathcal{C}$. Examples are the recoil jet-masses and broadenings

$$\rho_{X,\mathcal{R}} \equiv \rho_{X,\mathcal{C}} + \mathcal{R}_{\perp,\mathcal{C}} , \qquad B_{X,\mathcal{R}} \equiv B_{X,\mathcal{C}} + \mathcal{R}_{\perp,\mathcal{C}} , \ldots \tag{13}$$

It is clear that since these observables are defined in terms of central particles alone, the cut around the beam direction is not an issue here. There is however another potential problem. Due to the addition of the recoil term we lose direct exponentiation of the result in variable space. Exponentiation to NLL accuracy only holds in impact-parameter or b space .

The physical effect in question here is similar to Drell-Yan Q_T spectra where there are two competing mechanisms that lead to a given small Q_T, Sudakov suppression of soft emissions and vectorial cancellation between harder emissions. Where the latter effect takes over (typically in the region where

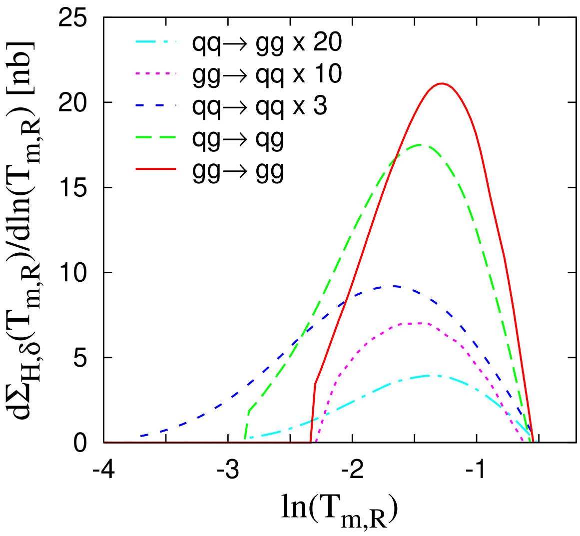


Fig. 5: The recoil thrust minor as predicted by CAESAR, with a cutoff before the divergence. Only a small fraction of the cross-section is beyond the cutoff.

single-logs are large $\alpha_s L \sim 1$) we get a breakdown of the Sudakov result generated by CAESAR. This result is of the general form:

$$P(V) = e^{Lg_1(\alpha_s L)+g_2(\alpha_s L)+\cdots}. \tag{14}$$

The result for recoil observables produced by CAESAR will contain a divergence in the single-log function g_2 and is cut before the divergence. Again for some variables this cut is at a position that significantly reduces the range of possible phenomenological studies. For other variables the divergence is at values of the observable that are sufficiently small so that only a few percent of the cross-section is beyond the cutoff. An example of the former is the recoil transverse thrust where 15% of the cross-section lies beyond the cut-off. For the recoil thrust minor, in contrast, the cutoff has only a moderate effect and much less of the cross-section is cutoff, due to the divergence in g_2.

Below we present a table of the different event shapes mentioned here and the impact of the two main limitations we discussed, the beam-cut $\eta_{\max}$ and the breakdown of resummation due to divergences of g_2. Additionally we mention the expected impact of hadronisation corrections (not yet computed in full) on the different observables as well as the form of the estimated contribution from the underlying event.

Event-shape	Impact of $\eta_{\max}$	Resummation breakdown	Underlying Event	Jet hadronisation
$\tau_{\perp,g}$	tolerable*	none	$\sim \eta_{\max}/Q$	$\sim 1/Q$
$T_{m,g}$	tolerable	none	$\sim \eta_{\max}/Q$	$\sim 1/(\sqrt{\alpha_s}Q)$
y_{23}	tolerable	none	$\sim \sqrt{y_{23}}/Q^*$	$\sim \sqrt{y_{23}}/Q^*$
$\tau_{\perp,\mathcal{E}}$, $\rho_{X,\mathcal{E}}$	negligible	none	$\sim 1/Q$	$\sim 1/Q$
$B_{X,\mathcal{E}}$	negligible	none	$\sim 1/Q$	$\sim 1/(\sqrt{\alpha_s}Q)$
$T_{m,\mathcal{E}}$	negligible	serious	$\sim 1/Q$	$\sim 1/(\sqrt{\alpha_s}Q)$
$y_{23,\mathcal{E}}$	negligible	none	$\sim 1/Q$	$\sim \sqrt{y_{23}}/Q^*$
$\tau_{\perp,\mathcal{R}}$, $\rho_{X,\mathcal{R}}$	none	serious	$\sim 1/Q$	$\sim 1/Q$
$T_{m,\mathcal{R}}$, $B_{X,\mathcal{R}}$	none	tolerable	$\sim 1/Q$	$\sim 1/(\sqrt{\alpha_s}Q)$
$y_{23,\mathcal{R}}$	none	intermediate*	$\sim \sqrt{y_{23}}/Q^*$	$\sim \sqrt{y_{23}}/Q^*$

In the above the entries marked * are subject to uncertainty at present.

Further work is needed before the resummed expressions presented here can be compared with data including the matching to fixed order and computation of the power corrections for the various observables. This is currently in progress.

Having discussed the hadron-hadron event shapes we now move on to describe resummed studies concerning dijet production at HERA which can also be straightforwardly extended to hadron-hadron collisons.

3 Dijet p_t and angular spectra

It has been known for some time that dijet total rates cannot be predicted within fixed-order QCD if symmetric cuts are applied to the two highest p_t dijets [18]. While it was understood that the problems are to do with constraints on soft gluon emission, the exact nature of this constraint was only made clear in Ref. [11]. There it was pointed out that there are large double logarithms (aside from single logarithms and less singular pieces) in the slope $\sigma'(\Delta)$ of the total rate, as a function of Δ the difference in minimum p_t values of the two highest p_t jets. These logarithms were resummed and it was shown that the slope of the total rate $\sigma' \to 0$ as $\Delta \to 0$. This leads to a physical behaviour of the total rate as reflected by the data [10].

To perform the comparison to data accurately however, requires two improvements to be made to the calculations of Ref. [11]. Firstly the exact same jet algorithm has to be employed in the theoretical calculations and experimental measurements. The current algorithm used by H1 and ZEUS experiments is the inclusive k_t algorithm. At hadron colliders variants of the cone algorithm are used and it is in fact a cone algorithm that was employed in Ref. [11]. However the details of the calculation need to be ammended to define the cones in η, ϕ space as is done experimentally and calculations concerning this were presented at the working group meeting. The second important step is matching to fixed order estimates. We report below on the leading order matching to DISENT [19] while a full NLO matching is still awaited.

We also introduce and study two variables of related interest, the first is the difference in p_t, between the highest p_t jets $\Delta p_{t,jj} = p_{t1} - p_{t2}$ (note that here we talk about the p_t difference rather than the difference in the minimum $E_{\rm cut}$, that we mentioned earlier. The resummation of this distribution $\frac{d\sigma}{d\Delta p_{t,jj}}$ is essentially identical to that carried out in Ref. [11], except that here we compute the next-to–leading logarithms in different versions of the jet algorithm, which should help with direct experimental comparsions. We also perform the leading-order matching to DISENT.

Having developed the calculational techniques for $d\sigma/d\Delta p_{t,jj}$ it is then straightforward to generate the results for the distribution in azimuthal angle between jets $d\sigma/d\phi_{jj}$ which requires resummation in the region $\phi_{jj} = \pi$. These distributions have been measured at HERA and the Tevatron (most recently by the D0 collaboration). Comparing the resummation with data would represent an interesting challenge for the theory insofar as the status of resummation tools is concerned, and is potentially very instructive.

3.1 The $\Delta p_{t,jj}$ and ϕ_{jj} distributions

We shall consider dijet production in the DIS Breit frame. For the jet definition we can consider either an η, ϕ cone algorithm (such as the infrared and collinear safe midpoint cone algorithm) or the inclusive k_t algorithm. We shall point out to what level the two algorithms would give the same result and where they can be expected to differ. We shall use a four-vector recombination scheme where the jet four-momentum is the sum of individual constituent hadron four-momenta. We also impose cuts on the highest p_t jets such that $|\eta_{1,2}| \leq 1$ and $p_{t1,t2} \geq E_{\rm min}$.

We then consider the quantity $\Delta p_{t,jj} = p_{t1} - p_{t2}$ which vanishes at Born order and hence the distribution at this order is just $\frac{d\sigma}{dp_{t,jj}} \propto \delta(p_{t,jj})$.

Beyond leading order the kinematical situation in the plane normal to the Breit axis is represented as before [11]:

$$
\begin{aligned}
\vec{p_{t1}} &= p_{t1}(1,0) && (15)\\
\vec{p_{t2}} &= p_{t2}\left(\cos(\pi \pm \epsilon), \sin(\pi \pm \epsilon)\right) && (16)\\
\vec{k_t} &= k_t\left(\cos\phi, \sin\phi\right) && (17)
\end{aligned}
$$

Thus we are considering a small deviation from the Born configuration of jets back-to–back in azimuth, induced by the presence of a soft gluon with transverse momentum $k_t \ll p_{t1,t2}$ (which is not recombined by the algorithm with either hard parton) and with azimuthal angle ϕ. In the above ϵ represents the recoil angle due to soft emission. We then have

$$\Delta p_{t,jj} = |p_{t1} - p_{t2}| \approx |k_t \cos\phi|, \tag{18}$$

which accounts for the recoil ϵ to first order and hence is correct to NLL accuracy. Thus for the emission of several soft gluons we have the p_t mismatch given by

$$\Delta p_{t,jj} = |\sum_{i\notin j} k_{xi}|, \tag{19}$$

where k_x denotes the single component of gluon transverse momentum, along the direction of the hard jets, which are nearly back-to–back in the transverse plane. The sum includes only partons not merged by the algorithm into the highest E_t jets.

Similarly for the dijet azimuthal angle distribution[5], we have :

$$\pi - \phi_{jj} \approx \frac{1}{p_t}|\sum_{i\notin j} k_{yi}|. \tag{20}$$

where ϕ_{jj} is the azimuthal angle between the two highest p_t jets. Note that in the above we have set $p_{t1} = p_{t2} = p_t$ since we are considering a small deviation from the Born configuration and this approximation is correct to NLL accuracy. We also introduced k_y, the component of soft gluon momentum normal to the jet axis in the transverse plane.

In either of the above two cases, i.e the $\Delta p_{t,jj}$ or ϕ_{jj} distributions, an identical resummation is involved , due to the similar role of soft partons not recombined into jets. Henceforth we shall proceed with just the $\Delta p_{t,jj}$ resummation results, it being understood that similar considerations apply to ϕ_{jj} in the region $\phi_{jj} \sim \pi$.

Assuming independent emission of soft gluons by the hard three-parton system (the incoming parton and the two outgoing partons that initiate the dijets) and factorising the phase-space Eq. 19 as below[6]:

$$\Theta\left(\Delta p_{t,jj} - |\sum_{i\notin j} k_{x,i}|\right) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{db}{b}\sin(b\Delta p_{t,jj})\prod_{i\notin j} e^{ibk_{xi}}, \tag{21}$$

the resummed result for the $\Delta p_{t,jj}$ distribution can be expressed as

$$\frac{d^3\sigma}{dx dQ^2 d\Delta p_{t,jj}}(E_{\min}, \Delta p_{t,jj}) = \sum_{\delta=q,g}\int_x^1 \frac{d\xi}{\xi}\int_0^1 dz \sum_{a=T,L} F_a(y) C_\delta^a(\xi, z, E_{\min}) w_\delta(Q, \Delta p_{t,jj}). \tag{22}$$

In the above ξ and z are phase-space variables that parametrise the Born dijet configuration, $F_{a=T,L}$ denotes the $y = Q^2/xs$ dependence associated to the transverse or longitudinal structure function while C^a is the Born matrix-element squared. The function w represents the result of resummation.

[5]Note that the kinematical relations we derive here would be equally valid for dijets produced in hadron-hadron collisions at the Tevatron or LHC and just the dynamics of multisoft gluon emission would be more complex.

[6]We compute here the cross-section for the observable to be less than $\Delta p_{t,jj}$ from which we can easily obtain the corresponding distribution.

The resummed expression w requires some explanation. Its form is as follows

$$w_\delta(p_{t,jj}) = \int_0^\infty \frac{db}{b}\sin(b\Delta p_{t,jj})\exp[-R_\delta(b)]\mathcal{S}(b)q_\delta\left(x/\xi, 1/b^2\right). \tag{23}$$

Note the fact that the exponentiation holds only in b space where b is the impact parameter. The function $R(b)$ (we ignore the subscript δ which describes either incoming quarks or gluons) is the Sudakov exponent which can be computed up to NLL accuracy,

$$R(b) = Lg_1(\alpha_s L) + g_2(\alpha_s L),\ L \sim \ln(bQ). \tag{24}$$

while $S(b)$ is the non-global contribution that arises from soft partons inside the jet emitting outside it. q_δ is the incoming quark or gluon density and its scale depends on the variable b. The functions g_1 and g_2 are the leading-logarithmic and next-to–leading logarithmic resummed quantities.

For the leading logarithms g_1 and a subset of next-to–leading logarithms g_2, generated essentially by exponentiation of the single-log result in b space, the cone and inclusive k_t algorithms would give the same result, which we have computed. Starting from terms that begin with $\alpha_s^2 \ln^2 b$ in g_2 (specifically two soft wide-angle gluons), the following two effects become important:

- For cone algorithms the implementation of the split/merge stage affects the g_2 piece. Present calculations [11] are valid to NLL accuracy if *all* the energy shared by overlapping jets is given to the jet that would have highest p_t. Note that this is different from merging the overlapping jets themselves. If other merging procedures are used the calculation becomes more complex but is still tractable.
- For the k_t algorithm it is just being realised that running the algorithm generates terms that start at $\alpha_s^2 \ln^2 b$ in the exponent, which are not correctly treated by naive Sudakov exponentiation. These terms, which are generated by the clustering procedure, can also be numerically accounted for in our case, but this is work in progress.

The effects that we mention above cause a similar impact on the final result as the non-global term $S(b)$ which was shown to be at around the 10% level in Ref. [11]. Hence the current results for the k_t algorithm that do not account for the recently found additional terms and only approximately for the non-global logs, can be expected to change by around 10% when these effects will be included correctly.

We present in Fig. 6 preliminary results for the $\Delta p_{t,jj}$ distribution matched to the leading order DISENT prediction, using the k_t algorithm. The matching at present combines quark and gluon channels wheras ideally one would like to separate the incoming quark and gluon channels with the right weights ($\mathcal{O}(\alpha_s)$ coefficient functions). This would be possible if, for instance, there was parton flavour information explicit in the fixed order codes, a limitation of the fixed-order codes that needs to be addressed also for hadron-hadron event shapes to be matched to NLO predictions.

We also present a comparison with HERWIG [20] results on the same quantity. The variable X in the figure merely refers to the effect of using the jet p_t as the hard scale rather than the photon virtuality Q^2, formally a NNLL effect. It is amusing to note the very good agreement of the resummation with HERWIG but not too much can be read into it at this stage. Given the minor role of non-global effects we would expect HERWIG and our predictions to indeed have a broad resemblence. However we should mention that the resummed result in Fig. 6 is at present subject to change pending proper inclusion of non-global logs and the effect of independent soft emission at large angles. The latter is partly included in the results shown, through exponentiation of the one-gluon result as we pointed out before, but the clustering procedure changes this result at about the same level as the non-global logs ($\mathcal{O}(\alpha_s^2 \ln^2 b)$ in the exponent), and this feature needs to be accounted for still. Secondly the matching to LO DISENT combines channels and this spoils control over the $\alpha_s^2 \ln^2 Q/\Delta p_{t,jj}$ term in the expansion of the resummation to NLO. A full NLO matching with proper separation of the channels is awaited. The HERWIG curve also includes an intrinsic k_t component that lowers the height of the result at small $p_{t,jj}$, which can be easily included in

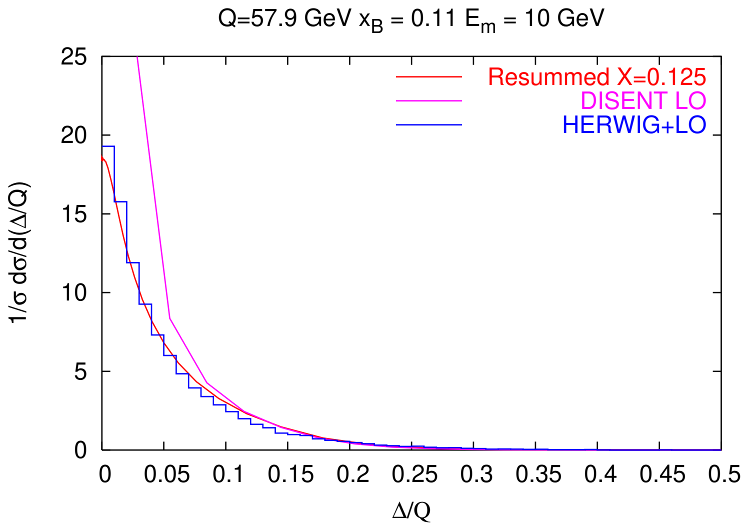


Fig. 6: Figure showing the resummed result matched to fixed-order DISENT results for the variable $\Delta = Q\Delta p_{t,jj}$. Also shown, for comparison, are HERWIG results with matrix-element corrections and the DISENT result alone.

the theoretical resummation but at present is excluded. Given these differences the very good agreement one sees with HERWIG is expected to change to some extent although broadly speaking the shapes of the two curves are expected to be similar. Similar conclusions apply for the ϕ_{jj} observable.

4 The vector Q_t of the current hemisphere

Next we examine a quantity that, as mentioned in the introduction, makes a very good analogy with Drell-Yan transverse momentum, Q_t, distributions. Comparison of the resummation of this observable with data could help to understand whether extra broadening of conventionally resummed Q_t spectra, is generated at small x. If so this will be a significant factor at the LHC. The observable in question is the (modulus of) the vectorially summed transverse momenta of all particles in the Breit frame current hemisphere:

$$Q_t = |\sum_{i \in \mathcal{H}_c} \vec{k}_{t,i}|. \tag{25}$$

Using momentum conservation this quantity is simply equal to the modulus of the transverse momenta of emissions in the remnant hemisphere. These emissions can all be ascribed to the incoming quark to NLL accuracy, *apart* from the soft wide-angle component where large-angle emissions in the current hemisphere can emit softer gluons into the remnant hemisphere (the by now familiar non-global logarithms).

The resummed result for this observable can be expressed as :

$$\frac{d\sigma}{dQ_T^2} \sim \sigma_0 \int_0^\infty bdbJ_0(bQ_t)\exp[-R(b)]S(b)q(x, 1/b^2) \tag{26}$$

where J_0 is the zeroth order Bessel function, $R(b)$ is the Sudakov exponent (the "radiator") , $S(b)$ the non-global contribution and q denotes the quark distribution summed over quark flavours with appropriate weights (charges).

The result for the radiator to NLL accuracy can be expressed, as before, in terms of a leading-log and next-to–leading log function:

$$R(b) = Lg_1(\alpha_s L) + g_2(\alpha_s L),\; L = \ln(bQ). \tag{27}$$

We have

$$g_1 = \frac{C_F}{2\pi\beta_0\lambda}[-\lambda - \ln(1-\lambda)], \tag{28}$$

$$g_2 = \frac{3C_F}{4\pi\beta_0}\ln(1-\lambda) + \frac{KC_F}{4\pi^2\beta_0^2}\left[\frac{\lambda}{1-\lambda} + \ln(1-\lambda)\right] + \frac{C_F}{2\pi}\left(\frac{\beta_1}{\beta_0^3}\right)\left[-\frac{1}{2}\ln^2(1-\lambda) - \frac{\lambda + (1-\lambda)}{1-\lambda}\right], \tag{29}$$

where we have $\lambda = \beta_0\alpha_s \ln[Q^2(\bar{b})^2]$, $\bar{b} = be^{\gamma_E}/2$ and $K = (67/18 - \pi^2/6)C_A - 5/9\, n_f$.

It is straightforward to express the result directly in Q_t space and one has for the pure NLL resummed terms:

$$\frac{d\sigma}{dQ_T^2} \sim \frac{d}{dQ_T^2}\left[e^{-R(Q/Q_t)-\gamma_E R'(Q/Q_t)}\frac{\Gamma(1-R'/2)}{\Gamma(1+R'/2)}q(x,Q_T^2)S(Q/Q_t)\right] \tag{30}$$

where $R' = dR/d\ln(Q/Q_t)$. The result has a divergence at $R' = 2$ which is due to retaining just NLL terms and is of the same nature as that discussed before for certain hadron-hadron event shapes and the Drell-Yan Q_t distribution. However in the present case the divergence is at quite low values of Q_t, e.g for $Q = 100$ GeV, the divergence is at around 0.5 GeV (depending on the exact choice for Λ_{QCD}). Thus it is possible to safely study the distribution down to Q_t values of a few GeV using the simple form Eq. 30. We note that is is also possible to eliminate the divergence if one defines the radiator such that $R(b) \to R(b)\theta(\bar{b}Q - 1)$, which is a restriction that follows from leading-order kinematics (that one assumes to hold at all orders). The resultant modification has only a negligible impact in the Q_t range that we expect to study phenomenologically.

After the matching to fixed-order is performed, we can probe the non-perturbative smearing e^{-gb^2} that one can apply to the b space resummed result. Comparisons with data should hopefully reveal whether the NLL resummed result + 'intrinsic k_t' smearing, mentioned above, is sufficient at smaller values of x or whether extra broadening is generated in the small x region, that has a significant effect on the result. Data from H1 are already available for this distribution [21] and this should enable rapid developments concerning the above issue.

5 Conclusions

In this article we have provided a summary of the developments discussed at the HERA-LHC workshop working group 2, concerning the topic of all-order QCD resummations. Specifically we have mentioned recent work carried out for hadronic dijet event shapes, dijet E_t and angular spectra and resummation of the current-hemisphere transverse momentum distribution in the DIS Breit frame.

We have stressed the important role of HERA studies in the development of the subject from the LEP era and the fact that, in this regard, HERA has acted as a bridge between LEP studies of the past (although LEP analysis of data continues and is an important source of information) and future studies at both the Tevatron and the LHC.

We have particularly tried to stress the continuing crucial role of HERA in testing all-order QCD dynamics, especially in the context of multi-hard parton observables where studies are currently ongoing. Careful experimental and theoretical collaborative effort is needed here in order to confirm the picture developed for NLL resummations and power corrections. If this program is successful it will greatly ease the way for accurate QCD studies at more complex hadronic environments, such as the LHC.

Acknowledgements

We would like to thank the convenors and organisers of the series of meetings which were a part of the HERA-LHC workshop, for their skillful organisation and for providing us all the necessary facilities needed to present and develop our respective contributions here.

References

[1] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. C **14**, 255 (2000).

[2] H1 Collaboration, G. J. McCance (2000). `hep-ex/0008009`.

[3] M. Dasgupta and G. Salam, Phys. Lett. B **512**, 323 (2001);
M. Dasgupta and G. Salam, JHEP **0203**, 017 (2002).

[4] Y. Dokshitzer, G. Marchesini, and B. Webber, Nucl. Phys. B **469**, 093 (1996).

[5] M. Dasgupta and B. Webber, Eur. Phys. J. C **1**, 539 (1998).

[6] M. Dasgupta and G. Salam, JHEP **0208**, 032 (2002).

[7] A. Banfi, Y. Dokshitzer, G. Marchesini, and G. Zanderighi, JHEP **0105**, 040 (2001);
A. Banfi, G. Marchesini, G. Smye, and G. Zanderighi, JHEP **0111**, 066 (2001).

[8] M. Dasgupta and G. P. Salam, J. Phys. **G30**, R143 (2004). `hep-ph/0312283`.

[9] A. Banfi, G. Salam, and G. Zanderighi, JHEP **0408**, 062 (2004).

[10] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C23**, 13 (2002). `hep-ex/0109029`;
ZEUS Collaboration, S. Chekanov *et al.*, Phys. Lett. **B551**, 226 (2003). `hep-ex/0210064`.

[11] A. Banfi and M. Dasgupta, JHEP **0401**, 027 (2004).

[12] N. Kidonakis, G. Oderda, and G. Sterman, Nucl. Phys. B. **525**, 299 (1998).

[13] S. Berge, P. Nadolsky, F. Olness, and C. P. Yuan, Phys. Rev. **D72**, 033015 (2005).
`hep-ph/0410375`.

[14] M. Dasgupta and G. Salam, Eur. Phys. J. C **24**, 323 (2001).

[15] D0 Collaboration, I. Bertram, Acta Phys. Polon. B **33**, 3141 (2002).

[16] A. Banfi, G. Salam, and G. Zanderighi, JHEP **0503**, 073 (2005).

[17] A. Banfi, G. Marchesini, G. Smye, and G. Zanderighi, JHEP **0108**, 047 (2001).

[18] S. Frixione and G. Ridolfi, Nucl. Phys. B **507**, 315 (1997);
M. Klasen and G. Kramer, Phys. Lett. B **366**, 385 (1996).

[19] S. Catani and M. Seymour, Nucl. Phys. B **485**, 291 (1997).

[20] G. Corcella *et al.*, JHEP01 **01**, 010 (2001).

[21] T. Kluge. Private Communication.

Matching Parton Showers and Matrix Elements

Stefan Höche[1], *Frank Krauss*[1], *Nils Lavesson*[2], *Leif Lönnblad*[2], *Michelangelo Mangano*[3], *Andreas Schälicke*[1], *Steffen Schumann*[1]
[1]Institut für Theoretische Physik, TU Dresden, Germany; [2]Department of Theoretical Physics, Lund University, Sweden; [3]CERN, Geneva, Switzerland.

Abstract
We compare different procedures for combining fixed-order tree-level matrix element generators with parton showers. We use the case of W-production at the Tevatron and the LHC to compare different implementations of the so-called CKKW scheme and one based on the so-called MLM scheme using different matrix element generators and different parton cascades. We find that although similar results are obtained in all cases, there are important differences.

1 Introduction

One of the most striking features of LHC final states will be the large number of events with several hard jets. Final states with 6 jets from $t\bar{t}$ decays will have a rate of almost 1Hz, with 10-100 times more coming from prompt QCD processes. The immense amount of available phase-space, and the large acceptance of the detectors, with calorimeters covering a region of almost 10 units of pseudorapidity (η), will lead to production and identification of final states with 10 or more jets. These events will hide or strongly modify all possible signals of new physics which involve the chain decay of heavy coloured particles, such as squarks, gluinos or the heavier partners of the top which appear in little-Higgs models. Being able to predict their features is therefore essential.

To achieve this, our calculations need to describe as accurately as possible both the full matrix elements for the underlying hard processes, as well as the subsequent development of the hard partons into jets of hadrons. For the complex final-state topologies we are interested in, no factorization theorem exists however to rigorously separate these two components, providing a constructive algorithm for the implementation of such separation. The main obstacle is the existence of several hard scales, like the jet transverse energies and dijet invariant masses, which for a generic multijet event will span a wide range. This makes it difficult to unambiguously separate the components of the event which belong to the "hard process" (to be calculated using a multiparton amplitude) from those developing during its evolution (described by the parton shower). A given $(N+1)$-jet event can be obtained in two ways: from the collinear/soft-radiation evolution of an appropriate $(N+1)$-parton final state, or from an N-parton configuration where hard, large-angle emission during its evolution leads to the extra jet. A factorization prescription (in this context this is often called a "matching scheme") defines, on an event-by-event basis, which of the two paths should be followed. The primary goal of a matching scheme is therefore to avoid double counting (by preventing some events to appear twice, once for each path), as well as dead regions (by ensuring that each configuration is generated by at least one of the allowed paths). Furthermore, a good matching scheme will optimize the choice of the path, using the one which guarantees the best possible approximation to a given kinematics. It is possible to consider therefore different matching schemes, all avoiding the double counting and dead regions, but leading to different results in view of the different ways the calculation is distributed between the matrix element and the shower evolution. As in any factorization scheme, the physics is independent of the separation between phases only if we have complete control over the perturbative expansion. Otherwise a residual scheme-dependence is left. Exploring different matching schemes is therefore crucial to assess the systematic uncertainties of multijet calculations.

In this work we present a first comparison of the three approaches which have been proposed so far, the so-called CKKW scheme, the Lönnblad scheme, and the MLM scheme. After shortly reviewing them, we present predictions for a set of W+multijet distributions at the Tevatron collider and at the LHC.

2 Matching procedures

In general, the different merging procedures all follow a similar strategy:

1. A jet measure is defined and all relevant cross sections including jets are calculated for the process under consideration. I.e. for the production of a final state X in pp-collisions, the cross sections for the processes $pp \to X + n$jets with $n = 0,\, 1,\, \dots n_{\max}$ are evaluated.
2. Hard parton samples are produced with a probability proportional to the respective total cross section, in a corresponding kinematic configuration following the matrix element.
3. The individual configurations are accepted or rejected with a dynamical, kinematics-dependent probability that includes both effects of running coupling constants and of Sudakov effects. In case the event is rejected, step 2 is repeated, i.e. a new parton sample is selected, possibly with a new number of jets.
4. The parton shower is invoked with suitable initial conditions for each of the legs. In some cases, like, e.g. in the MLM procedure described below, this step is performed together with the step before, i.e. the acceptance/rejection of the jet configuration. In all cases the parton shower is constrained not to produce any extra jet; stated in other words: Configurations that would fall into the realm of matrix elements with a higher jet multiplicity are vetoed in the parton shower step.

From the description above it is clear that the merging procedures discussed in this contribution differ mainly

- in the jet definition used in the matrix elements;
- in the way the acceptance/rejection of jet configurations stemming from the matrix element is performed;
- and in details concerning the starting conditions of and the jet vetoing inside the parton showering.

2.1 CKKW

In the original merging description according to [1, 2], which has been implemented [3] in SHERPA [4] in full generality, the acceptance/rejection of jet configurations from the matrix elements and the parton showering step are well-separated.

In this realisation of what is known as the CKKW-prescription the phase space separation for the different multijet processes is achieved through a $k_\perp$-measure [5, 6, 7]. For the case of hadron–hadron collisions, two final-state particles belong to two different jets, if their relative transverse momentum

$$k_\perp^{(ij)2} = 2 \min\left\{p_\perp^{(i)},\, p_\perp^{(j)}\right\}^2 \left[\cosh(\eta^{(i)} - \eta^{(j)}) - \cos(\phi^{(i)} - \phi^{(j)})\right] \tag{1}$$

is larger than a critical value, $k_{\perp,0}^2$. In addition, the transverse momentum of each jet has to be larger than $k_{\perp,0}$. The matrix elements are then reweighted by appropriate Sudakov and coupling weights. The task of the weight attached to a matrix element is to take into account terms that would appear in a corresponding parton shower evolution. Therefore, a "shower history" is reconstructed by clustering the initial and final state partons according to the $k_\perp$-algorithm. The resulting chain of nodal $k_\perp$-measures is interpreted as the sequence of relative transverse momenta of multiple jet production. The first ingredient of the weight are the strong coupling constants taken at the respective nodal values, divided by the value of α_S used during the matrix element evaluation. The other part of the correction weight is provided by NLL-Sudakov form factors defined by

$$\Delta_{q,g}(Q, Q_0) := \exp\left[-\int_{Q_0}^{Q} \mathrm{d}q\Gamma_{q,g}(Q, q)\right] , \tag{2}$$

where the integrated splitting functions $\Gamma_{q,g}$ are given by

$$\Gamma_{q,g}(Q, q) := \begin{cases} \frac{2C_F\alpha_s(q)}{\pi q}\left[\log\frac{Q}{q} - \frac{3}{4}\right] \\ \frac{2C_A\alpha_s(q)}{\pi q}\left[\log\frac{Q}{q} - \frac{11}{12}\right] \end{cases} \tag{3}$$

and contain the running coupling constant and the two leading, logarithmically enhanced terms in the limit when $Q_0 \ll Q$. The two finite, non-logarithmic terms $-3/4$ and $-11/12$, respectively emerge when integrating the non-singular part of the corresponding splitting function in the limits [0, 1]. Potentially, when q/Q is not going to zero, these finite terms are larger than the logarithmic terms and thus spoil an interpretation of the emerging NLL-Sudakov form factor as a non-branching probability. Therefore, without affecting the logarithmic order of the Sudakov form factors, these finite terms are integrated over the interval $[q/Q,\, 1 - q/Q]$ rather than over $[q,\, Q]$. This way a Sudakov form factor determines the probability for having no emission resolvable at scale Q_0 during the evolution from a higher scale Q to a lower scale Q_0. A ratio of two Sudakov form factors $\Delta(Q, Q_0)/\Delta(q, Q_0)$ then gives the probability for having no emission resolvable at scale Q_0 during the evolution from Q to q. Having reweighted the matrix element, a smooth transition between this and the parton shower region is achieved by choosing suitable starting conditions for the shower evolution of the parton ensemble and vetoing any parton shower emission that is harder than the separation cut $k_{\perp,0}$.

Within SHERPA the required matrix elements are provided by its internal matrix element generator AMEGIC++ [8] and the parton shower phase is handled by APACIC++ [9, 10]. Beyond the comparisons presented here the SHERPA predictions for W+multijets have already been validated and studied for Tevatron and LHC energies in [11, 12]. Results for the production of pairs of W-bosons have been presented in [13].

2.2 The Dipole Cascade and CKKW

The dipole model [14, 15] as implemented in the ARIADNE program [16] is based around iterating $2 \to 3$ partonic splitting instead of the usual $1 \to 2$ partonic splittings in a conventional parton shower. Gluon radiation is modeled as being radiated coherently from a color–anticolor charged parton pair. This has the advantage of eg. including first order correction to the matrix elements for $e^+e^- \to q\bar{q}$ in a natural way and it also automatically includes the coherence effects modeled by angular ordering in conventional showers. The process of quark antiquark production does not come in as naturally, but can be added [17]. The emissions in the dipole cascade is ordered according to invariant transverse momentum defined as

$$p_\perp^2 = \frac{s_{12}s_{23}}{s_{123}}, \tag{4}$$

where s_{ij} is the squared invariant mass of parton i and j, with the emitted parton having index 2.

When applied to hadronic collisions, the dipole model does not separate between initial and final state radiation. Instead all emissions are treated as coming from final state dipoles [18, 19]. To be able to extend the dipole model to hadron collisions, extended colored objects are introduced to model the hadron remnants. Dipoles involving hadron remnants are treated in a similar manner to the normal final-state dipoles. However, since the hadron remnant is considered to be an extended object, emissions with small wavelength are suppressed. This is modeled by only letting a fraction of the remnant take part in the emission. The fraction that is resolved during the emission is given by

$$a(p_\perp) = \left(\frac{\mu}{p_\perp}\right)^\alpha, \tag{5}$$

where μ is the inverse size of the remnant and α is the dimensionality.

There are two additional forms of emissions which need to be included in the case of hadronic collisions. One corresponds to an initial state $g \to q\bar{q}$ [20]. This does not come in naturally in the dipole model, but is added by hand in a way similar to that of a conventional initial-state parton shower [20]. The other corresponds to the initial-state $q \to gq$ (with the gluon entering into the hard sub-process) which could be added in a similar way, but this has not been implemented in ARIADNE yet.

When implementing CKKW for the dipole cascade, the procedure is slightly different from what has been described above [21, 22]. First, rather than just reconstructing emission scales using the $k_\perp$-algorithm, a complete dipole shower history is constructed for each state produced by the Matrix Element generator, basically answering the question *how would ARIADNE have generated this state*. This will produce a complete set of intermediate partonic states, S_i, and the corresponding emission scales, $p_{\perp i}$.

The Sudakov form factors are then introduced using the Sudakov veto algorithm. The idea is that we want to reproduce the Sudakov form factors used in Ariadne. This is done by performing a trial emission starting from each intermediate state S_i with $p_{\perp i}$ as a starting scale. If the emitted parton has a $p_\perp$ higher than $p_{\perp i+1}$ the state is rejected. This correspond to keeping the state according to the no emission probability in Ariadne, which is exactly the Sudakov form factor.

It should be noted that for initial-state showers, there are two alternative ways of defining the Sudakov form factor. The definition in eq. (2) is used in eg. HERWIG [23], while eg. PYTHIA [24, 25] uses a form which includes ratios of parton densities. Although formally equivalent to leading logarithmic accuracy, only the latter corresponds exactly to a no-emission probability, and this is the one generated by the Sudakov-veto algorithm. This, however, also means that the reconstructed emissions need not only be reweighted by the running α_S as in the standard CKKW procedure above, but also with ratios of parton densities, which in the case of gluon emissions correspond to the suppression due to the extended remnants in eq. (5) as explained in more detail in [22], where the complete algorithm is presented.

2.3 The MLM proceedure

In this approach we match the partons from the ME calculation to the jets reconstructed after the perturbative shower. Parton-level events are defined by a minimum E_T threshold E_T^{min} for the partons, and a minimum separation among them, $\Delta R_{jj} > R_{min}$. A tree structure is defined in analogy with the CKKW algorithm, starting however from the colour-flow extracted from the matrix-element calculation [26], thus defining the scales at which the various powers of α_s are calculated. However, no Sudakov reweighting is applied. Rather, events are showered, without any hard-emission veto during the shower. After evolution, a jet cone algorithm with cone size R_{min} and minimum transverse energy E_T^{min} is applied to the final state. Starting from the hardest parton, the jet which is closest to it in (η, ϕ) is selected. If the distance between the parton and the jet centroid is smaller than R_{min}, the parton and the jet match. The matched jet is removed from the list of jets, and matching for subsequent partons is performed. The event is fully matched if each parton has a matched jet. Events which do not match are rejected. A typical example is when two partons are so close that they cannot generate independent jets, and therefore cannot match. Rejection removes double counting of the leading double logarithms associated to the collinear behaviour of the amplitude when two partons get close. Another example is when a parton is too soft to generate its own jet, again failing matching. This removes double counting of some single logarithms. For events which satisfy matching, it is furthermore required that no extra jet, in addition to those matching the partons, be present. Events with extra jets are rejected, a suppression replacing the Sudakov reweighting used in the CKKW approach. Events obtained by applying this procedure to the parton level with increasing multiplicity can then be combined to obtain fully inclusive samples spanning a large multiplicity range. Events with extra jets are not rejected in the case of the sample with highest partonic multiplicity. The distributions of observables measured on this inclusive data set should not depend on the value of the parameters E_T^{min} and R_{min}, similar to the $k_{\perp,0}$ independence of the CKKW approach. This algorithm is encoded in the ALPGEN generator [27, 28], where evolution with both HERWIG and PYTHIA are enabled. In the following studies, the results quoted as "ALPGEN" employ the MLM matching scheme, and use ALPGEN for the generation of the parton-level matrix elements and HERWIG for the shower evolution and hadronisation.

3 Examples and comparisons

We present in this Section some concrete examples. We concentrate on the case of W+multijet production, which is one of the most studied final states because of its important role as a background to top quark studies at the Tevatron. At the LHC, W+jets, as well as the similar Z+jets processes, will provide the main irreducible backgrounds to signals such as multijet plus missing transverse energy, typical of Supersymmetry and of other manifestations of new physics. The understanding of W+multijet production at the Tevatron is therefore an essential step towards the validation and tuning of the tools presented here, prior to their utilization at the LHC.

For each of the three codes we calculated a large set of observables, addressing inclusive properties of the events (p_T spectrum of the W and of leading jets), geometric correlations between the jets, and intrinsic properties of the jets themselves, such as energy shapes. In view of the limited space available here we present only a subset of our studies, which will be documented in more detail in a future publication. An independent study of the systematics in the implementation of the CKKW prescription in HERWIG and PYTHIA was documented in [29].

The comparison between the respective results shows a reasonable agreement among the three approaches, but points also to differences, in absolute rates as well as in the shape of individual distributions, which underscore the existence of an underlying systematic uncertainty. The differences are nevertheless by and large consistent with the intrinsic systematic uncertainties of each of the codes, such as the dependence on the generation cuts or on the choice of renormalization scale. There are also differences due to the choice of parton cascade. In particular the ARIADNE cascade is quite different from a conventional parton shower, and it has been shown in this workshop [30] that ARIADNE eg. gives a much harder $p_{\perp\mathrm{W}}$ spectrum than does HERWIG or PYTHIA. Now, although the hard emissions in the matching proceedures should be described by the exact matrix element, the Sudakov formfactors in the ARIADNE matching (and indirectly in the MLM scheme) are generated by the cascade. In addition, the events in the ARIADNE matching are reweighted by PDF ratios in the same way as is done in the plain cascade. This means that some properties of the cascade may affect also the hard emissions in the matching procedure in these cases.

The existence in each of the codes of parameters specifying the details of the matching algorithms presents therefore an opportunity to tune each code so as to best describe the data. This tuning should be seen as a prerequisite for a quantitative study of the overall theoretical systematics: after the tuning is performed on a given set of final states (e.g. the W+jets considered here), the systematics for other observables or for the extrapolation to the LHC can be obtained by comparing the difference in extrapolation between the various codes. It is therefore auspicable that future analysis of Tevatron data will provide us with spectra corrected for detector effects in a fashion suitable to a direct comparison against theoretical predictions.

The following two sections present results for the Tevatron ($p\bar{p}$ collisions at 1.96 TeV) and for the LHC (pp at 14 TeV), considering events with a positively charged W. Jets are defined by Paige's `GETJET` cone-clustering algorithm, with a calorimeter segmentation of ($\Delta\eta$, $\Delta\phi$) = (0.1,6°) and a cone size of 0.7 and 0.4 for Tevatron and LHC, respectively. At the Tevatron (LHC) we consider jets with $E_T > 10(20)$ GeV, within $|\eta| < 2(4.5)$. We use the PDF set CTEQ6L, with $\alpha_S(M_Z) = 0.118$.

For our default distributions, the ALPGEN results for the Tevatron (LHC) were obtained using parton level cuts of $p_{T,min} = 10(20)$ GeV, $|\eta| < 2.5(5)$, $Rjj < 0.7(0.4)$ and matching defined by $E_{Tmin} = 10$ GeV and $R = 0.7$. The SHERPA samples have been generated using matrix elements with up to four extra jets and the value of the merging scale has been chosen to $k_{\perp,0} = 10(20)$ GeV,

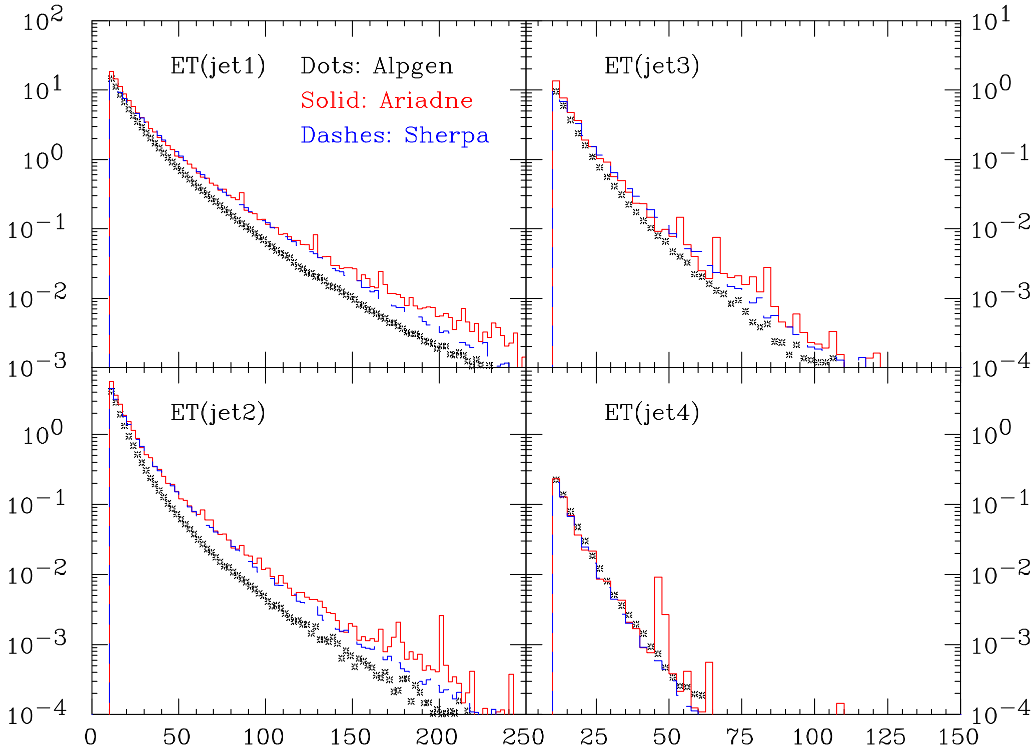


Fig. 1: Inclusive E_T spectra of the leading 4 jets at the Tevatron (pb/GeV).

respectively. Finally, for ARIADNE, the parton level cuts were $p_{T,min} = 10(20)$, $Rjj < 0.5(0.35)$ and, in addition, a cut on the maximum pseudorapidity of jets, $\eta_{j\,\max} = 2.5(5.0)$.

In all cases, the analysis is done at the hadron level, but without including the underlying event.

3.1 Tevatron Studies

We start by showing in fig. 1 the inclusive E_T spectra of the leading 4 jets. The absolute rate predicted by each code is used, in units of pb/GeV. We notice that the ALPGEN spectrum for the first two jets is softer than both SHERPA and ARIADNE, with the latter having even harder tails. The spectra for the third and fourth jet are instead in very good agreement, both in shape and normalization. As an indication of possible sources of systematics in these calculations, we rescaled the renormalization scale used in ALPGEN by a factor of 1/2. As seen in fig. 2 the distributions for the leading jets is now in perfect agreement with SHERPA, with an increase in rate for the third and fourth jet. These plots give us an idea of the level of flexibility which is intrinsic in the calculation of higher-order jet production. One should not forget that the rate for production of N jets is proportional to the Nth power of α_s, and the absence of the full set of virtual corrections unavoidably leads to a large scale uncertainty.

Figure 3 shows the inclusive η spectra of the leading 4 jets, all normalized to unit area. The asymmetry for the first two jets is due to the $W+$, which preferentially moves in the direction of the proton (positive η). This is partially washed out in the case of the third and fourth jet. There is a good agreement between the spectra of ALPGEN and SHERPA, while ARIADNE spectra appear to be broader, in particular for the subleading jets. This broadening is expected since the gluon emissions in ARIADNE are essentially unordered in rapidity, which means that the Sudakov form factors applied to the ME-generated states include also a $\log 1/x$ resummation absent in the other programs.

The top-left plot of fig. 4 shows the inclusive p_T distribution of the W^+ boson, with absolute normalization in pb/GeV. This distribution reflects in part the behaviour observed for the spectrum of the leading jet, with ALPGEN slightly softer, and ARIADNE slightly harder than SHERPA. The $|\eta|$ separation

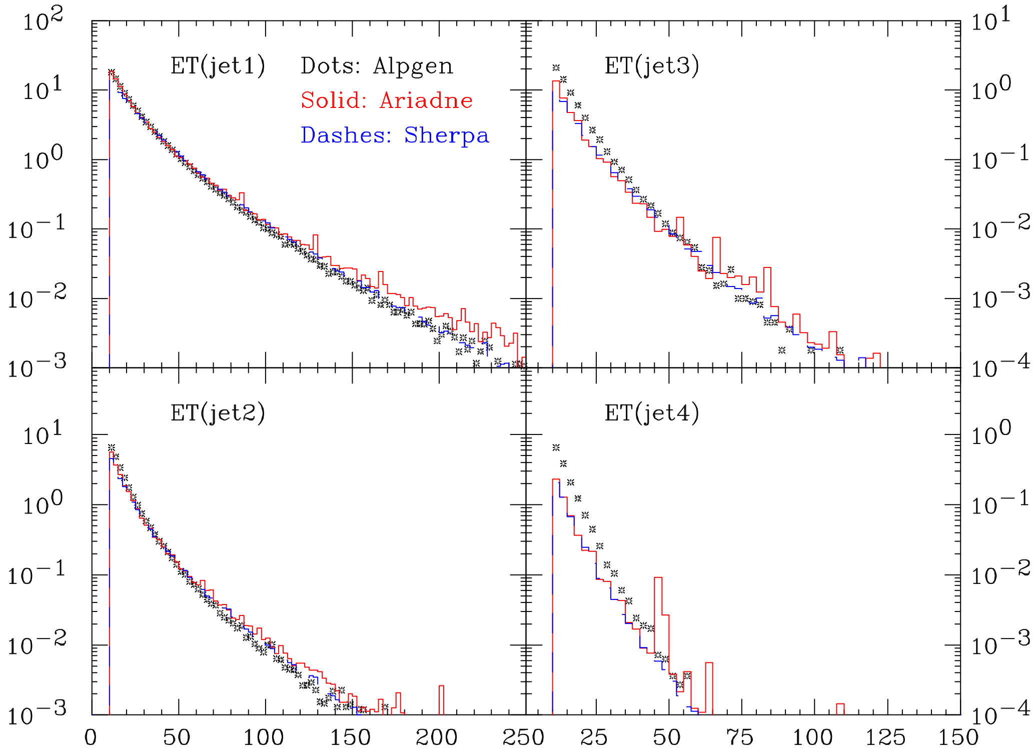


Fig. 2: Same as Fig. 1, but with the ALPGEN renormalization scale reduced by a factor 2.

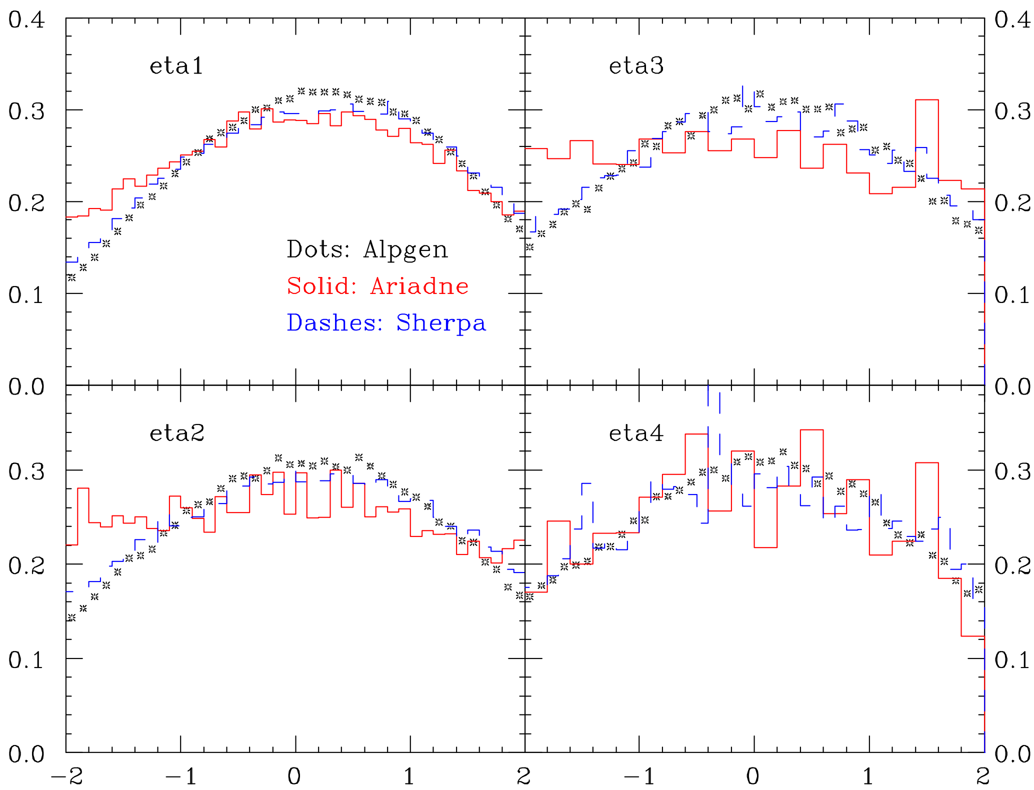


Fig. 3: Inclusive η spectra of the leading 4 jets at the Tevatron, normalized to unit area.

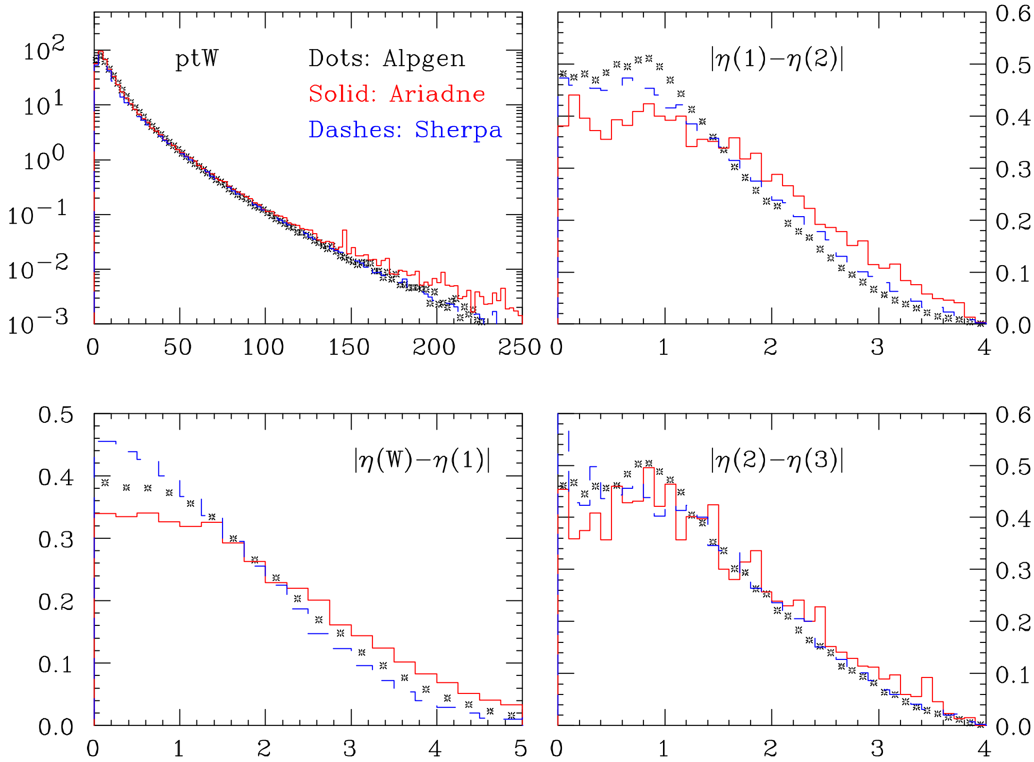


Fig. 4: Top left: inclusive $p_\perp(W^+)$ spectrum, pb/GeV. Bottom left: $|\eta(W^+) - \eta(\mathrm{jet}_1)|$ (unit area). Top right: $|\eta(\mathrm{jet}_1) - \eta(\mathrm{jet}_2)|$ and bottom right: $|\eta(\mathrm{jet}_2) - \eta(\mathrm{jet}_3)|$ (unit area).

between the W and the leading jet of the event is shown in the top-right plot. The two lower plots show instead the distributions of $|\eta(\mathrm{jet}_1) - \eta(\mathrm{jet}_2)|$ and $|\eta(\mathrm{jet}_2) - \eta(\mathrm{jet}_3)|$. These last three plots are normalized to unit area. In all these cases, we observe once more a reflection of the behaviour observed in the inclusive η distributions of the jets: ALPGEN is slightly narrower than SHERPA, and ARIADNE is slightly broader.

3.2 LHC Predictions

In this section we confine ourselves to ALPGEN and SHERPA. It turns out that ARIADNE has a problem in the reweighting related to the fact that initial-state $g \to q\bar{q}$ emissions, contrary to the gluon emissions, are ordered both in $p_\perp$ and rapidity. With the extra phase space available at the LHC this leads to unnatural reconstructions which, in turn, gives rise to a systematically too high reweighting. A solution for this problem is under investigation and a fuller comparison including ARIADNE will be documented in a future publication.

Following the same sequence of the Tevatron study, we start by showing in fig. 5 the inclusive E_T spectra of the leading 4 jets. The absolute rate predicted by each code is used, in units of pb/GeV. The relative behaviour of the predictions by ALPGEN and SHERPA follows the pattern observed in the Tevatron case, with ALPGEN being softer in the case of the leading two jets. We do not notice however a deterioration of the discrepancy going from the Tevatron to the LHC, suggesting that once a proper tuning is achieved at lower energy the predictions of two codes for the LHC should be comparable.

Figure 6 shows the inclusive η spectra of the leading 4 jets, all normalized to unit area. The asymmetry now is not present, because of the symmetric rapidity distribution of the W^+ in pp collisions. As in the case of the Tevatorn, jet production in ALPGEN is slightly more central than in SHERPA.

The top-left plot of fig. 7 shows the inclusive p_T distribution of the W^+ boson, with absolute

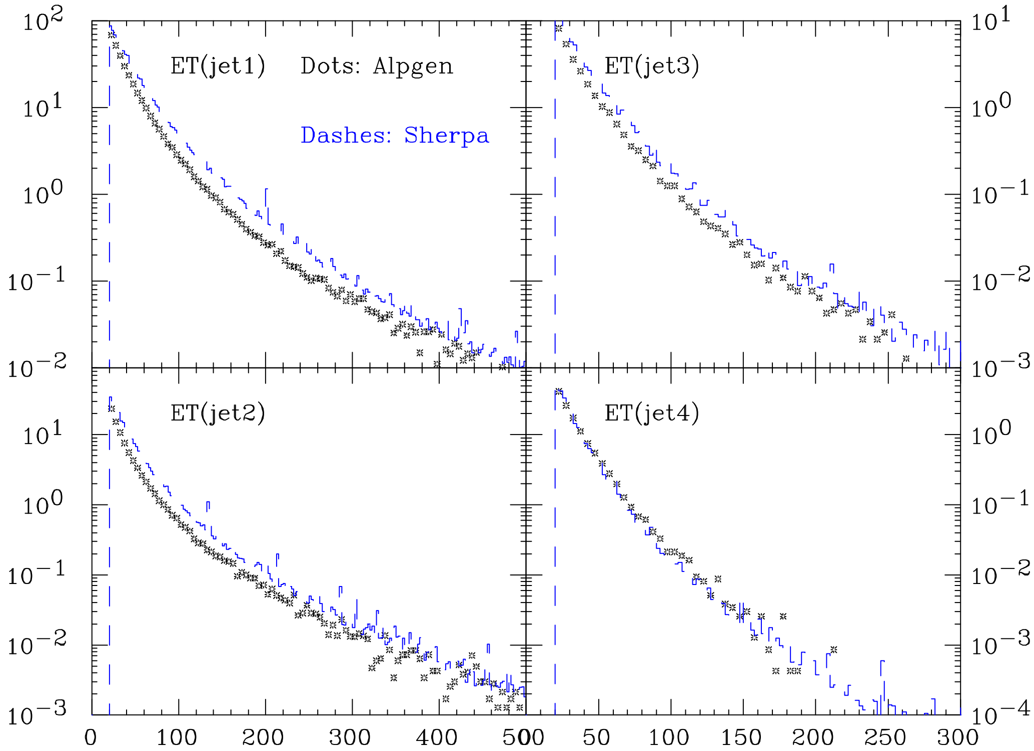


Fig. 5: Inclusive E_T spectra of the leading 4 jets at the LHC (pb/GeV).

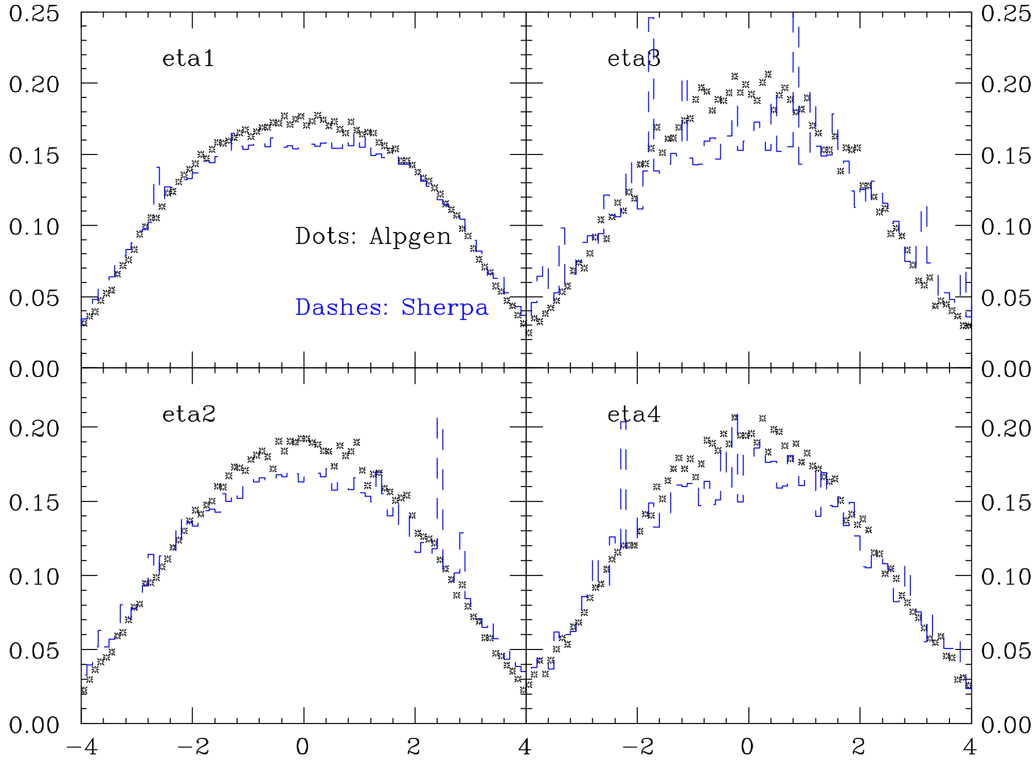


Fig. 6: Inclusive η spectra of the leading 4 jets at the LHC, normalized to unit area.

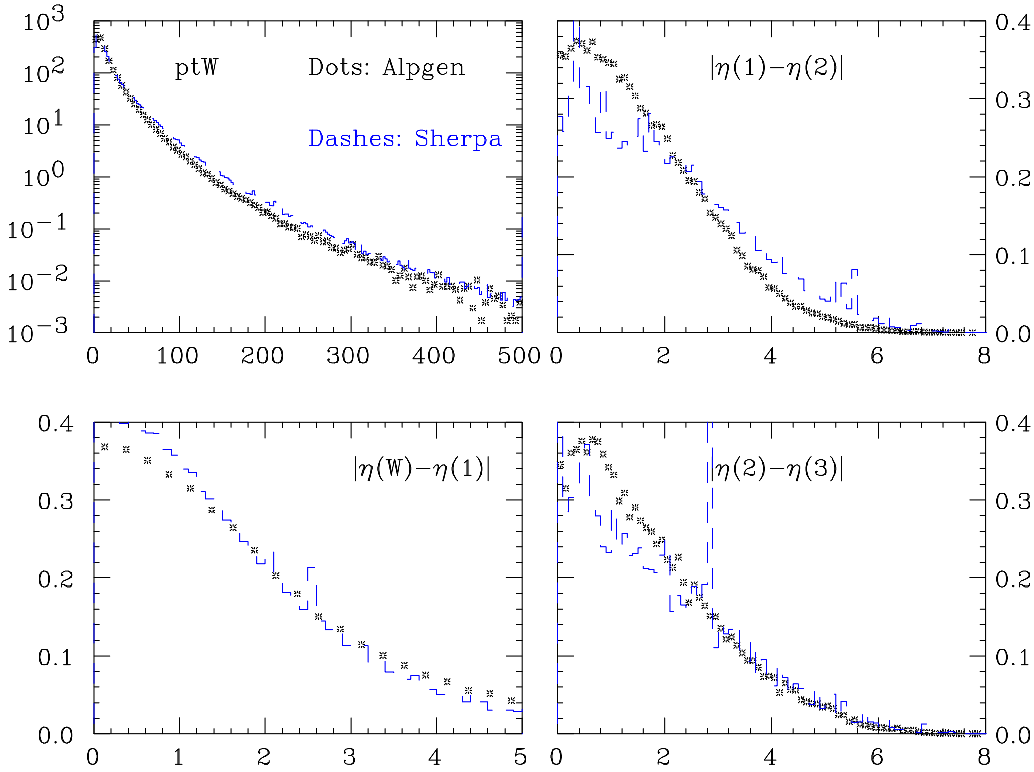


Fig. 7: Top left: inclusive $pt_T(W^+)$ spectrum, pb/GeV. Bottom left: $|\eta(W^+) - \eta(\text{jet}_1)|$ (unit area). Top right: $|\eta(\text{jet}_1) - \eta(\text{jet}_2)|$ and bottom right: $|\eta(\text{jet}_2) - \eta(\text{jet}_3)|$ (unit area).

normalization in pb/GeV. The $|\eta|$ separation between the W and the leading jet of the event is shown in the top-right plot. The two lower plots show instead the distributions of $|\eta(\text{jet}_1) - \eta(\text{jet}_2)|$ and $|\eta(\text{jet}_2) - \eta(\text{jet}_3)|$. These last three plots are normalized to unit area. As before, the features of these comparisons reflect what observed in the inclusive jet properties.

4 Conclusions

This document summarizes our study of a preliminary comparison of three independent approaches to the problems of merging matrix element and parton shower evolution for multijet final states. Overall, the picture shows a general consistency between the three approaches, although there are occasional differences. The origin of these differences is under study. It could be based on intrinsic differences between the matching schemes, as well as to differences between the different shower algorithms used in the three cases. We expect nevertheless that these differences be reconciled with appropriate changes in the default parameter settings for the matching schemes, as partly supported by the few systematic studies presented here. Validation and tuning on current Tevatron data is essential, and will allow to reduce the systematics.

It is also important to compare these models to HERA data. However, besides some preliminary investigations for ARIADNE [31], there is no program which properly implement a CKKW or MLM matching scheme for DIS. The energy of HERA is, of course, lower, as are the jet multiplicities and jet energies, but HERA has the advantage of providing a large phase space for jet production which is not mainly determined by the hard scale, Q^2, but rather by the total energy, giving rise to large logarithms of $x \approx Q^2/W^2$ which need to be resummed to all orders. This is in contrast to the Tevatron, where the phase space for additional jets in W-production mainly are determined by m_W. However, when going to the LHC there may also be important effects of the increased energy, and there will be large logarithms of $x \propto m_W/\sqrt{S}$ present, which may need to be resummed. The peculiar treatment of the available phase space in the plain ARIADNE cascade means that some logarithms of x are resummed in contrast to conventional initial-state parton cascades. This feature survives the matching procedure and is the reason for the broader rapidity spectra presented in the figures above. In DIS this is reflected by the increased rate of forward jets, and such measurements are known to be well reproduced by ARIADNE while conventional parton showers fail. It would be very interesting if the matching of these conventional showers with higher order matrix elements would improve the description of forward jets. In that case the extrapolation of the Tevatron results to the LHC would be on much safer grounds.

As our study of the LHC distributions suggests, the increase in energy exhibits the same pattern of discrepancies observed at the Tevatron. We therefore expect that if different algorithms are tuned on the same set of data, say Tevatron W+jets, they will extrapolate in the same way to the LHC or to different final states, for example multijet configurations without W bosons. While these systematics studies can be performed directly at the Monte Carlo level, only the availability of real measurements from the Tevatron can inject the necessary level or realism in these exploration. We look forward to the availability of such data.

References

[1] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, JHEP **0111**, 063 (2001). `arXiv:hep-ph/0109231`.

[2] F. Krauss, JHEP **0208**, 015 (2002). `arXiv:hep-ph/0205283`.

[3] A. Schälicke and F. Krauss, JHEP **07**, 018 (2005). `hep-ph/0503281`.

[4] T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann, and J. Winter, JHEP **0402**, 056 (2004). `arXiv:hep-ph/0311263`.

[5] S. Catani, Y. L. Dokshitser, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. **B 269**, 432 (1991).

[6] S. Catani, Y. L. Dokshitser, and B. R. Webber, Phys. Lett. **B 285**, 291 (1992).

[7] S. Catani, Y. L. Dokshitser, and B. R. Webber, Nucl. Phys. **B 406**, 187 (1993).

[8] F. Krauss, R. Kuhn, and G. Soff, JHEP **0111**, 044 (2002). `arXiv:hep-ph/0109036`.

[9] R. Kuhn, F. Krauss, B. Ivanyi, and G. Soff, Comput. Phys. Commun. **134**, 223 (2001). `hep-ph/0004270`.

[10] F. Krauss, A. Schälicke, and G. Soff, hep-ph/0503087.

[11] F. Krauss, A. Schälicke, S. Schumann, and G. Soff, Phys. Rev. **D70**, 114009 (2004). `hep-ph/0409106`.

[12] F. Krauss, A. Schalicke, S. Schumann, and G. Soff, Phys. Rev. **D72**, 054017 (2005). `hep-ph/0503280`.

[13] T. Gleisberg, F. Krauss, A. Schalicke, S. Schumann, and J.-C. Winter, Phys. Rev. **D72**, 034028 (2005). `hep-ph/0504032`.

[14] G. Gustafson and U. Pettersson, Nucl. Phys. **B306**, 746 (1988).

[15] G. Gustafson, Phys. Lett. **B175**, 453 (1986).

[16] L. Lönnblad, Comput. Phys. Commun. **71**, 15 (1992).

[17] B. Andersson, G. Gustafson, and L. Lönnblad, Nucl. Phys. **B339**, 393 (1990).

[18] B. Andersson, G. Gustafson, L. Lönnblad, and U. Pettersson, Z. Phys. **C43**, 625 (1989).

[19] L. Lönnblad, Nucl. Phys. **B458**, 215 (1996). `hep-ph/9508261`.

[20] L. Lönnblad, Z. Phys. **C65**, 285 (1995).

[21] L. Lönnblad, JHEP **05**, 046 (2002). `hep-ph/0112284`.

[22] N. Lavesson and L. Lönnblad, JHEP **07**, 054 (2005). `hep-ph/0503293`.

[23] G. Corcella *et al.*, JHEP **01**, 010 (2001). `hep-ph/0011363`.

[24] T. Sjöstrand, and others, Comput. Phys. Commun. **135**, 238 (2001). `arXiv:hep-ph/0010017`.

[25] T. Sjöstrand, L. Lönnblad, S. Mrenna, and P. Skands (2003). `hep-ph/0308153`.

[26] F. Caravaglios, M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. **B539**, 215 (1999). `hep-ph/9807570`.

[27] M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. **B632**, 343 (2002). `hep-ph/0108069`.

[28] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07**, 001 (2003). `hep-ph/0206293`.

[29] S. Mrenna and P. Richardson, JHEP **05**, 040 (2004). `hep-ph/0312274`.

[30] B. Kersevan and E. Richter-Was., *The Monte Carlo Event Generator AcerMC and package AcderDET*. These Proceedings.

[31] C. Aberg, *Correcting the colour dipole cascade with fixed order matrix elements in deep inelastic scattering*. Diploma thesis, Lund preprint LU-TP 04-25.

Constrained non-Markovian Monte Carlo modeling of the evolution equation in QCD*

S. Jadach and M. Skrzypek
Institute of Nuclear Physics, Academy of Sciences, ul. Radzikowskiego 152, 31-342 Cracow, Poland and CERN Department of Physics, Theory Division, CH-1211 Geneva 23, Switzerland

Abstract
A new class of the constrained Monte Carlo (CMC) algorithms for the QCD evolution equation was recently discovered. The constraint is imposed on the type and the total longitudinal energy of the parton exiting QCD evolution and entering a hard process. The efficiency of the new CMCs is found to be reasonable.

This brief report summarizes the recent developments in the area of the Monte Carlo (MC) techniques for the perturbative QCD calculations. Most of it was done at the time of the present HERA–LHC workshop, partial results being presented at several of its meetings. At present, two papers, [1] and [2], demonstrating the principal results are already available. Generally, these MC techniques concern the QCD evolution of the parton distribution functions (PDFs) $D_k(x, Q)$, where k denotes the type of the parton (quark, gluon), x the fraction of longitudinal momentum of the initial hadron carried by the parton, and the size of the available real/virtual emission phase space is Q. The evolution equation describes the response of the PDF to an increase of Q; $D_k(x, Q)$ is an inclusive distribution and can be measured almost directly in hadron–lepton scattering. On the other hand, it was always known that there exists in QCD an *exclusive* picture of the PDF, the so-called parton-shower process, in which $D_k(x, Q)$ is the distribution of the parton exiting the emission chain and entering the hard process (lepton–quark for example). The kernel functions $P_{kj}(Q, z)$, that govern the differential evolution equations of PDFs are closely related to distributions governing a single emission process $(i-1) \to i$ in the parton shower: $P_{k_i k_{i-1}}(Q_i, x_i/x_{i-1})$.

In other words, the evolution (Q-dependence) of PDFs and the parton shower represent two faces of the same QCD reality. The first one (inclusive) is well suited for basic precision tests of QCD at hadron–lepton colliders, while the second one (exclusive) provides realistic exclusive Monte Carlo modeling, vitally needed for experiments at high-energy particle colliders.

At this point, it is worth stressing that, so far, we were referring to DGLAP-type PDFs [3] and their evolution, and to constructing a parton-shower MC starting from them, as was done two decades ago and is still done today. This involves a certain amount of "backward engineering" and educated guesses, because the classical inclusive PDFs integrate over the p_T of the exiting parton. The so-called unintegrated PDFs (UPDFs) $D_k(x, p_T, Q)$ would be more suitable for the purpose, leading to higher-quality QCD calculations. UPDFs are, however, more complicated to handle, both numerically and theoretically. (It is still a challenge to construct a parton-shower MC based consistently on the theoretically well defined UPDFs.)

Another interesting "entanglement" of the evolution of PDFs on one side and of the parton shower (PS) MC on the other side is also present in the modeling of the showering of the incoming hadron – mostly for technical reasons and convenience. The Markovian nature of the QCD evolution can be exploited directly in the PS MC, where partons split/decay as long as there is enough energy to dissipate (final state) or the upper boundary Q of the phase space is hit (initial state). The multiparton distribution

*Supported in part by the EU grant MTKD-CT-2004-510126, in partnership with the CERN Physics Department.

in such a MC is a product of the evolution kernels. However, such a direct Markovian MC simulation of a shower is hopelessly inefficient in the initial state, because the hard process accepts only certain types and momenta of the incoming partons – most of the *shower histories* are rejected (zero MC weight) by the hard process. This is especially true for forming narrow resonances such as electroweak bosons or Higgs boson at the LHC. The well known "workaround" is the backward evolution MC algorithm of Sjöstrand, used currently in all PS MCs, such as HERWIG [4] and PYTHIA [5]. Contrary to the forward Markovian MC, where the physics inputs are PDFs at low $Q_0 \sim 1$ GeV and the evolution kernels, in the backward evolution MC one has to know PDFs in the entire range (Q_0, Q) from a separate non-MC numerical program solving the evolution equation and setting up look-up tables (or numerical parametrization) for them[1].

The following question has been pending in the parton-shower MC methodology for a long time: Could one invent an efficient "monolithic" MC algorithm for the parton shower from the incoming hadron, in which no external PDFs are needed and the only input are PDFs at Q_0 and the evolution kernel (the QCD evolution being a built-in feature of the parton shower MC)? Another question rises immediately: Why bother? Especially since this is a tough technical problem. This cannot still be fully answered before the above technique is applied in the full-scale (four-momentum level) PS MC. Generally, we hope that this technique will open new avenues in the development of the PS MC at the next-to-leading-logarithmic (NLL) level. In particular, it may help in constructing PS MCs closely related to unintegrated structure functions and, secondly, it may provide a better integration of the NLL parton shower (yet to be implemented!) with the NLL calculation for the hard process.

The first solution of the above problem of finding an efficient "constrained MC" (CMC) algorithm for the QCD evolution was presented in refs. [1, 6]. This solution belongs to what we call a CMC class II, and it relies on the observation that all initial PDFs at Q_0 can be approximated by const $\cdot\, x_0^{\eta-1}$; this is to be corrected by the MC weight at a later stage. This allows elimination of the constraint $x = \prod_i z_i$, at the expense of x_0, keeping the factorized form of the products of the kernels. Simplifying phase-space boundaries in the space of z_i is the next ingredient of the algorithm. Finally, in order to reach a reasonable MC efficiency for the pure bremsstrahlung case out of the gluon emission line, one has to generate a $1/z$ singularity in the $G \to G$ kernel in a separate branch of the MC. The overall efficiency of the MC is satisfactory, as is demonstrated in ref. [1] for the case of the pure bremsstrahlung out of the gluon and quark colour charge. Generalization to the quark–gluon transition is outlined, but not yet implemented. The main drawback of this method is its algebraic complexity. Further improvement of its relatively low MC efficiency is possible (even though it could lead to even more algebraic complexity).

The second, more efficient, CMC algorithm was presented in ref. [2] (as well as during the October 2004 meeting of the workshop). It belongs to what we call a CMC class I. The main idea is to project/map points from the hyperspace defined by the energy constraint $x = \prod_i z_i$, into a simpler hyperspace, defined by the hardest emission, $x = \min z_i$. This mapping is accompanied by the appropriate MC weight, which compensates exactly for the deformation of the distributions involved, and the bookkeeping of the hyperspace boundaries is rigorous. The above describes a CMC for the pure bremsstrahlung segment of the gluon emission out of a quark or gluon chain. Many such segments are interconnected by the quark–gluon transitions. The algebraic hierarchic reorganization of the emission chain into a super-level of the quark–gluon transitions and sub-level of the pure bremsstrahlung is an important ingredient in all CMC algorithms and will be published separately [7]. The basic observation made in ref. [8] is that the average number of super-level transitions is low, ~ 1; hence for precision of a 10^{-4} it is sufficient to limit it to three or four transitions. The integration/simulation of the super-level variables is done efficiently

[1]Backward evolution is basically a change in the order of the generation of the variables: Consider generating $\rho(x, y)$, where one generates first x according to $\rho(x) = \int dy\ \rho(x, y)$, and next y according to $\rho(x, y)$, by means of *analytical* mappings of x and y into uniform random numbers. However, such analytical mappings may not exist, if we insist on generating first x and next y! Nevertheless, we may still proceed with the same method by "brute force", if we pretabulate and invert numerically the functions $R(x) = \int^x \int dx' dy'\ \rho(x', y')$ and $R_x(y) = \int^y dy'\ \rho(x, y')$. This is what is done in a more dimensional case of the backward-evolution MC; it also explains why pretabulated PDFs are needed in these methods.

using the general-purpose MC tool FOAM [9, 10]. The above proof of the correctness of the CMC class I algorithm concept was given in ref. [2] for the full DGLAP-type QCD evolution with the LL kernels (including quark–gluon transitions).

Although our main aim is to construct the non-Markovian CMC class of algorithms, we have developed in parallel the family of Markovian MC (MMC) algorithms/programs, which provide numerical solutions of the QCD evolution equations with high precision, $\sim 10^{-3}$. We use them at each step of the CMC development as numerical benchmarks for the precision tests of the algorithms and their software implementations. The first example of MMC for DGLAP at LL was defined/examined in ref. [8] and tested using the non-MC program QCDnum16 [11][2]. In some cases our MMC programs stand ahead of their CMC brothers; for instance, they already include NLL DGLAP kernels. A systematic description of the MMC family of our MC toolbox is still under preparation [13].

The last development at the time of the workshop was an extension of the CMC type-I algorithm from DGLAP to CCFM one-loop evolution [14] (also referred to as HERWIG evolution [15]), in which the strong coupling constant gains z-dependence, $\alpha_s(Q) \to \alpha_s(Q(1-z))$, as advocated in ref. [16], confirmed by NLL calculations [17]. The above ansatz also compels introduction of a Q-dependent IR cutoff, $\varepsilon = Q_\varepsilon/Q$: another departure from DGLAP. This version of the CMC is still unpublished. Its version for the pure bremsstrahlung was presented at the March 2005 meeting of the workshop; in particular a perfect numerical agreement with the couterpartner MMC was demonstrated. Recently both CMC and MMC for the one-loop CCFM were extended to quark–gluon transitions, and again perfect agreement was found.

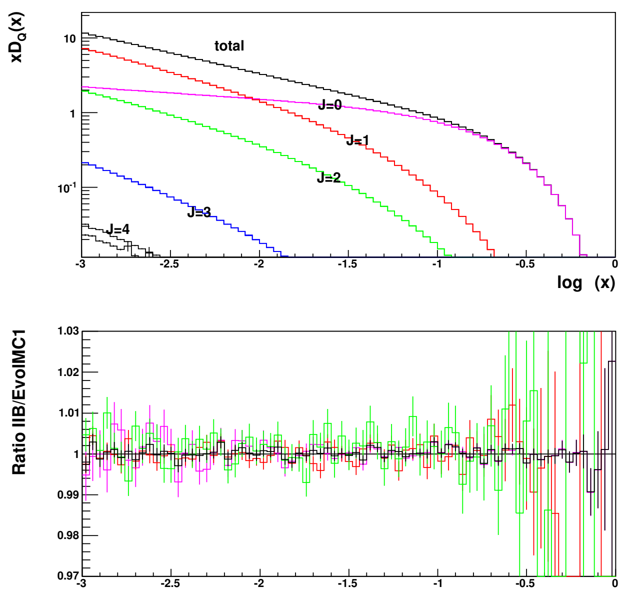


Fig. 1: CMC of the one-loop CCFM versus the corresponding MMC for quarks; number of quark–gluon transitions $J = 0, 1, 2, 3, 4$, and the total. The ratios in the lower plot are for $n = 0, 1$ and the total (blue).

For the detailed description of the new CMC algorithm, we refer the reader to the corresponding papers [1] and [2] and workshop presentations[3]. Here, let us only show one essential step in the development of the CMC for the one-loop CCFM model – the mapping of the Sudakov variables for the pure

[2]It was also compared with the non-MC program APCheb [12].

[3]To be found at `http://jadach.home.cern.ch/jadach/`.

bremsstrahlung:

$$
\begin{aligned}
I &= \int_{t_0}^{t_1} dt \int_0^{z_1} dz\; \alpha(Q(1-z))\; zP^{\Theta}_{GG}(z,t) \\
&= \frac{2}{\beta_0}\int_0^{z_1} dz \int_{t_0}^{t_1} dt\; \frac{1}{\hat{t}+\ln(1-z)} \frac{\theta_{\ln(1-z)>\hat{t}_\varepsilon-\hat{t}}}{1-z} = \frac{2}{\beta_0}\int_0^{y_{\max}} dy(z) \int_0^1 ds(t).
\end{aligned}
\tag{1}
$$

The short-hand notation $\hat{t}=\hat{t}(t)\equiv t-t_\Lambda$ and $v=\ln(1-z)$ supplements that of ref. [2] in use, and the mapping reads

$$
\begin{aligned}
&y(z) = \rho(v_1;\hat{t}_1,\hat{t}_0) = \rho(v_1+\hat{t}_1) - \theta_{v_1>t_\varepsilon-t_0}\rho(v_1+\hat{t}_0), \quad s(t) = \frac{\ln(\hat{t}+v)}{\rho'(v;\hat{t}_1,\hat{t}_0)}, \\
&\rho'(v;\hat{t}_1,\hat{t}_0) = \theta_{v<t_\varepsilon-t_0}\rho'(v+\hat{t}_1) + \theta_{v>t_\varepsilon-t_0}[\rho'(v+\hat{t}_1) - \rho'(v+\hat{t}_0)],
\end{aligned}
\tag{2}
$$

where $\rho(t) \equiv \hat{t}(\ln\hat{t} - \ln\hat{t}_\varepsilon) + \hat{t}_\varepsilon - \hat{t}$. Once the above mapping is set, the same algorithm, with the parallel shift $y_i \to y_i + Y$, can be used in this case. The super-level of quark–gluon transitions is again implemented using FOAM[4]. A numerical comparison of the corresponding CMC and MMC programs is shown in fig. 1. The MC efficiency is comparable with that of the DGLAP case.

Summary: We have constructed and tested new, efficient, constrained MC algorithms for the initial-state parton-emission process in QCD.

References

[1] S. Jadach and M. Skrzypek, report IFJPAN-V-04-06, hep-ph/0504205, Acta Phys. Pol. in press.

[2] S. Jadach and M. Skrzypek, report IFJPAN-V-04-07, `hep-ph/0504263`.

[3] L.N. Lipatov, *Sov. J. Nucl. Phys.* **20** (1975) 95;
V.N. Gribov and L.N. Lipatov, *Sov. J. Nucl. Phys.* **15** (1972) 438;
G. Altarelli and G. Parisi, *Nucl. Phys.* **126** (1977) 298;
Yu. L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 64.

[4] G. Corcella *et al.*, *JHEP* **01** (2001) 010, `hep-ph/0011363`.

[5] T. Sjöstrand *et al.*, *Comput. Phys. Commun.* **135** (2001) 238–259, `hep-ph/0010017`.

[6] S. Jadach and M. Skrzypek, *Nucl. Phys. Proc. Suppl.* **135** (2004) 338–341.

[7] S. Jadach, M. Skrzypek, and Z.Was, Report IFJPAN-V-04-09.

[8] S. Jadach and M. Skrzypek, *Acta Phys. Polon.* **B35** (2004) 745–756, `hep-ph/0312355`.

[9] S. Jadach, *Comput. Phys. Commun.* **152** (2003) 55–100, `physics/0203033`.

[10] S. Jadach and P. Sawicki, `physics/0506084`.

[11] M. Botje, ZEUS Note 97-066, http://www.nikhef.nl/ h24/qcdcode/.

[12] K. Golec-Biernat, APCheb33, the Fortran code to be obtained from the author; unpublished.

[13] K. Golec-Biernat, S. Jadach, M. Skrzypek, and W. Placzek, report IFJPAN-V-05-03.

[14] M. Ciafaloni, *Nucl. Phys.* **B296** (1988) 49;
S. Catani, F. Fiorani, and G. Marchesini, *Phys. Lett.* **B234** 339, *Nucl. Phys.* **B336** (1990) 18.

[15] P. Nason, *JHEP* **11** (2004) 040, `hep-ph/0409146`.

[16] D. Amati, A. Bassetto, M. Ciafaloni, G. Marchesini, and G. Veneziano, *Nucl. Phys.* **B173** (1980) 429.

[17] G. Curci, W. Furmanski, and R. Petronzio, *Nucl. Phys.* **B175** (1980) 27.

[4]The z-independent $\alpha_s(t)$ is set in front of the relevant flavour-changing kernels to simplify the program.

QED⊗QCD Exponentiation and Shower/ME Matching at the LHC*

B.F.L. Ward and S.A. Yost
Department of Physics, Baylor University, Waco, TX, USA

Abstract
We present the elements of QED⊗QCD exponentiation and its interplay with shower/ME matching in precision LHC physics scenarios. Applications to single heavy gauge boson production at hadron colliders are illustrated.

In the LHC environment, precision predictions for the effects of multiple gluon and multiple photon radiative processes will be needed to realize the true potential of the attendant physics program. For example, while the current precision tag for the luminosity at FNAL is at the $\sim 7\%$ level [1], the high precision requirements for the LHC dictate an experimental precision tag for the luminosity at the 2% level [2]. This means that the theoretical precision tag requirement for the corresponding luminosity processes, such as single W,Z production with the subsequent decay into light lepton pairs, must be at the 1% level in order not to spoil the over-all precision of the respective luminosity determinations at the LHC. This theoretical precision tag means that multiple gluon and multiple photon radiative effects in the latter processes must be controlled to the stated precision. With this objective in mind, we have developed the theory of $QED \otimes QCD$ exponentiation to allow the simultaneous resummation of the multiple gluon and multiple photon radiative effects in LHC physics processes, to be realized ultimately by MC methods on an event-by-event basis in the presence of parton showers in a framework which allows us to systematically improve the accuracy of the calculations without double-counting of effects in principle to all orders in both α_s and α.

Specifically, the new $QED \otimes QCD$ exponentiation theory is an extension of the QCD exponentiation theory presented in Refs. [3][1]. We recall that in the latter references it has been established that the following result holds for a process such as $q + \bar{q}' \to V + n(G) + X \to \bar{\ell}\ell' + n(g) + X$:

$$
\begin{aligned}
d\hat{\sigma}_{\rm exp} = \sum_n d\hat{\sigma}^n = e^{\rm SUM_{IR}(QCD)} \sum_{n=0}^{\infty} \int \prod_{j=1}^{n} \frac{d^3 k_j}{k_j} \\
\int \frac{d^4 y}{(2\pi)^4} e^{iy\cdot(P_1+P_2-Q_1-Q_2-\sum k_j)+D_{\rm QCD}} \\
* \tilde{\bar{\beta}}_n(k_1,\ldots,k_n) \frac{d^3P_2}{P_2^{\,0}} \frac{d^3Q_2}{Q_2^{\,0}}
\end{aligned}
\tag{1}
$$

where gluon residuals $\tilde{\bar{\beta}}_n(k_1,\ldots,k_n)$, defined by Ref. [3], are free of all infrared divergences to all orders in $\alpha_s(Q)$. The functions $SUM_{IR}(QCD), D_{\rm QCD}$, together with the basic infrared functions $B^{nls}_{QCD}, \tilde{B}^{nls}_{QCD}, \tilde{S}^{nls}_{QCD}$ are specified in Ref. [3]. Here $V = W^{\pm}, Z$,and $\ell = e, \mu,\ \ell' = \nu_e, \nu_\mu(e,\mu)$ respectively for $V = W^+(Z)$, and $\ell = \nu_e, \nu_\mu,\ \ell' = e, \mu$ respectively for $V = W^-$. We call attention to the essential compensation between the left over genuine non-Abelian IR virtual and real singularities between $\int dPh\bar{\beta}_n$ and $\int dPh\bar{\beta}_{n+1}$ respectively that really allows us to isolate $\tilde{\bar{\beta}}_j$ and distinguishes QCD from QED, where no such compensation occurs. The result in (1) has been realized by Monte

*Work partly supported by US DOE grant DE-FG02-05ER41399 and by NATO grant PST.CLG.980342.

[1]We stress that the formal proof of exponentiation in non-Abelian gauge theories *in the eikonal approximation* is given in Ref. [4]. The results in Ref. [3] are in contrast *exact* but have an exponent that only contains the leading contribution of the exponent in Ref. [4].

Carlo methods [3]. See also Refs. [5, 6, 7] for exact $\mathcal{O}(\alpha_s^2)$ and Refs. [8, 9, 10] for exact $\mathcal{O}(\alpha)$ results on the W,Z production processes which we discuss here.

The new $QED \otimes QCD$ theory is obtained by simultaneously exponentiating the large IR terms in QCD and the exact IR divergent terms in QED, so that we arrive at the new result

$$
\begin{aligned}
d\hat{\sigma}_{\rm exp} = e^{\rm SUM_{IR}(QCED)} \\
\sum_{n,m=0}^{\infty} \int \prod_{j_1=1}^{n} \frac{d^3 k_{j_1}}{k_{j_1}} \prod_{j_2=1}^{m} \frac{d^3 {k'}_{j_2}}{{k'}_{j_2}} \int \frac{d^4 y}{(2\pi)^4} \\
e^{iy\cdot(p_1+q_1-p_2-q_2-\sum k_{j_1}-\sum {k'}_{j_2})+D_{\rm QCED}} \\
\tilde{\bar{\beta}}_{n,m}(k_1,\ldots,k_n;k'_1,\ldots,k'_m)\frac{d^3p_2}{p_2^0}\frac{d^3q_2}{q_2^0},
\end{aligned}
\tag{2}
$$

where the new YFS [11, 12] residuals, defined in Ref. [13], $\tilde{\bar{\beta}}_{n,m}(k_1,\ldots,k_n;k'_1,\ldots,k'_m)$, with n hard gluons and m hard photons, represent the successive application of the YFS expansion first for QCD and subsequently for QED. The functions $\rm SUM_{IR}(QCED), D_{QCED}$ are determined from their analogs $\rm SUM_{IR}(QCD), D_{QCD}$ via the substitutions

$$
\begin{aligned}
B^{nls}_{QCD} &\to B^{nls}_{QCD} + B^{nls}_{QED} \equiv B^{nls}_{QCED},\\
\tilde{B}^{nls}_{QCD} &\to \tilde{B}^{nls}_{QCD} + \tilde{B}^{nls}_{QED} \equiv \tilde{B}^{nls}_{QCED},\\
\tilde{S}^{nls}_{QCD} &\to \tilde{S}^{nls}_{QCD} + \tilde{S}^{nls}_{QED} \equiv \tilde{S}^{nls}_{QCED}
\end{aligned}
\tag{3}
$$

everywhere in expressions for the latter functions given in Refs. [3]. The residuals $\tilde{\bar{\beta}}_{n,m}(k_1,\ldots,k_n;k'_1,\ldots,k'_m)$ are free of all infrared singularities and the result in (2) is a representation that is exact and that can therefore be used to make contact with parton shower MC's without double counting or the unnecessary averaging of effects such as the gluon azimuthal angular distribution relative to its parent's momentum direction.

In the respective infrared algebra (QCED) in (2), the average Bjorken x values

$$
\begin{aligned}
x_{avg}(QED) &\cong \gamma(QED)/(1+\gamma(QED))\\
x_{avg}(QCD) &\cong \gamma(QCD)/(1+\gamma(QCD))
\end{aligned}
$$

where $\gamma(A) = \frac{2\alpha_A \mathcal{C}_A}{\pi}(L_s-1)$, $A = QED, QCD$, with $\mathcal{C}_A = Q_f^2, C_F$, respectively, for $A = QED, QCD$ and the big log L_s, imply that QCD dominant corrections happen an order of magnitude earlier than those for QED. This means that the leading $\tilde{\bar{\beta}}_{0,0}$-level gives already a good estimate of the size of the interplay between the higher order QED and QCD effects which we will use to illustrate (2) here.

More precisely, for the processes $pp \to V + n(\gamma) + m(g) + X \to \bar{\ell}\ell' + n'(\gamma) + m(g) + X$, where $V = W^{\pm}, Z$,and $\ell = e, \mu,\ \ell' = \nu_e, \nu_\mu(e, \mu)$ respectively for $V = W^+(Z)$, and $\ell = \nu_e, \nu_\mu,\ \ell' = e, \mu$ respectively for $V = W^-$, we have the usual formula (we use the standard notation here [13])

$$
\begin{aligned}
d\sigma_{exp}(pp \to V + X \to \bar{\ell}\ell' + X') = \\
\sum_{i,j}\int dx_i dx_j F_i(x_i)F_j(x_j) d\hat{\sigma}_{exp}(x_i x_j s),
\end{aligned}
\tag{4}
$$

and we use the result in (2) here with semi-analytical methods and structure functions from Ref. [14]. A Monte Carlo realization will appear elsewhere [15].

We *do not* attempt in the *present* discussion to replace HERWIG [16] and/or PYTHIA [17] – we intend *here* to combine our exact YFS calculus with HERWIG and/or PYTHIA *by using the latter to generate a parton shower starting from the initial* (x_1, x_2) *point at factorization scale* μ *after this point is provided by the* $\{F_i\}$. This combination of theoretical constructs can be systematically improved with exact results order-by-order in α_s, where currently the state of the art in such a calculation is the work in Refs. [18] which accomplishes the combination of an exact $\mathcal{O}(\alpha_s)$ correction with HERWIG. We note that, even in this latter result, the gluon azimuthal angle is averaged in the combination. We note that the recent alternative parton distribution function evolution MC algorithm in Refs. [19] can also be used in our theoretical construction here. Due to its lack of the appropriate color coherence [20], we do not consider ISAJET [21] here.

To illustrate how the combination with Pythia/Herwig can proceed, we note that, for example, if we use a quark mass m_q as our collinear limit regulator, DGLAP [22] evolution of the structure functions allows us to factorize all the terms that involve powers of the big log $L_c = \ln \mu^2/m_q^2 - 1$ in such a way that the evolved structure function contains the effects of summing the leading big logs $L = \ln \mu^2/\mu_0^2$ where we have in mind that the evolution involves initial data at the scale μ_0. The result is therefore independent of m_q for $m_q \downarrow 0$. In the context of the DGLAP theory, the factorization scale μ represents the largest $p_\perp$ of the gluon emission included in the structure function. In practice, when we use these structure functions with an exact result for the residuals in (2), it means that we must in the residuals omit the contributions from gluon radiation at scales below μ. This can be shown to amount in most cases to replacing $L_s = \ln \hat{s}/m_q^2 - 1 \rightarrow L_{nls} = \ln \hat{s}/\mu^2$ but in any case it is immediate how to limit the p_T in the gluon emission [2] so that we do not double count effects. In other words, we apply the standard QCD factorization of mass singularities to the cross section in (2) in the standard way. We may do it with either the mass regulator for the collinear singularities or with dimensional regularization of such singularities – the final result should be independent of this regulator. This would in practice mean the following: We first make an event with the formula in (4) which would produce an initial beam state at (x_1, x_2) for the two hard interacting partons at the factorization scale μ from the structure functions $\{F_j\}$ and a corresponding final state X from the exponentiated cross section in $d\hat{\sigma}_{exp}(x_i x_j s)$; the standard Les Houches procedure [23] of showering this event (x_1, x_2, X) would then be used, employing backward evolution of the initial partons. If we restrict the p_T as we have indicated above, there would be no double counting of effects. Let us call this p_T matching of the shower from the backward evolution and the matrix elements in the QCED exponentiated cross section.

However, one could ask if it is possible to be more accurate in the use of the exact result in (2)? Indeed, it is. Just as the residuals $\tilde{\bar{\beta}}_{n,m}(k_1, \ldots, k_n; k'_1, \ldots, k'_m)$are computed order by order in perturbation theory from the corresponding exact perturbative results by expanding the exponents in (2) and comparing the appropriate corresponding coefficients of the respective powers of $\alpha^n \alpha_s^m$, so too can the shower formula which is used to generate the backward evolution be expanded so that the product of the shower formula's perturbative expansion, the perturbative expansion of the exponents in (2), and the perturbative expansions of the residuals can be written as an over-all expansion in powers of $\alpha^n \alpha_s^m$ and required to match the respective calculated exact result for given order. In this way, new shower subtracted residuals, $\{\hat{\tilde{\bar{\beta}}}_{n,m}(k_1, \ldots, k_n; k'_1, \ldots, k'_m)\}$, are calculated that can be used for the entire gluon p_T phase space with an accuracy of the cross section that should in principle be improved compared with the first procedure for shower matching presented above. Both approaches are under investigation.

Returning to the general discussion, we compute, with and without QED, the ratio $r_{exp} = \sigma_{exp}/\sigma_{Born}$, where we do not use the narrow resonance approximation; for, we wish to set a paradigm for precision heavy vector boson studies. The formula which we use for σ_{Born} is obtained from that in (4) by substituting $d\hat{\sigma}_{Born}$ for $d\hat{\sigma}_{\rm exp}$ therein, where $d\hat{\sigma}_{Born}$ is the respective parton-level Born cross section. Specifically, we have from (1) the $\tilde{\bar{\beta}}_{0,0}$-level result

$$\hat{\sigma}_{exp}(x_1 x_2 s) = \int_0^{v_{max}} dv \gamma_{QCED} v^{\gamma_{QCED}-1} F_{YFS}(\gamma_{QCED}) e^{\delta_{YFS}} \hat{\sigma}_{Born}((1-v)x_1 x_2 s) \qquad (5)$$

where we intend the well-known results for the respective parton-level Born cross sections and the value of v_{max} implied by the experimental cuts under study. What is new here is the value for the QED$\otimes$QCD

[2]Here, we refer to both on-shell and off-shell emitted gluons.

exponent

$$\gamma_{QCED} = \left\{ 2Q_f^2 \frac{\alpha}{\pi} + 2C_F \frac{\alpha_s}{\pi} \right\} L_{nls} \tag{6}$$

where $L_{nls} = \ln x_1 x_2 s/\mu^2$ when μ is the factorization scale. The functions $F_{YFS}(\gamma_{QCED})$ and $\delta_{YFS}(\gamma_{QCED})$ are well-known [12] as well:

$$\begin{aligned} F_{YFS}(\gamma_{QCED}) &= \frac{e^{-\gamma_{QCED}\gamma_E}}{\Gamma(1+\gamma_{QCED})}, \\ \delta_{YFS}(\gamma_{QCED}) &= \frac{1}{4}\gamma_{QCED} + (Q_f^2 \frac{\alpha}{\pi} + C_F \frac{\alpha_s}{\pi})(2\zeta(2) - \frac{1}{2}), \end{aligned} \tag{7}$$

where $\zeta(2)$ is Riemann's zeta function of argument 2, i.e., $\pi^2/6$, and γ_E is Euler's constant, i.e., 0.5772... . Using these formulas in (4) allows us to get the results

$$r_{exp} = \begin{cases} 1.1901 & \text{, QCED} \equiv \text{QCD+QED, LHC} \\ 1.1872 & \text{, QCD, LHC} \\ 1.1911 & \text{, QCED} \equiv \text{QCD+QED, Tevatron} \\ 1.1879 & \text{, QCD, Tevatron.} \end{cases} \tag{8}$$

We see that QED is at the level of .3% at both LHC and FNAL. This is stable under scale variations [13]. We agree with the results in Refs. [8, 9, 10, 5, 6] on both of the respective sizes of the QED and QCD effects. The QED effect is similar in size to structure function results found in Refs. [24, 25, 26, 27, 28], for further reference.

We have shown that YFS theory (EEX and CEEX) extends to non-Abelian gauge theory and allows simultaneous exponentiation of QED and QCD, QED$\otimes$QCD exponentiation. For QED$\otimes$QCD we find that full MC event generator realization is possible in a way that combines our calculus with Herwig and Pythia in principle. Semi-analytical results for QED (and QCD) threshold effects agree with literature on Z production. As QED is at the .3% level, it is needed for 1% LHC theory predictions. We have demonstrated a firm basis for the complete $\mathcal{O}(\alpha_s^2, \alpha\alpha_s, \alpha^2)$ results needed for the FNAL/LHC/RHIC/ILC physics and all of the latter are in progress.

Acknowledgments

One of us (B.F.L.W.) thanks Prof. W. Hollik for the support and kind hospitality of the MPI, Munich, while a part of this work was completed.

References

[1] S. Klimenko, in *Proc. HCP2002*, ed. M. Erdmann,(Karlsruhe, 2002) p.413.

[2] M. Dittmar, F. Pauss and D. Zurcher,*Phys. Rev.* D**56** (1997) 7284; M. Rijssenbeek, in *Proc. HCP2002*, ed. M. Erdmann,(Karlsruhe, 2002)p. 424; M. Dittmar, *ibid.*,p.431.

[3] B.F.L. Ward and S. Jadach, *Acta Phys.Polon.* **B33** (2002) 1543; in *Proc. ICHEP2002*, ed. S. Bentvelsen *et al.*,(North Holland, Amsterdam, 2003) p. 275 and references therein.

[4] J.G.M. Gatherall, Phys. Lett. B**133** (1983) 90.

[5] R. Hamberg, W. L. van Neerven and T Matsuura, *Nucl. Phys.*B**359** (1991) 343.

[6] W.L. van Neerven and E.B. Zijlstra, *Nucl. Phys.* B**382** (1992) 11; *ibid.* B**680** (2004) 513; and, references therein.

[7] C. Anastasiou *et al.*,*Phys. Rev.* D**69** (2004) 094008.

[8] U. Baur, S. Keller and W.K. Sakumoto, *Phys. Rev.* D **57** (1998) 199; U. Baur, S. Keller and D. Wackeroth, *ibid.***59** (1998) 013002; U. Baur *et al.*, *ibid.***65** (2002) 033007, and references therein.

[9] S. Dittmaier and M. Kramer, *Phys. Rev.* D**65** (2002) 073007, and references therein

[10] Z. A. Zykunov,*Eur. Phys. J.*C**3** (2001) 9, and references therein.

[11] D. R. Yennie, S. C. Frautschi, and H. Suura,*Ann. Phys.* **13** (1961) 379; see also K. T. Mahanthappa, *Phys. Rev.* **126** (1962) 329, for a related analysis.

[12] See also S. Jadach *et al.*,*Comput. Phys. Commun.* **102** (1997) 229; S. Jadach, M. Skrzypek and B.F.L. Ward,*Phys.Rev.* D**55** (1997) 1206; S. Jadach, B.F.L. Ward and Z. Was,*Phys. Rev.* D**63** (2001) 113009; S. Jadach, B.F.L. Ward and Z. Was, *Comp. Phys. Commun.* **130** (2000) 260; S. Jadach *et al.*, *ibid.***140** (2001) 432, 475.

[13] C. Glosser *et al.*,*Mod. Phys. Lett.* A **19** (2004) 2119.

[14] A.D. Martin et al., *Phys. Rev.* D**51** (1995) 4756.

[15] S. Jadach *et al.*, to appear.

[16] G. Corcella et al., hep-ph/0210213, and references therein.

[17] T. Sjostrand et al., hep-ph/0308153.

[18] S. Frixione and B. Webber, J. High Energy Phys.**0206**(2002) 029; S. Frixione, P. Nason and B. Webber, *ibid.***0308** (2003) 007, and references therein.

[19] S. Jadach and M. Skrzypek, *Acta Phys. Pol.* B**35** (2004) 735; hep-ph/0504263, 0504205.

[20] Michelangelo Mangano, private communication, 2004.

[21] F. Paige et al., hep-ph/0312045, and references therein.

[22] G. Altarelli and G. Parisi, Nucl. Phys. **B126** (1977) 298; Yu. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641; L. N. Lipatov, Yad. Fiz. **20** (1974) 181; V. Gribov and L. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 675, 938.

[23] E. Boos et al., hep-ph/0109068.

[24] S. Haywood, P.R. Hobson, W. Hollik and Z. Kunszt, in *Proc. 1999 CERN Workshop on Standard Model Physics (and more) at the LHC, CERN-2000-004*, eds. G. Altarelli and M.L. Mangano,(CERN, Geneva, 2000) p. 122.

[25] H. Spiesberger,*Phys. Rev.* D**52** (1995) 4936.

[26] W.J. Stirling,"Electroweak Effects in Parton Distribution Functions", talk presented at ESF Exploratory Workshop, *Electroweak Radiative Corrections to Hadronic Observables at TeV Energies* , Durham, Sept., 2003.

[27] M. Roth and S. Weinzierl,*Phys. Lett.* **B590** (2004) 190.

[28] W. J. Stirling et al., in *Proc. ICHEP04*, eds. H. Chen et al. (World Sci. Publ., Singapore, 2005) p. 527.

PHOTOS as a pocket parton shower: flexibility tests for the algorithm*

Piotr Golonka and Zbigniew Was
CERN, 1211 Geneva 23, Switzerland, and Institute of Nuclear Physics, ul. Radzikowskiego 152, 31-342 Kraków, Poland

Abstract
PHOTOS is widely used for generation of bremsstrahlung in decays of particles and resonances in LHC applications. We document here its recent tests and variants. Special emphasis is on those aspects which may be useful for new applications in QED or QCD.

Recently version 2.14 of the PHOTOS Monte Carlo algorithm, written for bremsstrahlung generation in decays became available. In ref. [1] detailed instructions on how to use the program are given. With respect to older versions [2, 3] of PHOTOS, it now features: improved implementation of QED interference and multiple-photon radiation. The numerical stability of the code was significantly improved as well. Thanks to these changes, PHOTOS generates bremsstrahlung corrections in Z and W decays with a precision of 0.1%. This precision was established in [4] with the help of a multitude of distributions and of a specially designed numerical test (SDP), see ref. [1], section 5 for the definition. The tests for other channels, such as semileptonic K decays and leptonic decays of the Higgs boson and the τ-lepton, are presented in [4] as well. In those cases the level of theoretical sophistication for the reference distribution was lower though.

In this note we will not repeat a discussion of the design properties, but we will recall the main tests that document robustness and flexibility of the PHOTOS design. The results of the comparisons of PHOTOS running with different options of separation of its physical content into functional parts of the algorithm will be shown. The design of the program, i.e. the relation between the parts of the algorithm remained unchanged for these tests. This aspect may be of broader use and may find extensions in future applications, also outside the simple case of purely QED bremsstrahlung in decays.

In the calculations that led to the construction of PHOTOS we had to deal with the diagrams generated by photon couplings to the charged fermions, scalars or vectors. They were definitely simpler than the ones required for the QCD, nonetheless they offered a place to develop solutions which may be of some use there as well. Having such possibility in mind, yet not having any extension to QCD at hand, we have called PHOTOS a *pocket parton shower*. We hope that the methods we developed would be useful for QCD at least as pedagogical examples.

We begin with a presentation of the components of the PHOTOS algorithm using operator language. The consecutive approximations used in the construction of the crude distribution for photon generation, and the correcting weights used to construct the physically complete distributions are listed, but can not be defined in detail here. Instead, we present the variations of the algorithm. Comparisons between different options of the algorithm provide an important class of technical tests, and also help to explore the limits of the universality of the PHOTOS solution. The results of some of these tests will be listed later in the contribution (for the remaining ones and the details we address the reader to refs. [1, 4]). In the comparisons we use the SDP universal test based on MC-TESTER [5] as in ref. [1]. We skip its definition here as well.

The starting point for the development of PHOTOS was the observation that, at first order, the bremsstrahlung corrections in the $Z \to \mu^+\mu^-$ process can be written as a convolution of the Born-level distribution with the single-photon emission kernels for the emission from μ^+ and μ^-.

*Supported in part by the EU grant MTKD-CT-2004-510126, in partnership with the CERN Physics Department, and the Polish State Committee for Scientific Research (KBN) grant 2 P03B 091 27 for the years 2004–2006.

The formulae for the emission kernels are 3-dimensional and can be parametrized using the angles and the invariant mass, which are the same variables as those used in the parametrization of the three-body phase space (the kernels use only a subset of the complete set of phase-space parametrization variables). The remaining two angular variables, not used in the kernels, can be identified as the angles defining the orientation of the μ^+ and/or μ^- directions (for a detailed definition, see e.g. [2]).

The principle of the single-photon algorithm working on n-body decay is to replace a point in the n-body phase space Ω_2, with either the point in the original Ω_2, or the point in the $(n+1)$-body phase space Ω_3 (with generated photon). The overall normalization of the decay rate has to change as well and, for example, in the case of $Z \to \mu^+\mu^-$, due to the action of the single-photon algorithm, it needs to be multiplied by a factor of $1 + \frac{3}{4}\frac{\alpha}{\pi}$.

Subsequent steps of the PHOTOS algorithm are described in terms of the evolution operators. Let us stress the relations of these operators to the matrix elements and phase-space parametrizations. We will present the decomposition of the operators in the top–down order, starting with the definition of R_α, the operator describing the complete PHOTOS algorithm for single emission (which at least in the case of Z and leptonic τ decays originates from field theory calculations without any approximation). Then, we will gradually decompose the operators (they differ from decay channel to decay channel) so that we will end up with the single well-defined, elementary operator for the emission from a single charged particle in the final state. By aggregation of these elementary operators, the R_α may be reconstructed for any decay channel. Let us point out that the expression of theoretical calculations in the form of operators is particularly suitable in computer programs implementation.

We skip here a separate discussion of the factorization properties, in particular to define/optimize the way the iteration of R's is performed in PHOTOS. Not only the first-order calculations are needed, but also higher-order ones, including mixed virtual–real corrections. For practical reasons, the R_α operator needs to be regularized with the minimum energy for the explicitly generated photons: the part of the real-photon phase space, under threshold, is integrated, and the resulting factor is summed with the virtual correction.

• 1

Let us define the five steps in R_α separation. In the first one, the R_α is replaced by (we use two-body decay as an example) $R_\alpha = R_I(R_S(\mu^+) + R_S(\mu^-))$, where R_I is a generalized interference operator and R_S is a generalized operator responsible for photon generation from a single, charged decay-product.

Let us point out here, that we use the word *interference* here having in mind its usual quantum-mechanical sense. The interference is introduced simultaneously for the real and the virtual photon correction. As a consequence, it changes, for instance, the hard-photon energy spectrum, and the action of R_I looks like kinematic reshuffling of events around the phase space. This interpretation of the interference was particularly clear in the case of the Z decays where the R_I operator can introduce *exact and complete* first-order radiative corrections.

It is important to firstly define the amplitudes, the sum of which is squared, in physically meaningful way, that is in gauge-invariant way, to produce interference. Our approach has changed with time, and we relaxed this requirement; at present we simply request that the action of R_I properly introduces interference effects. We also require that the generalized interference operator respects energy–momentum conservation, and also overall normalization of the distribution under construction. The freedom of choice in the separation of R_α into R_I and R_S we obtained this way is used to create different variants of the PHOTOS algorithm.

The R_S operator acts on the points from the Ω_2 phase space, and the results of its action belong either to Ω_2 or to Ω_3. The domain of the R_I operator has to be Ω_2 + Ω_3, and the results are also in $\Omega_2 + \Omega_3$. In our solution we required that R_I acts as a unit operator on the Ω_2-part of its domain and, with some probability, returns the points from Ω_3 back to the original points in Ω_2, thus reverting the action of the R_S.

Let us stress that in practical applications, to ease the extension of the algorithm to "any" decay mode, we used in PHOTOS a simplification for R_I. Obviously, the exact representation of the first-order result would require R_I to be decay-channel-dependent. Instead, we used an approximation that ensures the proper behaviour of the photon distribution in the soft limit. Certain deficiencies at the hard-photon limit of the phase space appear as a consequence, and are the subject of studies that need to be performed individually for every decay channel of interest. The comparisons with matrix-element formulae, as in [6], or experimental data, have to be performed for the sake of precision; they may result in dedicated weights to be incorporated into PHOTOS. In principle, there is no problem to install a particular decay-channel matrix element, but there has not been much need for this yet. So far, the precision of the PHOTOS algorithm could always be raised to a satisfactory level by implementing some excluded parts of formulae, being the case of W decay [6] an exception.

The density generated by the R_S operator is normally twice that of real photons at the end of generation and all over the phase space; it can also overpopulate only those regions of phase space where it is necessary for R_I. The excess of these photons is then reduced by Monte Carlo with the action of R_I.

• 2

In the next step of the algorithm construction, we have separated $R_S = R_B R_A$, where R_B was responsible for the implementation of the spin-dependent part of the emission, and the R_A part was independent of the spin of the emitting final-state particle. Note that this step of the algorithm can be performed at the earlier stage of generation as well, that is before the full angular construction of the event. R_B is again, as R_I, it moves the hard bremsstrahlung events in excess back to the original no-bremsstrahlung ones. R_B operates on the internal variables of PHOTOS rather than on the fully constructed events.

• 3

The definition of the R_I, R_B, R_A operators was initially based on the inspection of the first-order matrix elements for the two-body decays. In the general solution for R_A, the process of multiple-body decay of particle X is temporarily replaced by the two-body decay $X \to CY$, in which particle X decays to the charged particle C, which "emits" the photon, and the "spectator system" Y. The action of the operator is repeated for each charged decay product: the subsequent charged particle takes the role of the photon emitter C; all the others, including the photons generated in the previous steps, become a part of the spectator system Y. The independence of the emissions from each charged product then has to be ensured. This organization works well and can be understood with the help of the exact parametrization of multibody phase space. It is helpful for iteration in multiple-photon emission. It also helps to implement some genuine second-order matrix elements. This conclusion can be drawn from an inspection of the second-order matrix elements, as in [7].

• 4

In the next step, we decompose the R_A operator, splitting it in two parts: $R_A = R_a R_x$. The R_x operator generates the energy of the (to be generated) photon, and R_a generates its explicit kinematic configuration.

The R_x operator acts on points from the Ω_2 phase space, and generates a single real number x; the R_a operator transforms this point from Ω_2 and the number x to a point in Ω_3, or leaves the original point in Ω_2. Note that again, as R_I, the R_a operator has to be unitary and has to conserve energy–momentum[1].

[1]On the contrary, the R_x operator can not, in general, fulfill the unitarity requirement. For example, the part of R_α leading to $1 + \frac{3}{4}\frac{\alpha}{\pi}$ for the Z decay can not be placed elsewhere but in R_x. The energy–momentum conservation does not apply directly to R_x, as it does not change the kinematic configuration, but only supplements it with x– the energy of the photon to be generated. However, for multiple-photon generation, the limits for generated x for subsequent generated photons are the same as for the first photon, which may be in potential conflict with energy–momentum conservation constraint.

An analogy between R_x and the kernel for structure-function evolution should be mentioned. However, there are notable differences: the x variable is associated more with the ratio of the invariant mass of decay products of X, photon excluded, and the mass of X, than with the fraction of energy taken away by the photons from the outgoing charged product C. Also, R_x can be simplified by moving its parts to R_a, R_S or even R_I. Note that in R_x the contributions of radiation from all charged final states are summed.

• 5

The R_x operator is iterated, in the solutions for double, triple, and quartic photon emission. The iterated R_x can also be shifted and grouped at the beginning of the generation, because they are free from the phase-space constraints. The iterated R_x takes a form similar to a formal solution for structure-function evolution, but with exceptionally simple kernels. The phase-space constraints are introduced later, with the action of the R_a operators. Because of this, the iteration of R_x can go up to fixed or infinite order. The algorithm is then organized in two steps. At first, a crude distribution for the number of photon candidates is generated; then, their energies are defined. For that purpose we can perform a further separation: $R_x = R_f R_0 R_N$, where the R_0 operator determines whether a photon candidate has to be generated at all, and R_f defines the fraction of its energy (without energy–momentum-conservation constraint). From the iteration of R_0, we obtain a Poisson distribution, but any other analytically solvable distribution would be equally good.

The overall factor, such as $1 + \frac{3}{4}\frac{\alpha}{\pi}$ in Z leptonic partial width, does not need to be lost. It finds its way to the R_N, which is a trivial overall normalization constant in the case of the final-state radiation discussed here. In the cases where precision requirements are particularly high, the users of PHOTOS should include this (process-dependent) factor into the decay tables in their main generator for decays. However, until now, the effects on the normalization due to R_N are too small and were usually neglected. We rise the attention to this point, because it may be important for generalizations, when different organization of R_f, R_0 and R_N may be enforced by the properties of the matrix elements.

—

The input data for the algorithm are taken from the event record, the kinematic configurations of all particles, and the mother–daughter relations between particles in the decay process (which could be a part of the decay cascade) should be available in a coherent way.

This wraps up, a basic, presentation of the steps performed by the PHOTOS algorithm. For more details see [1, 8].

Tests performed on the algorithm:

1. The comparison of PHOTOS running in the quartic-photon emission mode and the exponentiated mode for the leptonic Z and W decays may be found on our web page which documents the results of the tests [4]. The agreement in branching ratios and shapes of the distributions is better than 0.07% for all the cases that were tested. It can be concluded that changing the relative order for the iterated R_0 and the rest of R_α operators does not lead to significant differences. This test, if understood as a technical test, is slightly biased by the uncontrolled higher-than-fourth-order terms which are missing in the quartic-emission option of PHOTOS. Also, the technical bias, due to the minimal photon energy in generation, present in the fixed-order options of PHOTOS may contribute to the residual difference.
2. The comparison of PHOTOS with different options for the relative separation between R_I and R_S. The tests performed for the fixed-order and exponentiated modes indicated that the differences in results produced by the two variants of the algorithm are below the level of statistical error for the runs of 10^8 events. In the code these two options are marked respectively as VARIANT-A and VARIANT-B.
3. The comparisons of PHOTOS with different algorithms for the implementation of the R_I operator. In PHOTOS up to version 2.12, the calculations were performed using internal variables in the

angular parametrization. This algorithm was limited to the cases of decays of a neutral particle into two charged particles. In later versions, the calculations are performed using the 4-momenta of particles, hence for any decay mode. The tests performed for leptonic Z decays indicated that the differences are below the statistical error of the runs of 10^8 events.

4. The comparisons of PHOTOS with different options for the relative separation between R_0 and R_x, consisting of an increase in the crude probability of hard emission at R_0. The tests performed for the exponentiated mode of PHOTOS indicated that the differences are below the statistical error of the runs of up to 10^8 events.
5. The remaining tests, including new tests for the effects of the interference weights in cascade decays, are more about the physics content of the program than on the technical or algorithmic aspects. They are presented in ref. [1] and the results are collected on the web page [4].

Multiple options for PHOTOS running and technical compatibility of results even for 10^8 event samples generated in a short CPU cycle time are encouraging. They indicate the potential for algorithm extensions. Note that PHOTOS was found to work for decays of up to 10 charged particles in the final state.

Acknowledgments: The authors are indebted to members of BELLE, BaBar, NA48, KTeV, ATLAS, CMS, D0, CDF collaborations for useful comments and suggestions.

References

[1] P. Golonka and Z. Was, hep-ph/0506026.

[2] E. Barberio, B. van Eijk, and Z. Was, *Comput. Phys. Commun.* **66** (1991) 115.

[3] E. Barberio and Z. Was, *Comput. Phys. Commun.* **79** (1994) 291–308.

[4] P. Golonka and Z. Was, see http://cern.ch/Piotr.Golonka/MC/PHOTOS-MCTESTER.

[5] P. Golonka, T. Pierzchala, and Z. Was, *Comput. Phys. Commun.* **157** (2004) 39–62, hep-ph/0210252.

[6] G. Nanava and Z. Was, *Acta Phys. Polon.* **B34** (2003) 4561–4570, hep-ph/0303260.

[7] Z. Was, hep-ph/0406045.

[8] P. Golonka, *In preparation*. PhD thesis, Institute of Nuclear Physics, Krakow, 2005/2006 Written under the supervision of Z. Was; a preliminary version will be available at http://cern.ch/Piotr.Golonka/MC/PhD.

3 Working group 3: Heavy Quarks (charm and beauty)

List of participants in the working group:

Baines John, Baranov Sergey, Barbagli Giuseppe, Behnke Olaf, Bertolin Alessandro, Bluemlein Johannes, Bruno Giuseppe, Buchmueller Oliver, Butterworth Jonathan, Cacciari Matteo, Carli Tancredi, Catmore James, Chiochia Vincenzo, Cholewa Axel, Cole Jo, Corcella Gennaro, Corradi Massimo, Dainese Andrea, Ellis Keith, Eskola Kari, Eynck, T.O., Fabris Daniela, Flucke Gero, Geiser Achim, Grab Christoph, Grindhammer Guenter, Guernane Rachid, Gutsche Oliver, Jung Hannes, Kiesling Christian, Klanner Robert, Kolhinen Vesa, Kramer Gustav, Kretzer Stefan, Kuhr Thomas, Kurepin Alexey, Kutak Krzysztof, Laenen Eric, Lagouri Theodota, Levin Eugene, Likhoded Anatoly, Lipatov Artem, Lunardon Marcello, Maltoni Fabio, Mangano Michelangelo, Marti-Magro Lluis, Meyer Andreas, Morando Maurizio, Morsch Andreas, Panikashvili Natalia, Pavel Nikolaj, Peters Krisztian, Piskounova Olga, Ranieri Riccardo, Ruiz Hugo, Schienbein Ingo, Sharafiddinov Rasulkhozha, Shulha Siarhei, Smizanska Maria, Sridhar Krishnamoorthy, Staykova Zlatka, Thompson P., Thorne Robert, Tonazzo Alessandra, Turrisi Rosario, Uwre Ulrich, Villa Mauro, Vogt Ramona, Vulpescu Bogdan, Was Z., Weiser Christian, Wing Matthew, Zoccoli Antonio, Zotov Nikolai, zur Nedden Martin

Introduction to Heavy Quarks (charm and beauty)

O. Behnke, M Cacciari, M. Corradi, A. Dainese A. Geiser, A. Meyer, M. Smizanska, U. Uwer C. Weiser

Perturbative QCD is expected to provide reliable predictions for the production of bottom and (to a lesser extent) charm quarks since their masses are large enough to assure the applicability of perturbative calculations. Anyway a direct comparison of perturbative QCD predictions to heavy flavour production data is not straightforward. Difficulties arise from the presence of scales very different from the quark masses that reduce the predictivity of fixed-order theory, from the non-perturbative ingredients needed to parametrize the fragmentation of the heavy quarks into the observed heavy hadrons and from the limited phase space accessible to present detectors. Moreover a breakdown of the standard collinear factorization approach can be expected at low-x. The study of heavy quark production in hadronic interactions and in ep collisions at HERA has been therefore an active field, in the effort to overcome these difficulties and to get a deeper understanding of hard interactions.

Besides its intrinsic interest, a precise understanding of heavy quark production is important at LHC because charm and beauty from QCD processes are relevant backgrounds to other interesting processes from the Standard Model (e.g. Higgs to $b\bar{b}$) or beyond. Moreover, theoretical and experimental techniques developed at HERA in the heavy quark field, such as heavy-quark parton densities or b-tagging, are also of great value for future measurements at the LHC.

The present status of the heavy-quark production theory is critically reviewed in the first contribution. The second contribution summarises the present heavy flavour data from HERA and gives an outlook of what can be expected from HERA-II. The potential of the LHC experiments for charm and beauty physics is reviewed in the 3rd contribution. Then the relevance of saturation and low-x effects to heavy quark production at HERA and at the LHC are discussed. The non-perturbative aspects of heavy-quark fragmentation and their relevance to HERA and LHC are discussed in the next contribution. Finally, a comparison of different theoretical predictions for HERA and LHC based on different approaches is presented.

Theoretical review of various approaches in heavy quark production

Coordinators: M. Cacciari[1], E. Laenen[2]
Contributing authors: S.P. Baranov[3], M. Cacciari[1], S. Diglio[4], T.O. Eynck[2], H. Jung[5], B.A. Kniehl[6], S. Kretzer[7], E. Laenen[2], A.V. Lipatov[8], F. Maltoni[9], F. Petrucci[4], O.I. Piskounova[3], I. Schienbein[6], J. Smith[2,10], A. Tonazzo[4], M. Verducci[11], N.P. Zotov[8]
[1]LPTHE - Université P. et M. Curie (Paris 6), France
[2]NIKHEF Theory Group, Kruislaan 409, 1098 SJ Amsterdam, The Netherlands
[3]P.N. Lebedev Physical Institute of Russian Academy of Science, Moscow, Russia
[4]Università Roma Tre, Dipartimento di Fisica "E.Amaldi" and
INFN Sezione Roma III, Via della Vasca Navale 84, 00146 Rome, Italy
[5]Deutsches Elektronen-Synchroton Hamburg, FRG
[6]II. Institut für Theoretische Physik, Universität Hamburg, Luruper Chaussee 149, 22761, Hamburg, Germany
[7] Brookhaven National Laboratory, Upton, NY 11973, USA
[8]D.V. Skobeltsyn Institute of Nuclear Physics, Moscow, Russia
[9]Institut de Physique Théorique, Université Catholique de Louvain,Chemin du Cyclotron, 2, B-1348, Louvain-la-Neuve, Belgium
[10]C.N. Yang Institute for Theoretical Physics, Stony Brook University, Stony Brook, NY 11794, USA
[11]CERN, CH-1211 Genève 23, Switzerland

Abstract
We review some of the main theoretical aspects of heavy quark production at HERA that will be important for understanding similar processes at the LHC.

1 Introduction

The value for the LHC physics program of heavy quark production studies at HERA consists not only of measured quantities such as parton distributions, heavy quark masses etc. but at least as much of the theoretical ideas on heavy quark production that were developed and refined in the course of these studies. The strong experimental interest in heavy quark observables at HERA has led to a significantly increased understanding of the benefits and limitations of finite order calculations. It has stimulated theorists to deepen their insight into the issue of when a heavy quark should be treated as a parton, and it has provoked novel proposals to explain the hadronization of heavy quarks. In what follows we review and critically assess some of these ideas.

2 Heavy quark production

The study of heavy quarks, historically plagued by low production rates and large uncertainties, has now entered the regime of 'precision physics'. On the one hand, the larger centre-of-mass energies of the colliders running now (Tevatron, HERA) and in the near future (LHC) lead to a much more copious production yield. On the other hand, technological advances such as the introduction of microvertex detectors based on semiconductor devices allow for much better tagging of the produced heavy flavours, and hence better measurements. Needless to say, an equally substantial improvement of the theoretical calculations has been needed in order to match this progress and therefore deliver predictions with an accuracy at least as good as that of the experimental measurements. Properly testing and constraining the theoretical calculations will in turn help in refining the predictions for the LHC.

One example for which a good theoretical accuracy at the LHC is desirable is in calculating the total Z boson production rate, a process which can be used as a luminosity candle and which we would like to have under control at the one per cent level. One channel contributing to this process is gluon-gluon fusion followed by bottom-antibottom annihilation, $gg \to b\bar{b} \to Z$. This channel provides about 5% of the total Z yield [1]: hence, it must be under control at the 20% level in order to achieve the sought-for final 1% accuracy.

As it turns out, it is more efficient and more reliable to rewrite this in terms of a perturbatively calculated parton distribution function (PDF) for the bottom quark, i.e. as the effective process $b\bar{b} \to Z$. The theoretical tools that we use to construct such heavy quark parton distribution functions must therefore be tested by employing them in other theoretical predictions, to be compared to the available experimental data. In the following section we shall list a number of examples where this is done.

From the point of view of 'standard' perturbative QCD calculations, the situation has not changed since the beginning of the '90s: fully massive next-to-leading order (NLO) calculations were made available for hadron-hadron [2, 3, 4, 5, 6], photon-hadron [7, 8, 9] (i.e. photoproduction) and electron-hadron [10, 11, 12, 13] (i.e. Deep Inelastic Scattering, DIS) collisions. These calculations still constitute the state of the art as far as fixed order results are concerned, and they form the basis for all modern phenomenological predictions.

Over the years, and with increasing experimental accuracies, it however became evident that perturbative QCD alone did not suffice. In fact, real particles - hadrons and leptons - are observed in the detectors, not the unphysical quarks and gluons of perturbative QCD. A proper comparison between theory and experiment requires that this gap be bridged by a description of the transition. Of course, the accuracy of such a description will reflect on the overall accuracy of the comparison. When the precision requirements were not too tight, one usually employed a Monte Carlo description to 'correct' the data, deconvoluting hadronization effects and extrapolating to the full phase space. The final 'experimental' result could then easily be compared to the perturbative calculation. This procedure has the inherent drawback of including the bias of our theoretical understanding (as implemented in the Monte Carlo) into an experimental measurement. This bias is of course likely to be more important when the correction to be performed is very large. It can sometimes become almost unacceptable, for instance when exclusive measurements are extrapolated by a factor of ten or so in order to produce an experimental result for a total photoproduction cross section or a heavy quark structure function.

The alternative approach is to present (multi)differential experimental measurements, with cuts as close as possible to the real ones, which is to say with as little theoretical correction/extrapolation as possible. The theoretical prediction must then be refined in order to compare with the real data that it must describe. This has two consequences. First, one has to deal with differential distributions which, in certain regions of phase space, display a bad convergence in perturbation theory. All-order resummations must then be performed in order to produce reliable predictions. Second, differential distributions of real hadrons depend unavoidably on some non-perturbative phenomenological inputs, fragmentation functions. Such inputs must be extracted from data and matched to the perturbative theory in a proper way, pretty much like parton distribution functions of light quarks and gluons are.

In the following sections we review the state of the art of theoretical calculations of heavy quark production in a number of high energy processes, pointing out similarities and differences. In particolar, resummations aimed at improving the theoretical description of heavy quark production at large transverse momentum or large photon virtuality in DIS (Section 3), small centre-of-mass energy (Section 5) and large centre of-mass energy (Section 6) are described in some detail.

3 Collinear resummations and heavy quark PDFs

Perturbative calculations of heavy quark production contain badly converging logarithmic terms of quasi-collinear origin in higher orders when a second energy scale is present and it is much larger than the heavy quark mass m. Examples are the (square root of the) photon virtuality Q^2 in DIS and the transverse momentum p_T in either hadroproduction or photoproduction. Naming generically E the large scale, we can write schematically the cross section for the production of the heavy quark Q as

$$\sigma_Q(E,m) = \sigma_0 \left(1 + \sum_{n=1} \alpha_s^n \sum_{k=0}^{n} c_{nk} \ln^k \left[\frac{E^2}{m^2} + \mathcal{O}\left(\frac{m}{E}\right) \right] \right) , \tag{1}$$

where σ_0 stands for the Born cross section, and the coefficients c_{nk} can contain constants as well as functions of m and E, vanishing as powers of m/E when $E \gg m$.

Resummation approaches bear many different names, (ZM-VFNS, ACOT, FONLL, BSMN to name but a few) but they all share the goal of resumming leading ($\alpha_s^n \ln^n(E^2/m^2)$, LL) and sometimes also next-to-leading ($\alpha_s^n \ln^{n-1}(E^2/m^2)$, NLL) logarithmic terms to all orders in the cross section above. This is achieved by discarding power suppressed m/E terms, and factoring all the logarithms into a resummation factor, to be obtained via Altarelli-Parisi evolution of an initial condition set at the heavy quark mass scale,

$$\sigma_Q^{res}(E,m) = \sigma_0 C(E,\mu) f(\mu,m) = \sigma_0 C(E,\mu) E(\mu,\mu_0) f(\mu_0,m) , \tag{2}$$

where μ and μ_0 represent artificial factorization scales, to be taken of order E and m respectively. The 'products' between the various functions actually hide convolution operations with respect to momentum fractions, not explicitly shown as arguments. $C(E,\mu)$ is a perturbatively calculable coefficient function, which does not contain large logarithms thanks to the choice $\mu \simeq E$. The function $f(\mu,m)$ can represent either a parton distribution or a fragmentation function for a heavy quark, and contains the resummation of the collinear logarithms. Due to the large heavy quark mass, its initial condition $f(\mu_0,m)$ can be calculated in perturbation theory [14, 15]: this is the distinctive feature that sets heavy quark parton and fragmentation functions apart from light flavour ones, whose initial conditions are instead entirely non-perturbative and must be fitted to data.

Once a massless but resummed result, valid in the $E \gg m$ region, is obtained, one would like to interpolate it with a fixed order cross section, valid instead in the $E \simeq m$ region, so as to retain predictivity over the whole E range.

The differences between the various approaches are then to be found essentially in two points:

- the perturbative order to which the initial condition $f(\mu_0,m)$ is evaluated, and the perturbative accuracy of the evolution;
- the way the matching with the fixed order calculation is performed.

We summarize below the features of the most commonly used implementations.

3.1 ACOT - Aivazis, Collins, Olness, Tung

This approach was the first to try to improve the prediction of the heavy quark structure functions $F_2^c(Q^2,m_c^2)$ and $F_2^b(Q^2,m_b^2)$ at large $Q^2 \gg m_c^2, m_b^2$, by moving potentially large logarithms $\ln(Q^2/m^2)$ into heavy quark parton densities [16, 17]. A general all-order analysis of factorization for the total inclusive $F_2(Q^2)$ in this context was presented in [18].

3.2 Simplified ACOT and ACOT(χ)[1]

The original ACOT prescription [16, 17] has been simplified in [19] along lines suggested in [20, 18]. In a nutshell, diagrams with initial state heavy quark legs can be treated as if they represented massless quarks. More generally, the diagrams can be manipulated by power suppressed terms provided that higher order diagrams are regularized consistently. ACOT(χ) [21, 22] explores this freedom to improve on the threshold behaviour of partonic heavy quark schemes by enforcing the physical pair-production

[1]Contributed by S. Kretzer

threshold on a term-by-term basis. Heuristically, it comes down to a simple re-scaling of Bjorken-x, i.e. in LO

$$F_2^{c\bar{c}} \ \propto \ c(\chi)|_{\chi=x_{\rm Bj}(1+4m^2/Q^2)} \quad . \tag{3}$$

Physical arguments –mostly kinematic– have been given in [21, 23, 22], here we will establish the connection with the FONLL terminology of Section 1.3.3 while focusing on the inclusive DIS process. Much of the following has appeared before, in one form or another, in the literature [16, 17, 18, 24, 25, 26, 27, 28, 19].

We formulate ACOT(χ) as an explicit manipulation of resummed terms of the perturbative series. We follow [24] in notation and add an $\mathcal{O}\left(\alpha_s^1\right)$ fixed order (FO) calculation to an all order collinearly resummed (RS) result. In RS heavy quark mass dependence other than logarithmic is neglected. When we remove double-counting terms from FO + RS the zero mass limit (FOM0) of the FO calculation will be required as an auxiliary quantity. Just as in RS, only asymptotic mass logarithms are retained in FOM0. We write therefore, as usual,

$$\sigma^{ACOT}(Q,m) = \mathrm{FO} + (\mathrm{RS} - \mathrm{FOM0}) \times G \tag{4}$$

where G is an arbitrary operation which behaves like $G = 1 + \mathcal{O}\left(\frac{m^2}{Q^2}\right)$. In [24] G was chosen to be an overall multiplicative factor. More generally, it can be seen as an operation which only modifies, with $\mathcal{O}(m^2/Q^2)$ power-suppressed terms, perturbative coefficients beyond those which have been explicitly calculated, and which are therefore unknown anyway. Any choice for G with this behaviour is therefore legitimate.

To motivate the ACOT(χ) choice for G we first re-write more explicitly the three terms given above in the case of inclusive DIS:

$$\begin{aligned} \mathrm{FO} &= \alpha_s\, g\, \tilde{\otimes} H(Q,m) && (5)\\ \mathrm{FOM0} &= \alpha_s \left(g \otimes P_{qg}^{(0)} \ln\frac{\mu^2}{m^2} + g \otimes C_g \right) && (6)\\ \mathrm{RS} &= c(x) + \alpha_s \left(g \otimes C_g + c \otimes C_q \right) && (7) \end{aligned}$$

where $H(Q,m)$ is the massive coefficient function for the FO gluon fusion process, C_g and C_q are the gluon and light quark coefficient functions (the $\overline{\mathrm{MS}}$ scheme is implied), and g and c are the gluon and charm (i.e. heavy quark) parton distribution functions (both the coefficient functions and the PDFs depend, of course, on the factorization scale $\mu \simeq Q$). $P_{qg}^{(0)}$ is the leading order Altarelli-Parisi splitting vertex. The symbol $\tilde{\otimes} \ \equiv \ \int_\chi^1 d\xi/\xi$ ... denotes a threshold-respecting convolution integral. One can convince oneself that the standard convolution $\otimes$, with $x \ \rightarrow \ \chi$ in the lower limit of integration, only differs by $\tilde{\otimes}$ by power-suppressed terms, $\tilde{\otimes} = \otimes \ + \ \mathcal{O}(m^2/Q^2)$.

The combined result (4) reads now

$$\begin{aligned} \sigma^{ACOT}(Q,m) &= \mathrm{FO} + (\mathrm{RS} - \mathrm{FOM0}) \times G \\ &= \alpha_s\, g\, \tilde{\otimes} H + \left[c(x) - \alpha_s\, g \otimes P_{qg}^{(0)} \ln\frac{\mu^2}{m^2} + \alpha_s\, c \otimes C_q \right] \times G \ , \end{aligned} \tag{8}$$

and we recognize the Krämer-Olness-Soper simplified ACOT framework of [19][2] if we set $G = 1$. Different choices for G can still be made, but natural demands are that:

- In kinematic regions where FO represents the relevant physics (i.e. $Q \sim m$), G should efficiently suppress uncontrolled spurious higher order terms in the square bracket of eq.(8).
- For computational efficiency, the simple $c(x)$ term alone should provide an optimized effective $\mathcal{O}\left(\alpha_s^0\right)$ approximation.

[2] See Eqs. (7), (8) there. General choices for G correspond to the discussion above these equations.

The ACOT(χ) scheme implements these requests by making an implicit choice for G which corresponds to writing

$$\begin{aligned}\sigma^{ACOT(\chi)}(Q,m) &= \mathrm{FO} + (\mathrm{RS} - \mathrm{FOM0}) \times G \\ &= \alpha_s\, g\, \tilde{\otimes} H + \left[c(\chi) - \alpha_s\, g\, \tilde{\otimes} P_{qg}^{(0)} \ln \frac{\mu^2}{m^2} + \alpha_s\, c\, \tilde{\otimes} C_q\right] . \end{aligned} \tag{9}$$

Further details on ACOT(χ) can be found in [21, 22, 23]. These articles also contain a more intuitive perspective of ACOT(χ). Moreover, [22] describes a PDF set that is consistent with ACOT(χ) applications.

3.3 BSMN - Buza, Smith, Matiounine, van Neerven

In Refs. [29, 30, 31, 32, 33] the treatment of heavy quarks as a parton was fully explored through next-to-next-leading order (NNLO), based on a precise two-loop analysis of the heavy quark structure functions from an operator point of view. This analysis yielded a number of results. One result is important beyond the observable at hand: the authors obtained the complete set of NNLO matching conditions for parton evolution across flavor thresholds. They found that, unlike at NLO, the matching conditions are *discontinuous* at the flavor thresholds. These conditions are necessary for any NNLO calculation at the LHC, and have already been implemented in a number of evolution packages [34, 35].

Furthermore, their two-loop calculations explicitly showed that the heavy *quark* structure functions in such a variable flavor approach are not infrared safe: one needs to either define a heavy *quark-jet* structure function, or introduce a fragmentation function to absorb the uncancelled divergence. In either case, a set of contributions to the inclusive light parton structure functions must be included at NNLO.

A dedicated analysis [36] for charm electroproduction showed that even at very large Q^2 one could not distinguish the fixed order NLO calculation of [10] and the NNLO VFNS calculations of [31], given the experimental data available in the year 2000. This demonstrates the possibility that the large logarithms $\ln(Q^2/m^2)$ together with small coefficients can in the end have little weight in the overall hadronic cross section.

3.4 FONLL - Fixed Order plus Next-to-Leading Log resummation

This approach was developed for improving the large-p_T differential cross section for heavy quark production in hadron-hadron collisions [37]. It was successively extended to photoproduction [38], and in a second phase a matching to the fixed order NLO calculations was performed [24, 39]. The FONLL acronym refers specifically to the matched version.

From the point of view of perturbative logarithms, it contains a NLO-accurate initial condition and full NLL evolution. It therefore reproduces the full NLL structure of the NLO calculation, and resums to all orders the large logarithms with NLL accuracy.

The matching with the fixed order result is performed according to the following master formula (see eq.(16) of [24]):

$$\sigma_Q^{\mathrm{FONLL}}(p_T, m) = \mathrm{FO} + (\mathrm{RS} - \mathrm{FOM0})G(m, p_T) \ , \tag{10}$$

where FO stands for the NLO fixed order massive calculation, FOM0 for its $m/p_T \to 0$ limit (where however $\ln p_T/m$ terms and non-vanishing terms are kept), and RS for the massless, resummed calculation[3]. The RS$-$FOM0 subtraction is meant to cancel the terms which are present in both RS and FO. This difference starts therefore at order α_s^2 with respect to the Born cross section: at large p_T it resums

[3]This term might also be referred to as a ‘zero-mass variable flavour number scheme’ (ZM-VFNS) contribution. However this name, while by itself completely general, has been used in the past for specific approaches with different overall perturbative accuracies. We shall therefore avoid its use. It will be understood that ‘RS’ in this approach has full NLL accuracy.

correctly the NLL terms, at small p_T it only contains spurious terms, which are suppressed by the function $G(m, p_T) = p_T^2/(p_T^2 + c^2 m^2)$, with $c = 5$, in order to ensure a physically correct behaviour. The choice of the suppression factor was motivated in [24] by the observation that the massless limit starts to approach the massive hadroproduction calculation at $\mathcal{O}(\alpha_s^3)$ only for $p_T > 5m$. Below this value the massless limit returns unreliable results, and its contribution must therefore be suppressed. It is important to realize that $G(m, p_T)$ only affects terms which are beyond the control of perturbation theory, and therefore it does not spoil the NLO+NLL accuracy. The choice to control such terms by means of an ad-hoc function might seem a somewhat unpleasant characteristic of this approach. However, it simply portraits the freedom one has in performing the matching, and does not represent a shortcoming of the approach: different matching procedures will simply make other implicit or explicit choices for $G(m, p_T)$.

For the sake of making comparisons with other approaches easier, the formula (10) can be rewritten with some more details as follows:

$$\sigma_Q^{\mathrm{FONLL}}(p_T, m) = \sum_{ij\in\mathcal{L}} F_i F_j\ \sigma_{ij\to QX}(p_T, m)$$
$$+ \left(\sum_{ijk\in\mathcal{L}+\mathcal{H}} F_i F_j\ \hat{\sigma}^{\overline{MS}}_{ij\to kX}(p_T) D_{k\to Q} - \sum_{ij\in\mathcal{L}} F_i F_j\ \sigma_{ij\to QX}(p_T, m; m\to 0) \right) G(m, p_T) \quad (11)$$

A number of ingredients needing definition have been introduced. The kernel cross sections $\sigma_{ij\to QX}(p_T, m)$ are the massive NLO calculations for heavy quark production of Refs. [2, 3, 4, 5, 6]. When convoluted with the PDFs for light flavours F_i ($i \in \mathcal{L}$) they yield the FO term in eq. (10). The $\sigma_{ij\to QX}(p_T, m; m\to 0)$ terms represent the $m \to 0$ limit of the massive NLO cross sections, performed by sending to zero m/p_T terms while preserving $\ln(p_T/m)$ contributions and non-vanishing constants. When convoluted with light flavour PDFs they give FOM0. Finally, $\hat{\sigma}^{\overline{MS}}_{ij\to kX}(p_T)$ are the massless $\overline{MS}$-subtracted NLO cross section kernels given in [40]. In addition to the light flavour PDFs, they are also convoluted with the perturbatively-calculated parton distribution functions for the heavy quarks ($i \in \mathcal{H}$) and with the fragmentation functions describing the transformation of a parton into a heavy quark, $D_{k\to Q}$ [15], to give the term RS.

The formula given above returns the differential cross section for heavy *quark* production, evaluated with NLO + NLL accuracy. In order to obtain the corresponding cross section for an observable heavy meson it must still be convoluted with the proper scale-independent non-perturbative fragmentation function, extracted from experimental data, describing the heavy quark → heavy hadron transition:

$$\sigma_H^{\mathrm{FONLL}}(p_T, m) = \sigma_Q^{\mathrm{FONLL}}(p_T, m) D^{NP}_{Q\to H}\ . \quad (12)$$

Phenomenological analyses of charm- and bottom-flavoured hadrons production within the FONLL approach have been given in [41, 42, 43, 44, 45].

3.5 GM-VFNS - General mass variable flavour number scheme

This approach also combines a massless resummed calculation with a massive fixed order one, for predicting p_T distributions in hadron-hadron collisions. One difference with respect to FONLL is that this approach does not include the perturbative NLO parton-to-heavy-quark fragmentation functions $D_{k\to Q}$. Rather, it directly convolutes a properly $\overline{MS}$ subtracted cross section (with mass terms also included, hence the 'general mass' name) with non-perturbative fragmentation functions for heavy mesons $D^{NP,\overline{MS}}_{Q\to H}$, fitted at LEP in a pure $\overline{MS}$ scheme. The cross section can be schematically written as

$$\sigma_H^{\mathrm{GM-VFNS}}(p_T, m) = \sum_{ij\in\mathcal{L}} F_i F_j\ \hat{\sigma}_{ij\to QX}(p_T, m) D^{NP,\overline{MS}}_{Q\to H} + \sum_{ijk\in\mathcal{L}+\mathcal{H}} F_i F_j\ \hat{\sigma}^{\overline{MS}}_{ij\to kX}(p_T) D^{NP,\overline{MS}}_{k\to H}\ , \quad (13)$$

where the 'massive-but-subtracted' cross section kernels $\hat{\sigma}_{ij\to QX}(p_T,m)$ are defined by

$$\hat{\sigma}_{ij\to QX}(p_T,m) \equiv \sigma_{ij\to Q}(p_T,m) - \sigma_{ij\to QX}(p_T,m;m\to 0) + \hat{\sigma}^{\overline{MS}}_{ij\to QX}(p_T)\ . \tag{14}$$

The new kernels $\hat{\sigma}_{ij\to QX}(p_T,m)$ defined by this operation (of the form FO-FOM0+RS) can be convoluted with an evolved $\overline{MS}$-subtracted fragmentation function, but they also retain power suppressed m/p_T terms. It should also be noted that the sum in the second term of (13) only runs over contributions not already included in the first.

Recalling the way the perturbative parton-to-heavy-quark $D_{k\to Q}$ fragmentation functions are defined in [15], setting

$$D^{NP,\overline{MS}}_{k\to H} = D_{k\to Q} D^{NP}_{Q\to H}, \qquad k \in \mathcal{L}+\mathcal{H}\ , \tag{15}$$

and comparing eqs.(13) and (11), it can be seen that the GM-VFNS master formula is a reshuffling of the FONLL one, up to higher-orders terms.

Two comments are worth making. The first is that due to the absence of the perturbative $D_{k\to Q}$ terms, eq. (13) cannot reproduce the NLO heavy *quark* production cross section: even the normalization must be extracted from the experimental data. Eq. (11), on the other hand, can reproduce the heavy quark spectrum, and only the heavy quark → heavy meson transition is fitted to data. The second remark concerns the higher order power suppressed terms: since GM-VNFS implicitly makes a different choice for the $G(m,p_T)$ function, the results from the two approaches might differ considerably in the $p_T \sim m$ region since, while formally suppressed, such terms can be numerically important.

An example of a phenomenological application of the GM-VFNS scheme is given below.

3.6 Hadroproduction of heavy mesons in a massive VFNS[4]

Various approaches for next-to-leading-order (NLO) calculations in perturbative QCD have been applied to one-particle-inclusive hadroproduction of heavy mesons. The general-mass variable-flavor-number scheme (GM-VFNS) devised by us in Ref. [46, 47] is closely related to the conventional massless variable-flavor-number scheme (ZM-VFNS), but keeps all m^2/p_T^2 terms in the hard-scattering cross sections, where m is the mass of the heavy quark and p_T the transverse momentum of the observed meson, in order to achieve better accuracy in the intermediate region $p_T \geq m$. The massive hard-scattering cross sections have been constructed in such a way that the conventional hard-scattering cross sections in the $\overline{\mathrm{MS}}$ scheme are recovered in the limit $p_T \to \infty$ (or $m \to 0$). The requirement to adjust the massive theory to the ZM-VFNS with $\overline{\mathrm{MS}}$ subtraction is necessary, since all commonly used PDFs and FFs for heavy flavors are defined in this particular scheme. In this sense, this subtraction scheme is a consistent extension of the conventional ZM-VFNS for including charm-quark mass effects. It should be noted that our implementation of a GM-VFNS is similar to the ACOT scheme [16, 17], which has been extended to one-particle-inclusive production of B mesons a few years ago [48]. As explained in the second paper of Ref. [46, 47], there are small differences concerning the collinear subtraction terms. Furthermore, in Ref. [48], the resummation of the final-state collinear logarithms has been performed only to leading logarithmic accuracy. The field-theoretical foundation of a GM-VFNS has been provided a few years ago by a factorization proof including heavy-quark masses [18]. Therefore, it is possible to extract improved universal parton distribution functions (PDFs) [22] and fragmentation functions (FFs) [49] from fits employing massive hard-scattering cross sections. From this perspective, it is important to compute massive hard-scattering cross sections in a given massive scheme for all relevant processes. Explicit calculations in the original ACOT scheme have been performed in Ref. [50, 51] for inclusive and semi-inclusive deep-inelastic scattering (DIS). Furthermore, our calculation in Ref. [46, 47] for hadronic collisions completes earlier work in the GM-VFNS on D-meson production in $\gamma\gamma$ and γp collisions [52, 53, 54], and it is planned to extend our analysis to the case of heavy-meson production in DIS.

[4]Contributed by B.A. Kniehl and I. Schienbein.

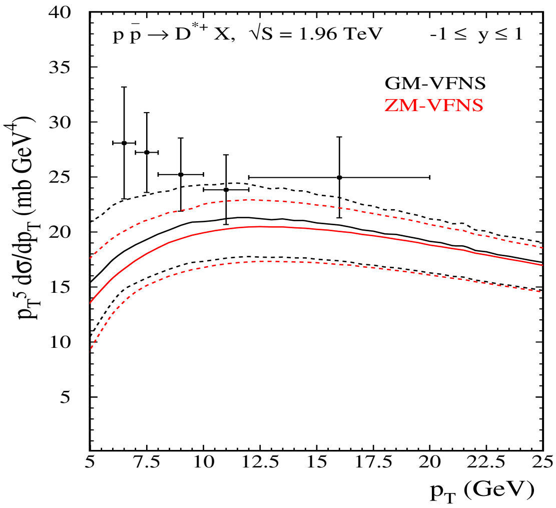


Fig. 1: QCD predictions for one-particle-inclusive production of $D^\star$ mesons at the Tevatron Run II in comparison with CDF data [57]. The results are shown for the average $D^\star = (D^{\star +} + D^{\star -})/2$. The solid lines have been obtained with $\mu_\mathrm{R} = \mu_\mathrm{F} = \mu'_\mathrm{F} = m_T$. The upper and lower dashed curves represent the maximum and minimum cross sections found by varying μ_R, μ_F, and μ'_F independently within a factor of 2 up and down relative to the central values while keeping their ratios in the range $0.5 \leq \mu_\mathrm{F}/\mu_\mathrm{R}, \mu'_\mathrm{F}/\mu_\mathrm{R}, \mu_\mathrm{F}/\mu'_\mathrm{F} \leq 2$.

Next, we show predictions for the cross section $d\sigma/dp_T$ of $D^\star$-meson production obtained in the GM-VFNS and the ZM-VFNS. The cross section has been scaled with p_T^5 in order to arrive at a flat p_T distribution, which is useful for visualizing the heavy-quark mass effects. The hard-scattering cross sections are convoluted with the (anti-)proton PDFs and FFs for the transition of the final-state parton into the observed $D^\star$ meson. We use the CTEQ6M PDFs [55] and the FFs for $D^\star$ mesons from Ref. [56]. As in the experimental analysis, the theoretical results are presented for the average $(D^{\star +} + D^{\star -})/2$. We consider $d\sigma/dp_T$ at $\sqrt{S} = 1.96$ TeV as a function of p_T with y integrated over the range $-1.0 < y < 1.0$. We take the charm mass to be $m = 1.5$ GeV and evaluate $\alpha_s^{(n_f)}(\mu_R)$ with $n_f = 4$ and scale parameter $\Lambda^{(4)}_{\overline{\mathrm{MS}}} = 328$ MeV, corresponding to $\alpha_s^{(5)}(m_Z) = 0.1181$. The results are presented in Fig. 1 for the GM-VFNS (black lines) and the ZM-VFNS (red lines) in comparison with CDF data [57]. The solid lines have been obtained with $\mu_\mathrm{R} = \mu_\mathrm{F} = \mu'_\mathrm{F} = m_T$. The upper and lower dashed curves represent the maximum and minimum cross sections found by varying μ_R, μ_F, and μ'_F independently within a factor of 2 up and down relative to the central values requiring for their ratios to satisfy the inequalities $0.5 \leq \mu_\mathrm{F}/\mu_\mathrm{R}, \mu'_\mathrm{F}/\mu_\mathrm{R}, \mu_\mathrm{F}/\mu'_\mathrm{F} \leq 2$. As can be seen, for large values of p_T, the predictions of the GM-VFNS nicely converge to the corresponding results in the ZM-VFNS. Both approaches lead to reasonable descriptions of the data, but the inclusion of the positive mass effects clearly improves the agreement with the data. It should be noted that the mass effects are largest for the upper curves of the uncertainty band, which have been obtained with the smaller value of the renormalization scale implying a larger $\alpha_s(\mu_R)$. At $p_T = 5$ GeV, one observes an increase of the massless cross section by about 35%. A more detailed comparison of the GM-VFNS with CDF data [57] including D^0, D^+, and D_s^+ mesons can be found in Refs. [58, 59].

Residual sources of theoretical uncertainty include the variations of the charm mass and the employed PDF and FF sets. A variation of the value of the charm mass does not contribute much to the theoretical uncertainty. Also, the use of other up-to-date NLO proton PDF sets produces only minor differences. Concerning the choice of the NLO FF sets, we obtain results reduced by a factor of 1.2–1.3 when we use the NLO sets from Ref. [60], which is mainly caused by a considerably different gluon FF. A more detailed discussion can be found in Ref. [56].

Name	LO Process	Interest	Accuracy
single-top t-channel	$qb \to qt$	top EW couplings	NLO
single-top tW-associated	$gb \to tW^-$	Higgs bckg, top EW couplings	NLO
Vector boson + 1 b-jet	$gb \to (\gamma, Z)b$	b-pdf, SUSY Higgs benchmark	NLO
Vector boson + 1 b-jet +1 jet	$qb \to (\gamma, Z, W)bq$	single-top and Higgs bckgs	NLO
Higgs inclusive	$b\bar{b} \to (h, H, A)$	SUSY Higgs discovery at large $\tan\beta$	NNLO
Higgs + 1 b-jet	$gb \to (h, H, A)b$	SUSY Higgs discovery at large $\tan\beta$	NLO
Charged Higgs	$gb \to tH^-$	SUSY Higgs discovery	NLO

Table 1: Process relevant for SM measurements and SUSY discoveries at the LHC which entail the use of bottom in the initial state. All of them are known at least at NLO accuracy.

4 A case study in collinear resummation: b-quark PDF from $Z + b$ production at LHC[5]

4.1 Introduction

The discovery of new physics at LHC will probably rely on the detailed understanding of standard-model background processes. Outstanding among these is the production of weak bosons (W, Z) in association with jets, one or more of which contains a heavy quark ($Q = c, b$). The prime example is the discovery of the top quark at the Fermilab Tevatron, which required a thorough understanding of the W+jets background, with one or more heavy-quark jets. The discovery of single-top-quark production via the weak interaction will require an even more sophisticated understanding of this background [61, 62].

For many processes involving production of heavy quarks, there are two ways (schemes) to perform the calculation in QCD: the fixed-flavor-scheme (FFS) and variable-flavor-scheme (VFS). The main practical difference between the two approaches is simple: in the VFS the heavy-quark can also be in the initial state, and in that case is assumed to be massless, while in the FFS it appears only as a final state (massive) particle. QCD factorisation tells us that if calculations could be performed at arbitrary high order, the two schemes would be equivalent. At fixed order, on the other hand, differences arise and one should choose that describing more effectively the kinematics of the process of interest. This freedom has sometimes created intense and fruitful debates among the QCD practitioners (see, *e.g.*, Ref. [63] for a detailed comparison of Higgs boson production in association with bottom quarks). Here we just recall the main two reasons for using a heavy-quark distribution function. First, it resums collinear logarithms of the form $\ln Q/m_Q$ to all orders, where Q is the scale of the hard scattering and m_Q is the mass of the heavy quark. Second, it simplifies the leading-order process, which often makes a higher-order calculation feasible. There are many processes in the standard model and in models beyond it, such as SUSY, that are better described using a bottom in the initial state. In Table 1, we give a non-exhaustive list of processes that will be relevant for QCD, EW and SUSY studies at the LHC, and the QCD order at which they are known.

At present the b distribution function is derived perturbatively from the gluon distribution function [17, 18, 55, 34]. Recently, direct, albeit not very precise, measurements of F_2^b have become available that are compatible with the perturbative determination [64, 65]. In the light of its phenomenological importance, a better direct determination of the b distribution function is certainly desirable.

To this aim it has been proposed to use the associated production of a photon and a b-jet via

[5] Contributed by S. Diglio, F. Maltoni, F. Petrucci, A. Tonazzo and M. Verducci

$gb \to \gamma b$ at the LHC [66]. This measurement suffers from two main limitations. The first is the large contamination from charm which has a much larger cross section due to both the pdf and the electromagnetic coupling. The second is that the theoretical prediction at NLO for an isolated photon is uncertain, due to necessity of introducing a photon fragmentation function, which is at present poorly known.

In this note we follow the suggestion of Ref. [67] to use Z production in association with a b-jet to extract information on the b-pdf. At leading order, it proceeds via $gb \to Zb$, as shown in Fig. 2. This process is known at NLO, including γ/Z interference effects. The advantages of using a γ/Z decaying into leptons with respect to a real photon are noticeable. The NLO cross section is theoretically very well known and, apart from the PDF's, free of non-perturbative inputs. In addition, the competing process $gc \to Zc$ is suppressed by the ratio of the couplings of the charm and the bottom to the Z, and makes the b-pdf determination much cleaner.

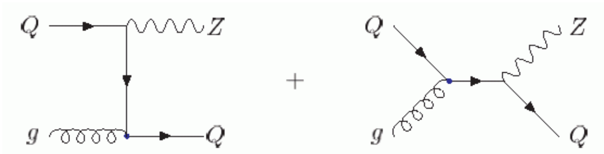


Fig. 2: Leading Order Feynman diagrams for associated production of a Z boson and a single high-p_T heavy quark ($Q = c, b$).

The D0 Experiment at Tevatron has recently measured the cross-section ratio $\sigma(Z + b)/\sigma(Z + jet)$ [68], and their result is consistent with the NLO calculation.

As pointed out in [67], the measurement of this process at the LHC should be even more interesting because the contribution of the leading order process, sensitive to the b content of the proton, is more relevant than at the Tevatron. In addition, the total cross-section is larger by a factor 50, and the relative contribution of background processes, mainly $Z+c$, is smaller. These features are summarised in table 2, taken from Ref. [67].

Cross sections (pb)	Tevatron	LHC
Process	ZQ inclusive	
$gb \to Zb$	$13.4 \pm 0.9 \pm 0.8 \pm 0.8$	$1040^{+70+70\ +30}_{-60-100-50}$
$gb \to Zb\bar{b}$	6.83	49.2
$gc \to Zc$	$20.3^{+1.8}_{-1.5} \pm 0.1^{+1.3}_{-1.2}$	$1390 \pm 100^{+60+40}_{-70-80}$
$gc \to Zc\bar{c}$	13.8	89.7
	Zj inclusive	
$q\bar{q} \to Zg, gq \to Zq$	$1010^{+44+9+7}_{-40-2-12}$	$15870^{+900+60\ +300}_{-600-300-500}$

Table 2: Next-to-leading-order inclusive cross sections (pb) for Z-boson production in association with heavy-quark jets at the Tevatron ($\sqrt{s}$ = 1.96 TeV $p\bar{p}$) and the LHC ($\sqrt{s}$ = 14 TeV pp). A jet lies in the range $p_T > 15$ GeV/c and $|\eta| < 2$ (Tevatron) or $|\eta| < 2.5$ (LHC). ZQ indicates events containing a heavy quark, Zj events which do not contain a heavy quark.

Besides the possibility of extracting the b-pdf, $Z+b$ represents also a benchmark and in some cases a background to the search of the Higgs boson, when it is produced in association with a single high-p_T b quark [63]: the dominant leading-order subprocess for the production of a Higgs boson via its coupling to the b is $b\bar{b} \to h$; however, if the presence of a single b with high p_T is demanded, the dominant process becomes $gb \to hb$, with cross-sections of the order of tens of fb. The h can then decay to the same final states as the Z; in particular, the decay $h \to \mu^+\mu^-$ is enhanced in some models [69, 70, 71].

A preliminary analysis on the potential of the ATLAS experiment to measure the Z+b-jet production at the LHC is presented in the following.

4.2 A study of LHC measurement potential

A sample of Z+jet events generated using the PYTHIA Monte Carlo [72] was processed with a fast simulation of the ATLAS detector, the ATLFAST package [73]. Only decays of the Z boson to $\mu^+\mu^-$ were taken into account. The signal was defined as the sample events containing a b quark with $p_T > 15$ GeV/c and $|\eta| < 2.5$. The background samples containing respectively a c quark within the same cuts, or a jet originating from a light quark or a gluon in the same range, were considered separately. The NLO cross-sections computed in [67] were used for the signal and for these two classes of background, while the cross-section given by PYTHIA was taken for the other types of events.

The experimental selection of Z+jet events with $Z \to \mu^+\mu^-$ required the detection of two muons of opposite charge with $p_T > 20$ GeV/c and $|\eta| < 2.5$ and one hadronic jet. The presence of two high-p_T muons ensures the possibility to have high trigger efficiency on this type of events. In addition, to reject the contribution from virtual photons, the invariant mass $M_{\mu\mu}$ of the muon pair was required to be close to the Z mass (80 GeV/c$^2 < M_{\mu\mu} <$ 105 GeV/c^2). About 50% of signal events are retained after applying these cuts, the loss being equally due to the η acceptance and to the p_T cut.

The selection of events where the jet originates from a b quark was based on two different tagging methods, as described in the following. Their complementarity is still to be studied in detail, however the comparison of two independent selections will be important to control the systematic uncertainties.

The first method to select $Z + b$ events was based solely on the presence of a third muon. Hadrons containing a b quark give origin to prompt muon decays in about 12% of the cases. The efficiency of this method, therefore, cannot exceed this value, however the background is also expected to be small. The "third muon", considered to be the muon from the b hadron decay, will in general be softer and closer to a jet than the muons from the Z decay. The distribution of the transverse momentum of the third muon in $Z + j$ events is shown in Fig. 3. Different thresholds on the third muon p_T were considered for the final selection.

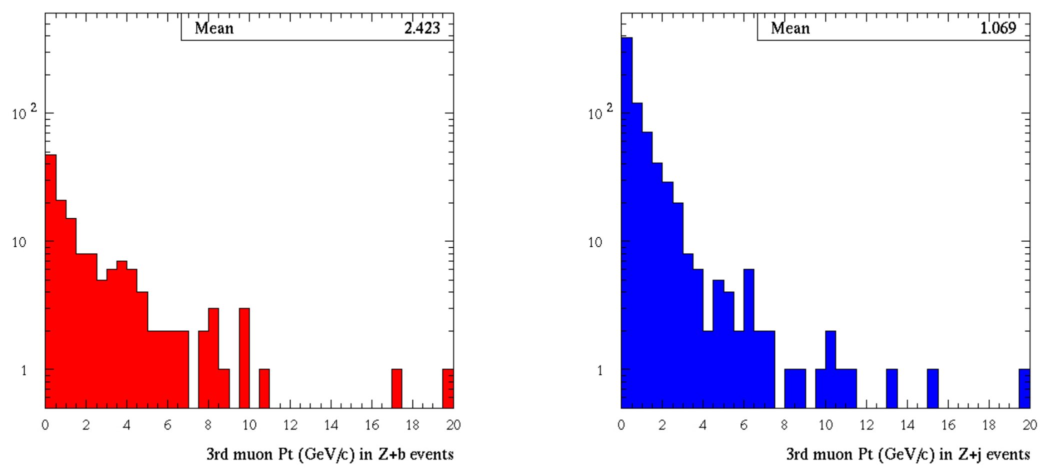


Fig. 3: Distribution of the transverse momentum of the third muon in a Z+jet sample, for signal events (left) and for events with no b quark (right).

The second analysis used an inclusive method for b-tagging, based on the presence of secondary vertices and of tracks with high impact parameter with respect to the primary vertex, originated from the decay of the long-lived b hadrons. The ATLFAST package reproduces the ATLAS b-tagging capabilities by applying the tagging efficiency on b jets and a mis-tag rate on non-b jets on a statistical basis, according to the values set by the user to reproduce the actual detector performance. The efficiency of the inclusive b-tagging on signal events, after the selection described above, is about 30%. The mistagging probability is about 4% on c-quark jets and 0.5% on light jets.

The overall efficiency on signal events, the expected number of signal and background events with an integrated luminosity of 10 fb^{-1} and the expected purity of the selected samples are reported in table 3. With the fast simulation, the soft muon tagging capabilities are optimistic, in that full efficiency and no mis-tag are assumed for the lepton identification; more realistic assumptions will be made when the study is carried on with the full detector simulation. The efficiency on signal events achieved with the inclusive b-tagging method, where the results of the fast simulation are more realistic, is higher than with the soft muon tagging, while the purity of the selected sample is still quite good. Consistent results were obtained with a full simulation of the ATLAS experiment, on a small statistics sample.

A better determination of the signal component in the selected sample will eventually be achieved by exploiting the information on the transverse momentum of the b-jet or of the third muon.

Cut	Efficiency	$N_b^{p_T>15\text{ GeV},\ \lvert\eta\rvert\leq 2.5}$	$N_c^{p_T>15\text{ GeV},\ \lvert\eta\rvert\leq 2.5}$	N_{other}	Purity
$p_T^\mu > 4$ GeV/c	4%	13990	6270	0	69%
$p_T^\mu > 5$ GeV/c	3%	11090	5210	0	69%
$p_T^\mu > 6$ GeV/c	2.5%	8430	4180	0	67%
incl. b-tag	14%	49500	17400	49600	43%

Table 3: Expected efficiency, statistics and purity in a data sample corresponding to an integrated luminosity of 10 fb^{-1}, using the soft muon tagging with different thresholds on the muon transverse momentum and the inclusive b-tagging. N_b denotes the number of expected signal events as defined in the text, N_c the number of selected events with a c jet with $p_T > 15$ GeV and $|\eta| < 2.5$, N_{other} the selected events from other processes. The statistical error on efficiencies and purities, due to the limited size of the simulated sample, is at the level of 1-2%.

Given the large statistics of the available data samples, the measurement will be limited by systematic effects.

The possibility to control the systematic effects directly from data samples has been explored, in particular the evaluation of b-tagging performance and of the residual background.

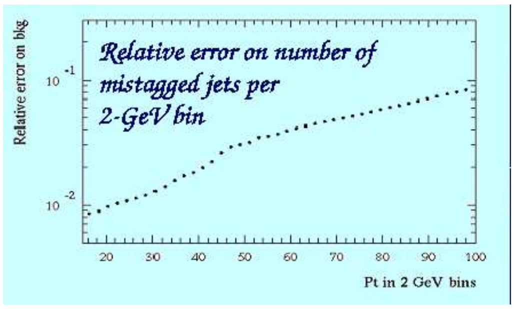
The b-tagging efficiency can be checked using b-enriched samples. Based on previous experience at Tevatron and LEP, we can expect a relative uncertainty of about 5%.

The background in the selected sample is mainly due to mis-tagged jets from c and light quarks. This can be controlled by looking at the number of b-tagged jets in data samples that in principle should contain no b-jets at first order. W+jet events, for example, will be available with large statistics and with jets covering the full p_T range of the signal. It can therefore be expected to estimate the background from mis-tagging with a relative uncertainty at the level of few percent, as shown by the plots in figure 4.

4.3 Conclusions and outlook

Z boson production in association with a b-jet can provide information on the b-pdf.

A preliminary study of the $Z + b$ channel using a fast simulation of the ATLAS detector has shown that this type of event will be observed with very high statistics and good purity at the LHC. Given the large statistics of the samples, the precision of the $Z + b$ cross-section measurement will be limited by



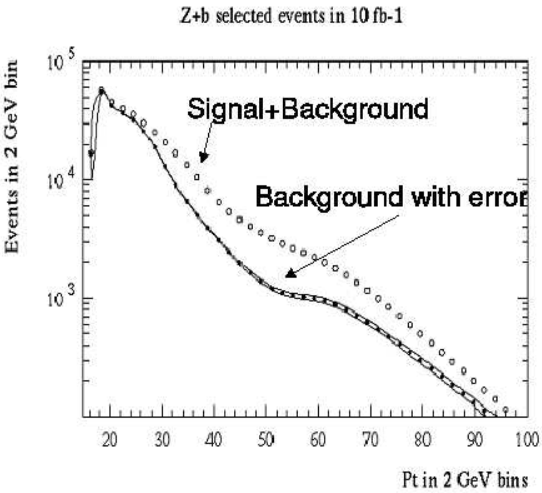


Fig. 4: Systematics due to mis-tagging of b-jets as evaluated from W+jet events. Left: relative error on background level per jet p_T bin. Right: p_T distribution of jets in event selected as $Z + b$; the error band on the background contribution represents the systematic uncertainty, as derived from the previous plot.

systematic effects. Some possibilities to evaluate systematic uncertainties directly form the data have been considered. An overall accuracy on the measurement at the level of 5% can be expected.

The availability of large samples opens interesting possibilities for the study of differential distributions: for instance, measuring the cross-section as a function of the η and p_T of the Z boson would allow for the measurement of the b PDF as a function of the momentum fraction carried by the quark inside the proton. These items are an additional topic for further studies.

5 Soft-gluon resummation[6]

QCD factorizes long- and short distance dynamics in inclusive cross sections with initial state hadrons into non-perturbative, but universal parton distribution functions, and perturbatively calculable hard scattering functions. Large remnants of the long-distance dynamics occur near the threshold edge of phase space in the form of logarithmic distributions that are singular at the edge. Resummation [74, 75] of these effects organizes them to all orders in perturbation theory, and thereby extends the predictive power of QCD.

Threshold resummation is now a well-established calculational scheme with systematically improvable accuracy. It allows organization of all subleading powers of the logarithmic enhancements, and can be consistently matched to finite order perturbation theory. Resummed expressions, which take the form of exponentiated integrals over functions of the running coupling, require however a prescription for their numerical evaluation to handle a Landau pole singularity of the coupling. But for this intrinsic ambiguity (which must cancel against ambiguities in power corrections), threshold resummation is just as systematically improvable as the standard coupling constant expansion.

As stated earlier, the more differential a cross section, the better suited it is for phenomenology, because one may incorporate detector-specific acceptance cuts and thereby reduce the need for extrapolation. Therefore we should like to better understand the behavior of threshold-resummed expressions for double-differential cross sections. A study for the inclusive threshold-resummed heavy quark structure function can be found in Ref. [76]. Here we examine the differential structure function for the reaction

$$\gamma^*(q) + \mathrm{P}(p) \to \mathrm{Q}(p_1) + X'(p_2') \tag{16}$$

[6]Contributed by T.O. Eynck and E. Laenen.

which we write as

$$\frac{d^2 F_2^Q(S,T_1,U_1)}{dT_1\, dU_1} \tag{17}$$

We define the invariants

$$\begin{aligned} S &= (p+q)^2 \equiv S' - Q^2, \qquad T_1 = (p-p_1)^2 - m^2, \\ U_1 &= (q-p_1)^2 - m^2, \qquad S_4 = S' + T_1 + U_1 \,. \end{aligned} \tag{18}$$

The invariant mass squared of the final state X' is given by

$$M_{X'}^2 = m^2 + S_4 \tag{19}$$

so that the elastic (threshold) limit for the subprocess (16) is approached by $S_4 \to 0$. It may be converted to the double-differential structure function in terms of the heavy quark transverse momentum and rapidity, e.g.

$$\frac{d^2 F_k^Q}{d(p_T^Q)^2\, dy^Q} = S' \frac{d^2 F_k^Q}{dT_1\, dU_1}\,, \tag{20}$$

where e.g. [11]

$$p_T^Q = \left[\frac{S'T_1U_1 + Q^2T_1^2 + Q^2S'T_1}{S'^2} - m^2\right]^{(1/2)} . \tag{21}$$

At the parton level one may define invariants equivalent to those in (18), which we will denote by using lower case. The order-by-order perturbation theory expansion for the partonic version of this distribution $\omega(s_4, t_1, u_1)$ and its all-order resummation have the following schematic forms

$$\begin{aligned} \omega &= 1 + \alpha_s(L^2 + L + 1) + \alpha_s^2(L^4 + L^3 + L^2 + L + 1) + \ldots \\ &= \exp\left(\underbrace{\underbrace{Lg_1(\alpha_s L)}_{LL} + g_2(\alpha_s L)}_{NLL} + \ldots\right) \underbrace{C(\alpha_s)}_{\text{constants}} \\ &\quad + \text{ suppressed terms} \end{aligned} \tag{22}$$

with

$$g_1(\lambda) = \frac{C_F}{\pi b_0 \lambda}\Big[\lambda + (1-\lambda)\ln(1-\lambda)\Big], \quad \lambda = b_0 \alpha_s \ln N\,. \tag{23}$$

(We have also computed $g_2(\lambda)$; by including ever more g_i functions in the exponent in Eq. (22) we can increase the parametric accuracy of the resummation.) The symbol L^i represents, in this case, the logarithmically singular plus-distributions

$$\left[\frac{\ln^{i-1}(\rho)}{\rho}\right]_+ \tag{24}$$

with $\rho = s_4/m^2$, or, after a Laplace transform $\int d\rho \exp(-N\rho)$ by $\ln^i N$. The conversion to momentum space then requires a numerical inverse Laplace transform. For the case at hand one needs to compute

$$S'^2 \frac{d^2 F_2^Q(S_4,T_1,U_1)}{dT_1\, dU_1} = \int\limits_{c-i\infty}^{c+i\infty} \frac{dN}{2\pi i} e^{NS_4/m^2} \bar{\phi}_g\left(N\frac{S'+T_1}{m^2}\right) \omega\Big(N, T_1, U_1\Big)\,, \tag{25}$$

with c the intercept of the contour with the real N axis, and $\phi_g(N)$ the gluon density in moment space. We chose a toy density for the gluon PDF, and the minimal prescription [77] to perform the N integral.

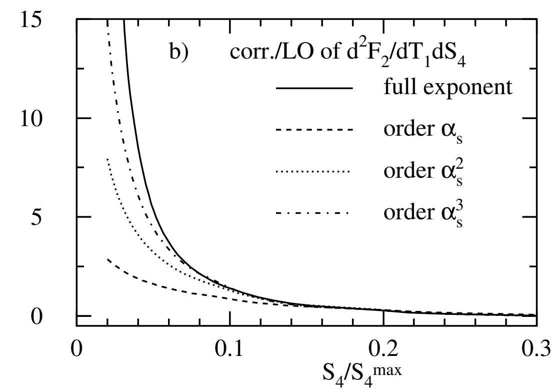


Fig. 5: Expandability of the resummed expressions for $d^2F_2^c/dT_1dS_4$ with NLL exponent (ratio to LO)

In Fig. 5 we evaluate this expansion as a function of the recoil mass S_4, and compare it to its finite order expansions. We keep the variable T_1 fixed at the average of its minimum and maximum allowed value. Clearly, for reasonable values of S_4 the resummed result is already well-approximated by its 2nd and 3rd order expansions.

Another way to evaluate the resummed expression is in terms of towers of [78] $L = \ln N$.

$$\omega = h_{00}(\alpha_s)\left[1 + \sum_{k=1}^{\infty}\left(\frac{\alpha_s}{\pi}\right)^k \left(c_{k1}\,L^{2k} + c_{k2}\,L^{2k-1} + c_{k3}\,L^{2k-2} + \ldots\right)\right]. \tag{26}$$

where the indicated coefficients c_{kj} can be determined exactly. More accuracy here means including more subleading towers. This method is equivalent, but not identical to the minimal prescription method. In practice, one need only include the first 4 terms in each tower, the higher terms are vanishingly small. The ambiguities mentioned earlier are shifted to far-subleading towers in this approach. To exhibit the convergence of terms in the towers, it will be useful and illustrative to exhibit contributions of a particular order in the strong coupling and the large logarithms. We will employ the notation

$$\mathrm{N}^k\mathrm{LO} - l\mathrm{L}_N \tag{27}$$

for finite order results, where k indicates the order in the strong coupling, the subscript N denotes moments, and l expresses if only the leading term ($l = 1$, L^2k_N), or also the next-to-leading term ($l = 2$, L_N^{2k-1}) is included, etc. In Fig. 6 we see also in this approach a rapid convergence toward the

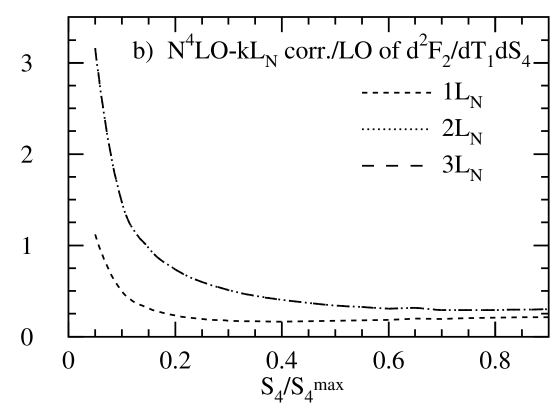


Fig. 6: Tower resummation at $\mathrm{N}^4\mathrm{LO} - k\mathrm{L}_N$, $k \in \{1, 2, 3\}$ ($\mathrm{N}^4\mathrm{LO} - 2\mathrm{L}_N$ and $\mathrm{N}^4\mathrm{LO} - 3\mathrm{L}_N$ almost coincide).

tower-resummed result.

A more complete study of the relevance of threshold resummation for electroproduction of heavy quarks at HERA still awaits. We note that even if the size of the corrections does not cause much concern for the perturbative analysis of an observable, threshold resummation or its finite order approximations, often lead to a reduction of scale dependence [79], indeed also seen in Ref. [76].

6 k_t - factorization[7]

6.1 Introduction

The transverse momenta of the partons initiating a hard scattering process, like heavy quark production via $\gamma g \to Q\bar{Q}$ or $gg \to Q\bar{Q}$ in lepto- (hadro-) production, respectively, is mainly generated by the QCD evolution, which can reach large values, in DGLAP up to the factorization scale, in BFKL/CCFM/LDC even larger.

The typical transverse momenta of the gluons in a process $gg \to X$ for different masses M of the system X are shown in Fig. 7 as a function of the momentum fraction x of one of the gluons for LHC energies. The transverse momenta can become large, so that they cannot be neglected. A

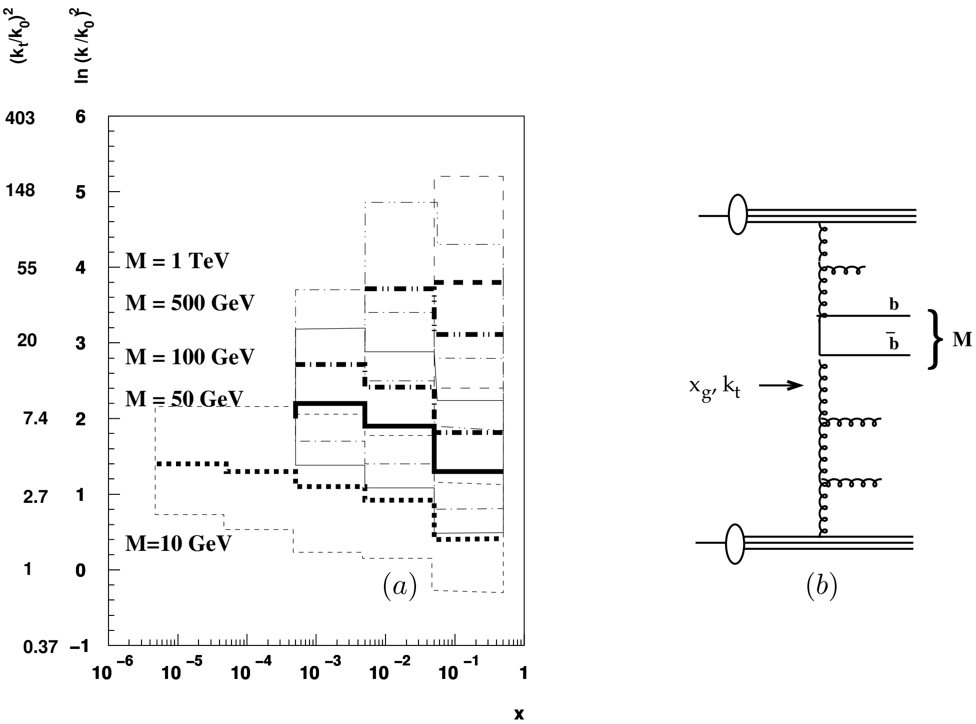


Fig. 7: Average tranverse gluon momenta k_t in processes $gg \to X$ for different masses M of the system X as a function of the momentum fraction of one of the gluons x. The thin lines indicate the RMS spread of the distributions. In (b) is shown the definition of x, k_t and M for a gluon induced process.

theoretical approach, formulated for small x, which takes into account the tranverse momenta is the k_t-factorization [80, 81] or semi-hard [82] approach.

In k_t-factorization the cross section for any process $pp \to X$ can be written as:

$$\sigma = \int dx_1 dx_2 \int dk_{t\,1} dk_{t\,2} \mathcal{A}(x_1, k_{t\,1}, q) \mathcal{A}(x_2, k_{t\,2}, q) \hat{\sigma}(x_1, x_2, k_{t\,1} k_{t\,2}, q) \tag{28}$$

with $\mathcal{A}(x, k_t, q)$ being the un-integrated (k_t-dependent) parton density function uPDF, q defines the factorization scale and $\hat{\sigma}$ is the partonic cross section. The off-shell matrix-elements $\hat{\sigma}$ are calculated in [80, 81].

[7]Contributed by S.P. Baranov, H. Jung, A.V. Lipatov and N.P. Zotov

The effects of finite transverse momenta are present independent of the evolution scheme: uPDFs can be defined also for the DGLAP evolution. A more detailed discussion on these effects can be found in [83, 84].

It is interesting to note, that the k_t-factorization approach (in LO α_s) agrees very well with calculations performed in the collinear approach in NLO α_s, which is shown in [85]. The main effect comes from a more realistic treatment of the kinematics already in LO, which otherwise has to be covered in NLO. The k_t factorization approach, however, is strictly valid only at small x, where the virtuality of the exchanged gluons can be identified with its tranverse momentum $k^2 \sim -k_t^2$. The full expression for the virtuality is [86]:

$$k^2 = \frac{-k_t^2}{1-x} - \frac{x \cdot m^2}{1-x} \tag{29}$$

with m being the recoiling mass of the hadronic system except the hard scattering process, taking into account the history of the evolution process. For finite x the mass effects can be substantial.

6.2 Open $b\bar{b}$ production and correlations at the LHC

Heavy quark production in the k_t-factorization approach at HERA and the Tevatron was considered already in many papers (see, for example, [82, 87, 88, 89, 90]). In Ref. [91] the k_t-factorization approach was used for a more detailed analysis of the D0 and CDF experimental data. The effect of the initial gluon tranverse momenta on the kinematics of the $b\bar{b}$ production at the LHC were investigated [92]. The renormalization and factorization scales were set equal to either the initial gluon virtualities, $\mu_R^2 = \mu_F^2 = q_{T1,2}^2$, or $\mu_F^2 = m_{bT}^2$, as is in the standard collinear QCD, and the quark mass of $m_b = 4.75$ GeV was used.

In Fig. 8a we show the transverse momentum distributions of B mesons at LHC energies. The calculation was performed in the range $|\eta^B| < 1$ and the Peterson fragmentation with $\epsilon = 0.006$ using the KMS [93] parameterization for the un-integrated gluon density (see [83, 84]). The prediction for the azimuthal correlations between the muons coming from B meson decays are shown in Fig. 8b with $p_t^\mu > 6$ GeV and $|\eta^\mu| < 2.5$. The azimuthal correlations indicate an important theoretical difference

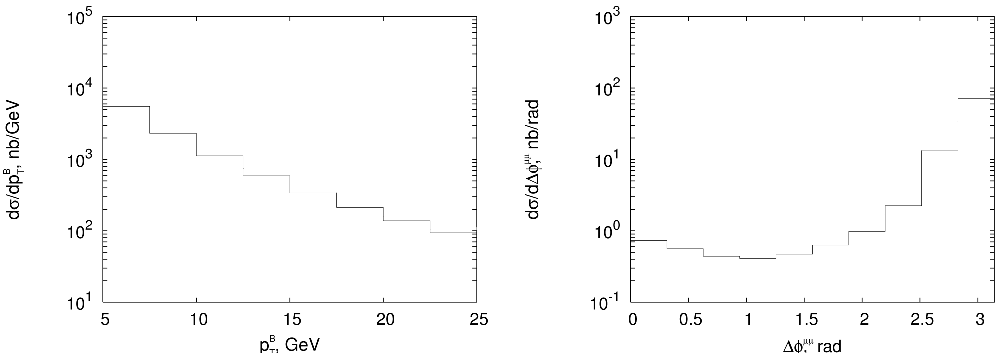


Fig. 8: Prediction for B-meson production at the LHC using the KMS un-integrated gluon density. In a the p_T distribution of B-mesons is shown. In b the azimuthal $\mu\mu$ correlation coming from the B decays is shown.

between the collinear and k_t-factorization approaches. In the collinear approximation at parton level and leading order, the b quarks are be produced exactly back-to-back, which is clearly unphysical when the gluon is evolved up to a large enough scale. Only starting with NLO a significant deviation from the back-to-back scenario is found. Thus the NLO calculation has to correct for the wrong kinematics in LO together with higher order corrections, leading to large K factors. In the k_t-factorization, the transverse momenta of the gluons are correctly treated already in LO. In the k_t - factorization approach the NLO corrections are therefore expected to be much smaller, since here only the purely dynamical corrections have to be applied, whereas the kinematics are already correctly treated in LO.

6.3 Quarkonium production and polarization at the LHC

Since the initial gluons have non-zero transverse momenta, they are off-shell, and they have a longitudinal component in their polarization vector. Typically, the k_t values of the two colliding gluons are much different, as the parton evolution is equivalent to the random walk in the $\ln|k_t|$ plane, not in k_t plane. Roughly speaking, the k_t of one of the gluons can be neglected in comparison with that of the other. So, in the initial state we have one nearly on-shell (transversely polarized) gluon and one off-shell (longitudinally polarized) gluon. After the interaction, they convert into one on-shell gluon and a heavy vector meson. Simple helicity conservation arguments show that the polarization of vector meson must be longitudinal, in contrast with the ordinary parton model, where the initial gluons are both on-shell. This effect has been already studied for the HERA [94] and Tevatron [95] conditions. Fig.9a shows the predictions for the LHC energy obtained with KMS [93] parameterization for un-integrated gluon densities. The calculations are restricted to the pseudorapidity interval $-2.5 < \eta_\Upsilon < 2.5$ and assume ATLAS "$\mu 6\mu 3$" trigger cut, which means one muon with $p_t > 6$ GeV and another muon with $p_t > 3$ GeV.

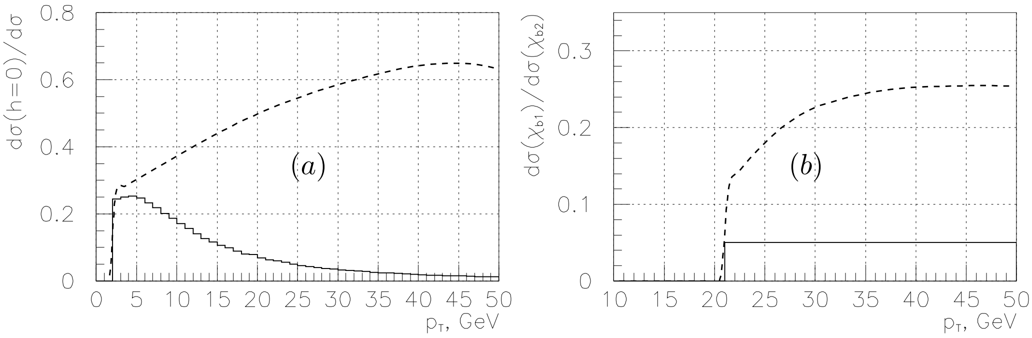


Fig. 9: Predictions of different theoretical approaches for quarkonium production. In (a) the fraction of longitudinally polarized Υ mesons is shown: solid histogram - collinear parton model, singlet + octet; dashed - k_t-factorization with KMS u.g.d.. In (b) the ratio of the production rates χ_{b1}/χ_{b2} is shown: solid histogram - collinear parton model, singlet + octet; dashed - k_t factorization with KMS u.g.d.

Important effects are also seen in the production of P-wave bottomium states with different spins χ_{b1} and χ_{b2}. At the Tevatron energies, this process has been considered in Ref. [96], and the predictions for the LHC are presented in Fig.9b. The P-wave states are assumed to be detected via the decay $\chi_b \to \Upsilon + \gamma$, with an additional requirement that the energy of the decay photon be greater than 2 GeV. The ratio of the production rates $\sigma(\chi_1)/\sigma(\chi_2)$ is qualitatively different in the k_t-factorization and the collinear parton model. The underlying physics is clearly connected with gluon off-shellness. In the collinear parton model, the relative suppression of χ_1 states becomes stronger with increasing p_T because of the increasing role of the color-octet contribution: in this approach, the leading-order fragmentation of an on-shell transversely polarized gluon into a vector meson is forbidden. In contrast with that, in the k_t-factorization approach, the increase in the final state p_T is only due to the increasing transverse momenta (and virtualities) of the initial gluons, and, consequently, the suppression motivated by the Landau-Yang theorem becomes weaker at large p_T.

6.4 Associated Higgs + jets production at the LHC

The dominant mechanism for Higgs production at the LHC is gluon-gluon fusion, and the calculations can be significantly simplified in the large top mass limit ($M_H \leq 2M_{top}$) [97]. The differential cross section of the inclusive Higgs production $p\bar{p} \to H + X$ in the k_t-factorization approach has been calculated in [98, 99] and can be written as:

$$\frac{d\sigma(p\bar{p}\to H+X)}{dy_H} = \int \frac{\alpha_s^2(\mu^2)}{288\pi}\frac{G_F\sqrt{2}}{x_1x_2m_H^2s}\left[m_H^2+\mathbf{p}_T^2\right]^2\cos^2(\Delta\varphi)\times$$
$$\times\mathcal{A}(x_1,\mathbf{k}_{1T}^2,\mu^2)\mathcal{A}(x_2,\mathbf{k}_{2T}^2,\mu^2)d\mathbf{k}_{1T}^2d\mathbf{k}_{2T}^2\frac{d(\Delta\varphi)}{2\pi}, \tag{30}$$

where G_F is the Fermi coupling constant, $\mathcal{A}(x,\mathbf{k}_T^2,\mu^2)$ is the un-integrated gluon distribution, $\Delta\varphi$ the azimuthal angle between the momenta $\mathbf{k}_{1T}$ and $\mathbf{k}_{2T}$, and the transverse momentum of the produced Higgs boson is $\mathbf{p}_T = \mathbf{k}_{1T}+\mathbf{k}_{2T}$. It should be noted, that this process is particularly interesting in k_t-factorization, as the transverse momenta of the gluons are in the same order as their longitudinal momenta ($\sim\mathcal{O}(10$ GeV$)$) [100].

The total inclusive Higgs production cross section at the LHC energies ($\sqrt{s}=14$ TeV) is plotted in Fig. 10(a) as a function of the Higgs mass in the mass range $m_H = 100-200$ GeV. The solid line is obtained by fixing both the factorization and renormalization scales at the default value $\mu = m_H$ with the J2003 (set 1) CCFM un-integrated gluon distribution [101]. In order to estimate the theoretical uncertainties, we take $\mu = \xi m_H$ and vary the scale parameter ξ between $1/2$ and 2 about the default value $\xi = 1$. The uncertainty band is presented by the upper and lower dashed lines. We find that our central values agree very well with recent NNLO results [102].

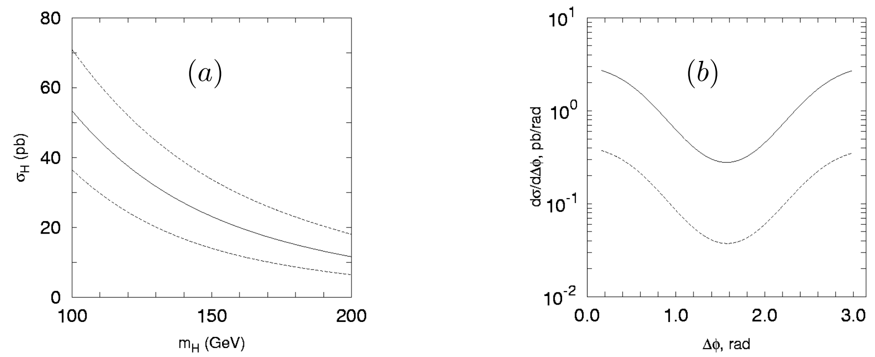


Fig. 10: Higg production at the LHC. In (a) the total cross section for Higgs boson production as a function of Higgs mass is shown: the solid curve corresponds to the default scale $\mu = m_H$, upper and lower dashed curves - $\mu = m_H/2$ and $\mu = 2m_H/2$ respectively. In (b) the jet-jet azimuthal angle distribution in the Higgs+jet+jet production at $\sqrt{s} = 14$ TeV. The kinematical cut $|\mathbf{p}_{\mathrm{jet}\,T}| > 20$ GeV was applied for both jets. Solid and dashed lines correspond to the J2003 (set 1) and J2003 (set 2) [101] u.g.d. respectively.

To demonstrate the capabilities of the k_t-factorization approach, we calculate the azimuthal angle $\Delta\phi$ distribution between the two final jets transverse momenta in the Higgs+jet+jet production process. Our results are shown in Fig. 10(b). The dip at $\Delta\phi = \pi/2$ comes from the $\cos(\Delta\varphi)$ in eq.(30). In the approach presented here, the k_t of the initial gluons is approximately compensated by the transverse momenta of the jets [103]: $\mathbf{k}_T \simeq -\mathbf{p}_{T,jet}$, and, consequently, $\Delta\phi\simeq\Delta\varphi$ applying a cut-off $|\mathbf{p}_{\mathrm{jet}\,T}| > 20$ GeV. This dip is characteristic for the loop-induced Higgs coupling to gluons in the framework of fixed-order perturbative QCD calculations [102]. Thus, we illustrate that the features usually interpreted as NNLO effects are reproduced in the k_t-factorization with LO matrix elements.

However, we see a very large difference (about one order of magnitude) between the predictions based on the J2003 gluon densities set 1 and set 2 [101], showing the sensitivity to the shape of the un-integrated gluon density.

6.5 Conclusions

The finite k_t of the initial state gluons significantly modifies the kinematics of the gluon-gluon fusion process and leads to nontrivial angular correlations between the final state heavy quarks. The longitudinal polarization of the initial off-shell gluons manifests in the longitudinal polarization of J/ψ and Υ mesons at moderate p_T and, also, affects the production rates of P-wave quarkonia.

The predictions in k_t-factorization are very close to NNLO pQCD results for the inclusive Higgs production at the LHC, since the main part of high-order collinear pQCD corrections is already included in the k_t-factorization. In the k_t-factorization approach the calculation of associated Higgs+jets production is much simpler than in the collinear factorization approach. However, the large scale dependence of our calculations (of the order of $20 - 50\%$) probably indicates the sensitivity to the unintegrated gluon distributions.

7 Baryon charge transfer and production asymmetry of $\Lambda^0/\bar{\Lambda}^0$ in hadron interactions[8]

7.1 Introduction to the QGSM

The phenomenon of nonzero asymmetry of baryon production with nonbaryonic beams (π,μ,e) was mentioned and explained in a few theoretical papers. Baryon charge can be transferred from proton or nucleus targets through the large rapidity gap with the string junction. In baryonic beam reactions (p,A, etc.) this effect is displayed in the valuable baryon/antibaryon spectrum asymmetry at $y = 0$. Every theoretical discourse on baryon charge transfer appeals to the value of the intercept, $\alpha_{SJ}(0)$, that is an intercept of the Regge-trajectory of imaginary particles which consists only of string junctions from baryon and antibaryon. Practically, the models that can account for this effect are only non perturbative QCD phenomenological models: the Dual Parton Model (DPM) [104] and the Quark Gluon String Model (QGSM) [105] as well as the DPMJET Monte Carlo expansion of these two models. Both analytical models are similar and they were based on the common Regge asymptotic presentation of constituent quark structure functions and string (quark) fragmentation functions. Here we are considering QGSM. In the comparison to the other models, QGSM accounts for many Pomeron exchanges. This approach works very well to give us the correct description of particle production cross sections at very high energies. The QGSM procedures of constructing of quark/diquark structure functions and fragmentation functions were presented in many previous publications. We take into consideration the π-p reaction that gives similar asymmetries as the γ-p reaction. The spectra in this reaction are more sensitive to the baryon excess in the region of positive x_F than the spectra of baryons in p-p collisions.

7.2 Comparison with Experimental Data

The asymmetry $A(y)$ between the spectra of baryons and antibaryons is defined as:

$$A(x) = \frac{dN^{\Lambda^0}/dx - dN^{\bar{\Lambda}^0}/dx}{dN^{\Lambda^0}/dx + dN^{\bar{\Lambda}^0}/dx}, \tag{31}$$

The EHS and the NA49 experiments have presented evidence for a nonzero baryon production asymmetry in proton-proton fixed target interactions, measuring at $y = 0$ values of the order of 0.5 - 0.3. In pion-proton interactions (E769) we can see the y dependence of the asymmetry and the measured asymmetry, which was multiplied by a factor of 2 in order to be compared with the pp asymmetry.

The data from these experiments can be presented in one plot for $A(\Delta y)$, where Δy is the rapidity distance from interacting target-proton (see Fig. 11). It is seen that the points are situated on the same line. If we add the data from proton-nucleus experiments (HERA-B and RHIC) they are still approximately on this line. Only the STAR asymmetry point at $\sqrt{s} = 130$ GeV can be interpreted as a sign that the curve is bent. And the result of the H1 experiment at HERA [106] calls certainly for a steeper curve.

[8] Contributed by O.I. Piskounova.

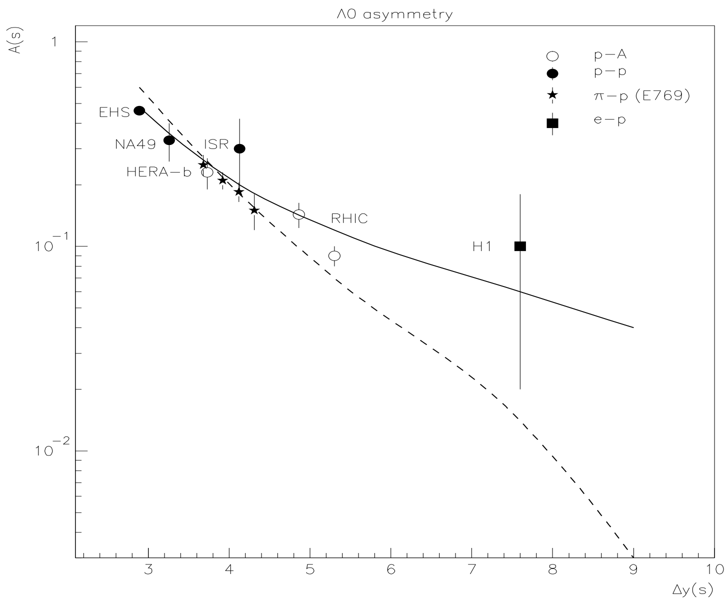


Fig. 11: Asymmetry in Λ^0 and $\bar{\Lambda}^0$ production and QGSM curves: α_{SJ}(0)=0.5(dashed line) and α_{SJ}(0)=0.9 (solid line).

What means do we have in QGSM to describe this dependence? In Ref. [107] it was shown that the data of the E769 and H1 experiments can not be described with the same value of $\alpha_{SJ}(0)$: the points at lower energy correspond to α_{SJ}(0)=0.5, while the H1 point requires α_{SJ}(0)=0.9.

7.3 Summary

The purpose of this contribution is to show the band of asymmetries that can be predicted for the LHC experiments between the two possibilities given above for α_{SJ}(0). The results are shown in Fig. 11. The solid line represents the case of α_{SJ}(0)=0.9. This curve fits the data at low energies (small Δy) due to varying the energy splitting between string junction and diquark configuration: 0.1*SJ+0.9*DQ. What we had to tune also was the fragmentation parameter af=0.15 instead of 0.2 accepted in previous papers. Also the curve for α_{SJ}(0)=0.5 is shown in Fig. 11 with a dashed line. This line certainly doesn't fit the H1 point and gives a negligible asymmetry at the energy of the LHC experiments. Finally, we have the prediction for strange baryon asymmetries at the LHC within the range: $0.003 < A < 0.04$. The same procedure has to be applied to charmed baryon asymmetry to get the predictions at LHC energy.

References

[1] W. Stirling. Private communication.

[2] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988).

[3] W. Beenakker, H. Kuijf, W. L. van Neerven, and J. Smith, Phys. Rev. **D40**, 54 (1989).

[4] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B327**, 49 (1989). Erratum-ibid. B **335** (1989) 260.

[5] W. Beenakker, W. L. van Neerven, R. Meng, G. A. Schuler, and J. Smith, Nucl. Phys. **B351**, 507 (1991).

[6] M. L. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B373**, 295 (1992).

[7] R. K. Ellis and P. Nason, Nucl. Phys. **B312**, 551 (1989).

[8] J. Smith and W. L. van Neerven, Nucl. Phys. **B374**, 36 (1992).

[9] S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B412**, 225 (1994). hep-ph/9306337.

[10] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, Nucl. Phys. **B392**, 162 (1993).

[11] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, Nucl. Phys. **B392**, 229 (1993).

[12] S. Riemersma, J. Smith, and W. L. van Neerven, Phys. Lett. **B347**, 143 (1995). hep-ph/9411431.

[13] B. W. Harris and J. Smith, Nucl. Phys. **B452**, 109 (1995). hep-ph/9503484.

[14] J. C. Collins and W.-K. Tung, Nucl. Phys. **B278**, 934 (1986).

[15] B. Mele and P. Nason, Nucl. Phys. **B361**, 626 (1991).

[16] M. A. G. Aivazis, F. I. Olness, and W.-K. Tung, Phys. Rev. **D50**, 3085 (1994). hep-ph/9312318.

[17] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, Phys. Rev. **D50**, 3102 (1994). hep-ph/9312319.

[18] J. C. Collins, Phys. Rev. **D58**, 094002 (1998). hep-ph/9806259.

[19] M. Kramer, F. I. Olness, and D. E. Soper, Phys. Rev. **D62**, 096007 (2000). hep-ph/0003035.

[20] J. C. Collins, F. Wilczek, and A. Zee, Phys. Rev. **D18**, 242 (1978).

[21] W.-K. Tung, S. Kretzer, and C. Schmidt, J. Phys. **G28**, 983 (2002). hep-ph/0110247.

[22] S. Kretzer, H. L. Lai, F. I. Olness, and W. K. Tung, Phys. Rev. **D69**, 114005 (2004). hep-ph/0307022.

[23] S. Kretzer and F. I. Olness (2005). hep-ph/0508216.

[24] M. Cacciari, M. Greco, and P. Nason, JHEP **05**, 007 (1998). hep-ph/9803400.

[25] R. S. Thorne and R. G. Roberts, Phys. Rev. **D57**, 6871 (1998). hep-ph/9709442.

[26] R. S. Thorne and R. G. Roberts, Phys. Lett. **B421**, 303 (1998). hep-ph/9711223.

[27] A. Chuvakin, J. Smith, and W. L. van Neerven, Phys. Rev. **D61**, 096004 (2000). hep-ph/9910250.

[28] M. Gluck, E. Reya, and M. Stratmann, Nucl. Phys. **B422**, 37 (1994).

[29] M. Buza, Y. Matiounine, J. Smith, R. Migneron, and W. L. van Neerven, Nucl. Phys. **B472**, 611 (1996). hep-ph/9601302.

[30] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, Phys. Lett. **B411**, 211 (1997). hep-ph/9707263.

[31] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, Eur. Phys. J. **C1**, 301 (1998). hep-ph/9612398.

[32] Y. Matiounine, J. Smith, and W. L. van Neerven, Phys. Rev. **D57**, 6701 (1998). hep-ph/9801224.

[33] A. Chuvakin, J. Smith, and W. L. van Neerven, Phys. Rev. **D62**, 036004 (2000). hep-ph/0002011.

[34] A. Chuvakin and J. Smith, Comput. Phys. Commun. **143**, 257 (2002). hep-ph/0103177.

[35] A. Vogt, Comput. Phys. Commun. **170**, 65 (2005). hep-ph/0408244.

[36] A. Chuvakin, J. Smith, and B. W. Harris, Eur. Phys. J. **C18**, 547 (2001). hep-ph/0010350.

[37] M. Cacciari and M. Greco, Nucl. Phys. **B421**, 530 (1994). hep-ph/9311260.

[38] M. Cacciari and M. Greco, Z. Phys. **C69**, 459 (1996). hep-ph/9505419.

[39] M. Cacciari, S. Frixione, and P. Nason, JHEP **03**, 006 (2001). hep-ph/0102134.

[40] F. Aversa, P. Chiappetta, M. Greco, and J. P. Guillet, Nucl. Phys. **B327**, 105 (1989).

[41] S. Frixione and P. Nason, JHEP **03**, 053 (2002). hep-ph/0201281.

[42] M. Cacciari and P. Nason, Phys. Rev. Lett. **89**, 122003 (2002). hep-ph/0204025.

[43] M. Cacciari and P. Nason, JHEP **09**, 006 (2003). hep-ph/0306212.

[44] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, JHEP **07**, 033 (2004). hep-ph/0312132.

[45] M. Cacciari, P. Nason, and R. Vogt, Phys. Rev. Lett. **95**, 122001 (2005). hep-ph/0502203.

[46] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, Phys. Rev. **D71**, 014018 (2005). hep-ph/0410289.

[47] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, Eur. Phys. J. **C41**, 199 (2005). hep-ph/0502194.

[48] F. I. Olness, R. J. Scalise, and W.-K. Tung, Phys. Rev. **D59**, 014506 (1999). hep-ph/9712494.

[49] B. A. Kniehl, T. Kneesch, G. Kramer, and I. Schienbein, work in progress.

[50] S. Kretzer and I. Schienbein, Phys. Rev. **D58**, 094035 (1998). hep-ph/9805233.

[51] S. Kretzer and I. Schienbein, Phys. Rev. **D59**, 054004 (1999). hep-ph/9808375.

[52] G. Kramer and H. Spiesberger, Eur. Phys. J. **C22**, 289 (2001). hep-ph/0109167.

[53] G. Kramer and H. Spiesberger, Eur. Phys. J. **C28**, 495 (2003). hep-ph/0302081.

[54] G. Kramer and H. Spiesberger, Eur. Phys. J. **C38**, 309 (2004). hep-ph/0311062.

[55] J. Pumplin *et al.*, JHEP **07**, 012 (2002). hep-ph/0201195.

[56] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, in preparation.

[57] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. **91**, 241804 (2003). hep-ex/0307080.

[58] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, *Hadroproduction of d and b mesons in a massive vfns*, 2005. Proceedings of the 13th International Workshop on Deep Inelastic Scattering (DIS 2005), Madison, Wisconsin, USA, April 27 - May 1, 2005., hep-ph/0507068.

[59] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, *Reconciling open charm production at the fermilab tevatron with qcd*, 2005. `hep-ph/0508129`.

[60] J. Binnewies, B. A. Kniehl, and G. Kramer, Phys. Rev. **D58**, 014014 (1998). `hep-ph/9712482`.

[61] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. **D71**, 012005 (2005). `hep-ex/0410058`.

[62] D0 Collaboration, V. M. Abazov *et al.*, Phys. Lett. **B622**, 265 (2005). `hep-ex/0505063`.

[63] J. Campbell *et al.*, *Higgs boson production in association with bottom quarks* (unpublished). 2004.

[64] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C40**, 349 (2005). `hep-ex/0411046`.

[65] H1 Collaboration, A. Aktas *et al.* (2005). `hep-ex/0507081`.

[66] M. Dittmar and K. Mazumdar, *Measuring the parton Distribution functions of charm, beauty and strange quarks and anti-quarks at the LHC* (unpublished). CMS note 2001/002.

[67] J. Campbell, R. K. Ellis, F. Maltoni, and S. Willenbrock, Phys. Rev. **D69**, 074021 (2004). `hep-ph/0312024`.

[68] D0 Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. **94**, 161801 (2005). `hep-ex/0410078`.

[69] C. Kao and N. Stepanov, Phys. Rev. **D52**, 5025 (1995). `hep-ph/9503415`.

[70] V. D. Barger and C. Kao, Phys. Lett. **B424**, 69 (1998). `hep-ph/9711328`.

[71] S. Dawson, D. Dicus, and C. Kao, Phys. Lett. **B545**, 132 (2002). `hep-ph/0208063`.

[72] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001). `hep-ph/0010017`.

[73] E. Richter-Was, *AtlfastTemp package- ATLAS fast simulation (fortran) integrated with Athena framework* (unpublished). ATLAS internal note ATL-COM-PHYS-2000-023.

[74] G. Sterman, Nucl. Phys. **B281**, 310 (1987).

[75] S. Catani and L. Trentadue, Nucl. Phys. **B327**, 323 (1989).

[76] E. Laenen and S.-O. Moch, Phys. Rev. **D59**, 034027 (1999). `hep-ph/9809550`.

[77] S. Catani, M. L. Mangano, P. Nason, and L. Trentadue, Nucl. Phys. **B478**, 273 (1996). `hep-ph/9604351`.

[78] A. Vogt, Phys. Lett. **B471**, 97 (1999). `hep-ph/9910545`.

[79] G. Sterman and W. Vogelsang, hep-ph/0002132 (2000). `hep-ph/0002132`.

[80] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B 366**, 135 (1991).

[81] J. Collins and R. Ellis, Nucl. Phys. **B 360**, 3 (1991).

[82] E. Levin, M. Ryskin, Y. Shabelski, and A. Shuvaev, Sov. J. Nucl. Phys. **53**, 657 (1991).

[83] Small x Collaboration, B. A. et al., Eur. Phys. J. **C 25**, 77 (2002). `hep-ph/0204115`.

[84] Small x Collaboration, J. R. Andersen *et al.*, Eur. Phys. J. **C35**, 67 (2004). `hep-ph/0312333`.

[85] M. C. et al, *Benchmark cross sections for heavy flavor producton*, 2005. This proceedings.

[86] J. Collins and H. Jung, *Need for fully unintegrated parton densities*, 2005. hep-ph/0508280 and these proceedings.

[87] M. Ryskin, A. Shuvaev, and Y. Shabelski, Phys. Atom. Nucl. **64**, 120 (2001). hep-ph/9907507.

[88] P. Hagler, R. Kirschner, A. Schafer, L. Szymanowski, and O. Teryaev, Phys. Rev. **D62**, 071502 (2000). hep-ph/0002077.

[89] A. V. Lipatov, L. Lonnblad, and N. P. Zotov, JHEP **01**, 010 (2004). hep-ph/0309207.

[90] H. Jung, Phys. Rev. **D 65**, 034015 (2002). DESY-01-136, hep-ph/0110034.

[91] S. P. Baranov, N. P. Zotov, and A. V. Lipatov, Phys. Atom. Nucl. **67**, 837 (2004).

[92] S. P. Baranov, A. V. Lipatov, and N. P. Zotov, Nucl. Phys. Proc. Suppl. **146**, 228 (2005).

[93] J. Kwiecinski, A. Martin, and A. Stasto, Phys. Rev. **D 56**, 3991 (1997).

[94] A. V. Lipatov and N. P. Zotov, Eur. Phys. J. **C27**, 87 (2003). hep-ph/0210310; S. P. Baranov and N. P. Zotov, J. Phys. **G29**, 1395 (2003). hep-ph/0302022.

[95] S. P. Baranov, Phys. Rev. **D66**, 114003 (2002).

[96] S. P. Baranov, Phys. Lett. **B594**, 277 (2004).

[97] J. Ellis, M. Gaillard, and D. Nanopoulos, Nucl. Phys. **B106**, 292 (1976); M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, Sov. J. Nucl. Phys. **30**, 711 (1979).

[98] F. Hautmann, Phys. Lett. **B535**, 159 (2002). hep-ph/0203140.

[99] A. V. Lipatov and N. P. Zotov, *Higgs boson production at hadron colliders in the k(t)-factorization approach*, 2005. hep-ph/0501172, DESY-05-020.

[100] H. Jung, Mod. Phys. Lett. **A19**, 1 (2004). hep-ph/0311249.

[101] M. Hansson and H. Jung, *The status of ccfm unintegrated gluon densities*, 23-27 April 2003. DIS 2003, St. Petersburg, Russia, hep-ph/0309009.

[102] R. Harlander and W. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002). hep-ph/0201206; C. Anastasiou and K. Melnikov, Nucl. Phys. **B646**, 220 (2002). hep-ph/0207004; V. Ravindran, J. Smith, and W. van Neerven, Nucl. Phys. **B665**, 325 (2003). hep-ph/0302135.

[103] S. P. Baranov. and N. P. Zotov, Phys. Lett. **B491**, 111 (2000).

[104] A. Capella and J. Tran Thanh Van, Z. Phys. **C10**, 249 (1981).

[105] A. B. Kaidalov and K. A. Ter-Martirosian, Sov. J. Nucl. Phys. **39**, 979 (1984).

[106] H1 Collaboration, C. Adloff *et al.*, Z. Phys. **C76**, 613 (1997). hep-ex/9708016.

[107] G. H. Arakelian, A. Capella, A. B. Kaidalov, and Y. M. Shabelski, Eur. Phys. J. **C26**, 81 (2002).

*

Experimental overview of heavy quark measurements at HERA

O. Behnke[1], *A. Geiser*[2], *A. Meyer*[2]
[1] Universität Heidelberg, Germany
[2] DESY Hamburg, Germany

Abstract
Experimental results on heavy flavour production at HERA are reviewed in the context of their relevance for future measurements at the LHC.

1 Introduction

Measurements of heavy flavour production at HERA can have significant impacts on the preparation and understanding of heavy flavour measurements at the LHC, and on the understanding of background processes to new physics discoveries [1]. The purpose of this contribution is to summarize the current status of heavy flavour measurements at HERA, and provide an outlook on how they might improve in the near future. The relation of these measurements to measurements at the LHC will be covered in more detail in subsequent contributions [2, 3, 4]. Since the top quark is too heavy to be produced at HERA with a significant rate, the term "heavy flavour" refers to b and c quarks only. The dominant diagram for heavy flavour production at HERA is shown in Fig. 1. The theory of heavy quark production at HERA is covered in the theoretical review section [5].

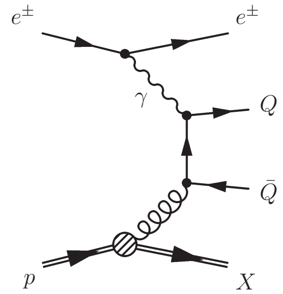


Fig. 1: Feynman graph for the production of a heavy quark pair via the leading order boson-gluon-fusion (BGF) process.

The interest in heavy flavour production arises from several aspects.

- Tagging a heavy flavour particle, e.g. inside a jet, establishes that this jet arises from a quark rather than a gluon. The number of possible QCD diagrams is thus reduced, and specific QCD final states can be studied more precisely than in inclusive measurements. This is even more true when *both* quarks of a heavy flavour quark pair can be tagged.
- The charm and beauty masses ($m_b, m_c \gg \Lambda_{QCD}$) provide energy scales which are large enough to allow perturbative calculations using a "massive" scheme [6, 7]. *All* QCD-processes involving heavy quarks should thus be reliably calculable. However, these mass scales often compete with other scales occurring in the same process, such as the transverse momentum (p_T) of the heavy quarks, or the virtuality of the exchanged photon, Q^2. Since the perturbative expansion can not be optimized for all scales at once, additional theoretical uncertainties enter which reduce the reliability of the predictions. If one of the competing scales (p_T, Q^2) is much larger than the quark mass, approximations in which the heavy quarks are treated as massless [8, 9, 10, 11, 12, 13, 14] can improve the reliability. Mixed schemes [15, 16, 17] are also possible. Understanding and resolving these difficulties should contribute to the understanding of multi-scale problems in general.

- Tagging the final state also constrains the initial state. Therefore, heavy flavour measurements can be used to measure or constrain parton density functions. In particular, Fig. 1 illustrates the direct sensitivity to the gluon density in the proton. Alternatively, in appropriate kinematic ranges, the initial state splitting of the gluon or photon into a heavy quark pair can be absorbed into the parton density definition. If the mass can be neglected, the same diagram (or higher order variants of it) can be reinterpreted as a way to measure the heavy flavour content of the proton or of the photon. These can in turn be used to calculate cross sections for other processes, such as Higgs production at the LHC.
- The production of "hidden" heavy flavour states (quarkonia) yields further insights into the interplay of (perturbative) heavy quark production and (non-perturbative) binding effects.

At HERA, the fraction of charm production vs. inclusive QCD processes is of order 10% in the perturbative QCD regime. Reasonably large samples can therefore be obtained despite the partially rather low tagging efficiency. Beauty production is suppressed with respect to charm production by the larger b mass, and by the smaller coupling to the photon. The resulting total cross section is two orders of magnitude smaller than the one for charm. Beauty studies at HERA are thus often limited by statistics. Kinematically, beauty production at HERA is similar to top production at LHC ($m_b/\sqrt{s_{HERA}} \sim m_t/\sqrt{s_{LHC}}$). On the other hand, in the "interesting" physics region beauty is produced as copiously at the LHC as charm is at HERA.

2 Open charm production

Charmed mesons are tagged at HERA in different ways. A typical mass distribution for the "golden" channel $D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$ (+ c.c.) is shown in Fig. 2 [20]. Despite the low branching ratio, this channel yields large statistics charm samples of high purity. Fig. 3 [21] shows a corresponding D* production cross section in photoproduction for different kinematic variables. In general, D* production

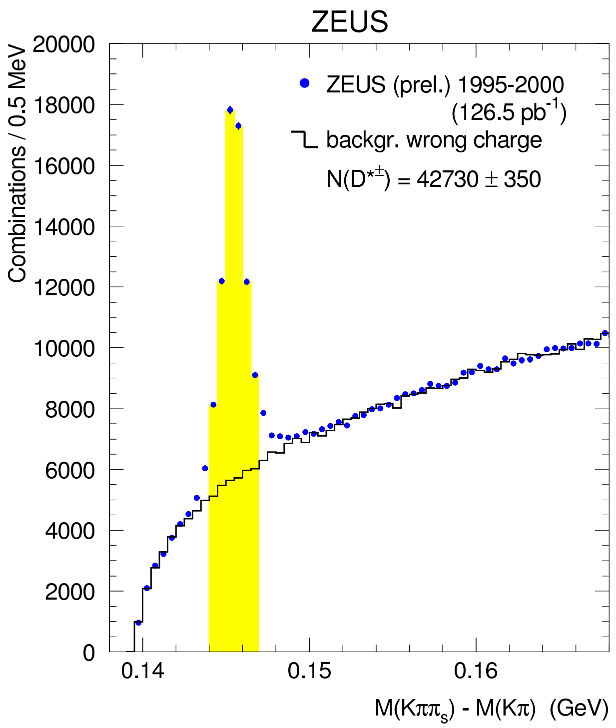


Fig. 2: Total inclusive $D^{*\pm}$ sample obtained by ZEUS for the HERA I data period in the golden decay channel $D^{*+} \to D^0\pi_s^+ \to (K^-\pi^+)\pi_s^+$.

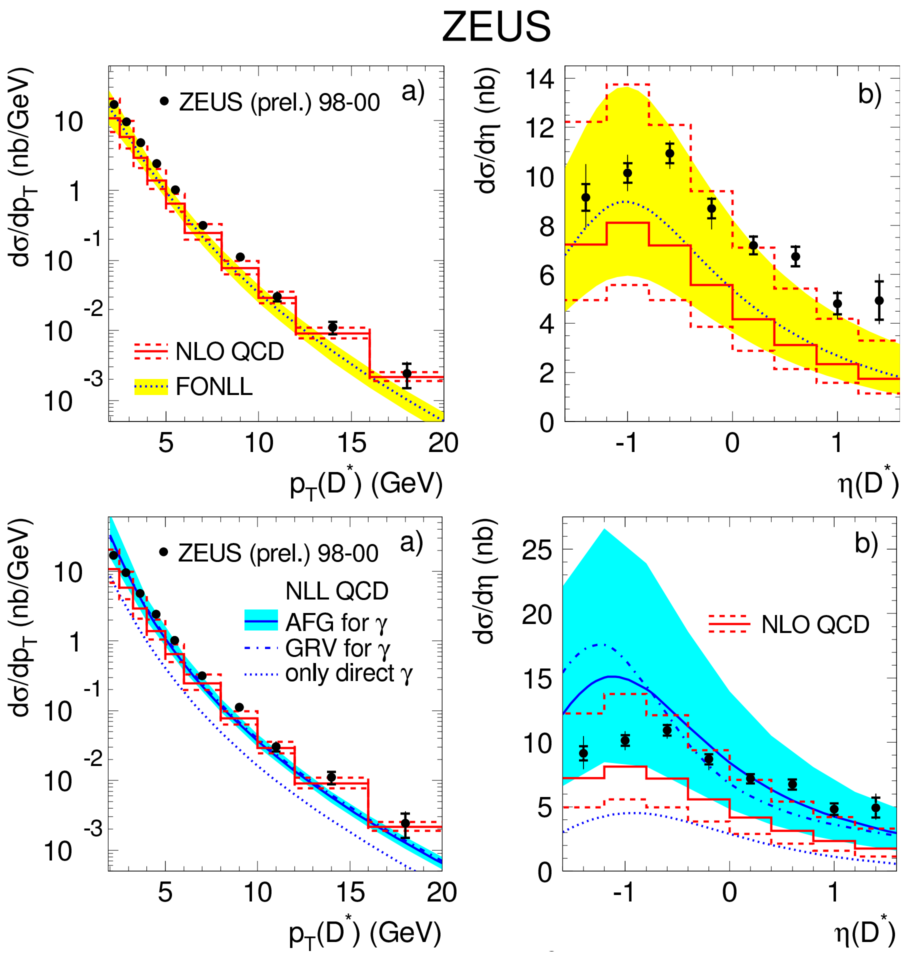


Fig. 3: $D^{*\pm}$ single differential cross sections in photoproduction as function of the $D^{*\pm}$ transverse momentum and pseudo-rapidity. The measurements are compared to NLO calculations in the massive (NLO), massless (NLL) and mixed (FONLL) scheme.

is well described by next-to-leading order QCD predictions, although the data often prefer the upper edge of the theoretical error band. Some deviations are observed in particular regions of phase space. For instance, there are indications that forward (i.e. in the direction of the proton) charm production might be slightly larger than theoretical expectations (Fig. 3b). Also, there are regions of phase space which effectively require four-body final states which are not covered by NLO calculations (see Fig. 5 in [1]). In order to describe such phase space regions, either NNLO calculations, or parton shower extensions to NLO calculations such as MC@NLO [18, 19] will be needed.

Other ways to tag charm include the reconstruction of a secondary vertex in a high resolution Micro-Vertex-Detector (MVD) in addition to the reconstruction of a charmed meson mass (Fig. 4) [22], or the reconstruction of inclusive multiple impact parameters resulting from the finite charm meson lifetime. A resulting cross section for D^+ production is shown in Fig. 5.

Since the charm mass of approximately 1.5 GeV is not very much above the threshold at which perturbative calculations are believed to produce reliable results, the generally good agreement of perturbative QCD predictions with the data is highly nontrivial, and encouraging concerning the validity of corresponding predictions for the LHC.

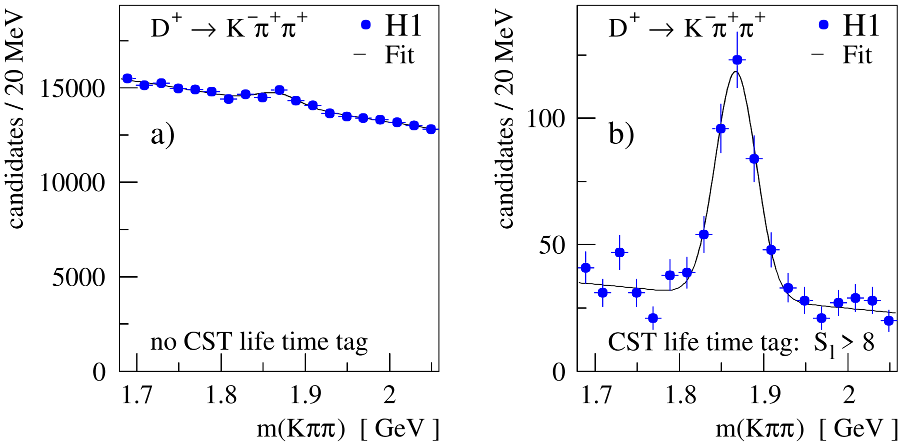


Fig. 4: D^+ mass peak in H1 before (left) and after (right) a cut on the decay length significance.

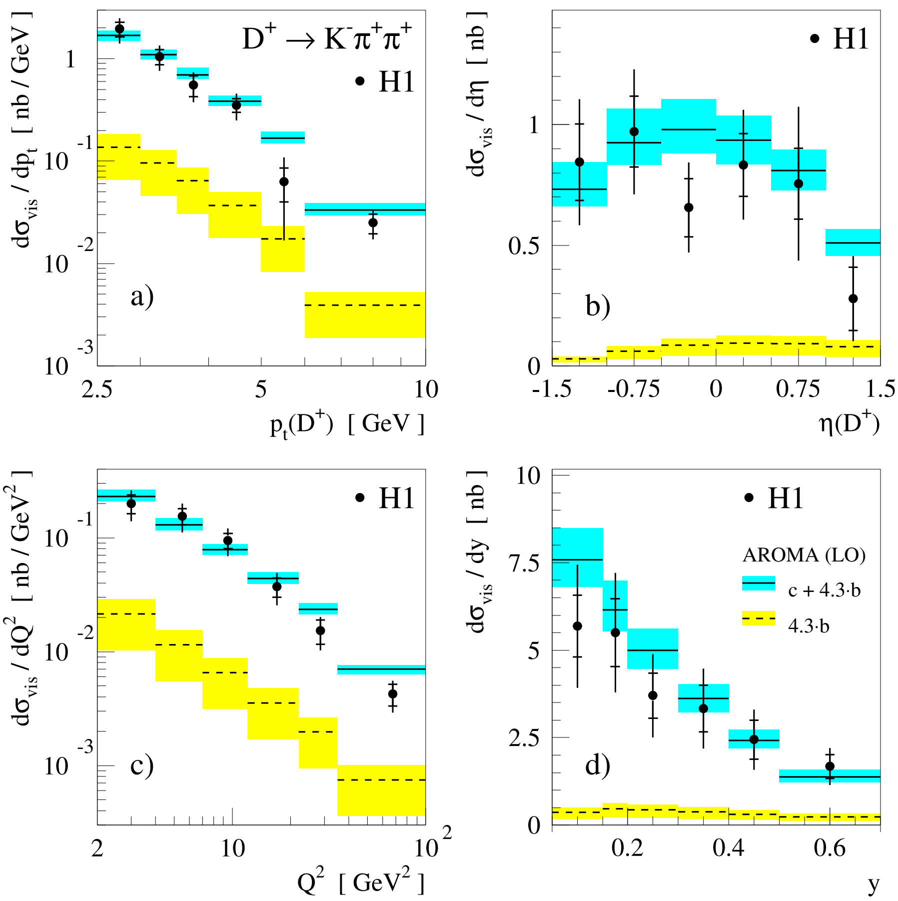


Fig. 5: Cross sections for D^+ production in DIS. A leading order + parton shower QCD prediction (AROMA) is shown for comparison.

3 Open beauty production

Open beauty production is detected at HERA using essentially three different methods, related to the large b mass or long b lifetime.

- If a jet is built out of the fragmentation and decay products of a b meson/quark, the transverse momentum of the decay products with respect to the jet axis will be of order half the b mass. This is significantly larger than for decay products of charm particles, or the transverse momenta induced by non-perturbative fragmentation effects, which are both of order 1 GeV or less. This distribution of this transverse momentum, called p_T^{rel}, can thus be used to measure the beauty contribution to a given jet sample.
- Due to the CKM-suppressed weak deacy of the b quark, the lifetime of b hadrons is longer than that of charmed particles. Furthermore, the larger decay angle due to the larger mass results in a higher significance of the decay signature. Detectors with a resolution in the 100 μ region or better can thus separate the beauty contribution from charm and light flavour contributions.
- Again due to their mass, b hadrons take a larger fraction of the available energy in the fragmentation process. Furthermore, they produce decay products with sizeable transverse momentum even when produced close to the kinematic threshold. Simple lower cuts on the transverse momentum of such decay products therefore enrich the beauty content of a sample. Applying such cuts on two different decay products (double tagging) often sufficiently enriches the beauty content such that the remaining background can be eliminated or measured by studying the correlation between these decay products.

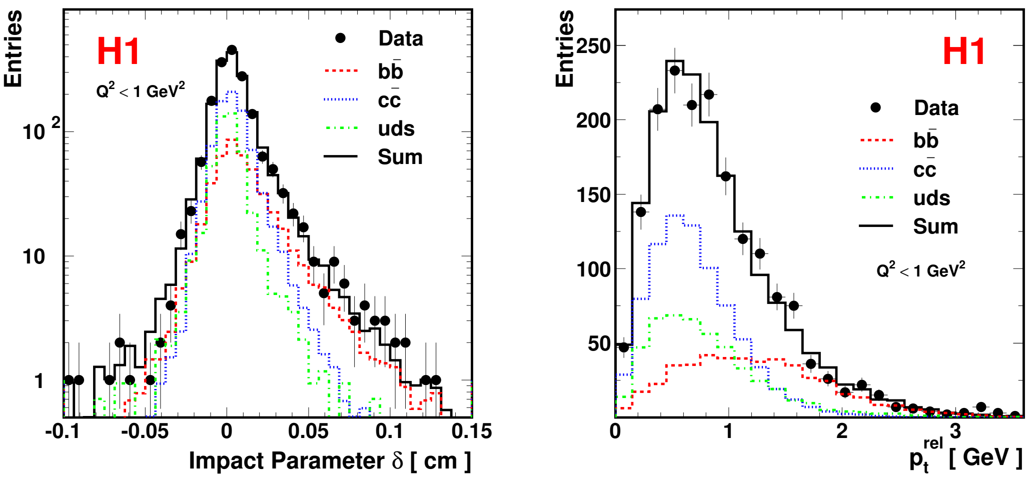


Fig. 6: Distributions of the impact parameter δ of the muon track (left) and the transverse muon momentum p_t^{rel} relative to the axis of the associated jet (right) in H1. Also shown are the estimated contributions of events arising from b quarks (dashed line), c quarks (dotted line) and the light quarks (dash-dotted line).

An example for an analysis using the first two methods with muons from semileptonic b decays is shown in Fig. 6 [23]. Some cross sections resulting from this type of analysis are shown in Fig. 3 of [1]. In general, reasonable agreement is observed between the data and corresponding NLO QCD predictions, although, as in the charm case, the data tend to prefer the upper edge of the theoretical error band. In some regions of phase space, e.g. at low p_T^{μ} or high η^{μ} differences of up to two standard deviations are observed. More precise measurements (section 8) will be needed to decide whether these deviations are really significant.

An example for an analysis using the 2nd method only is shown in Fig. 7 [24], while an example for an analysis using the third method is shown in Fig. 8 [25].

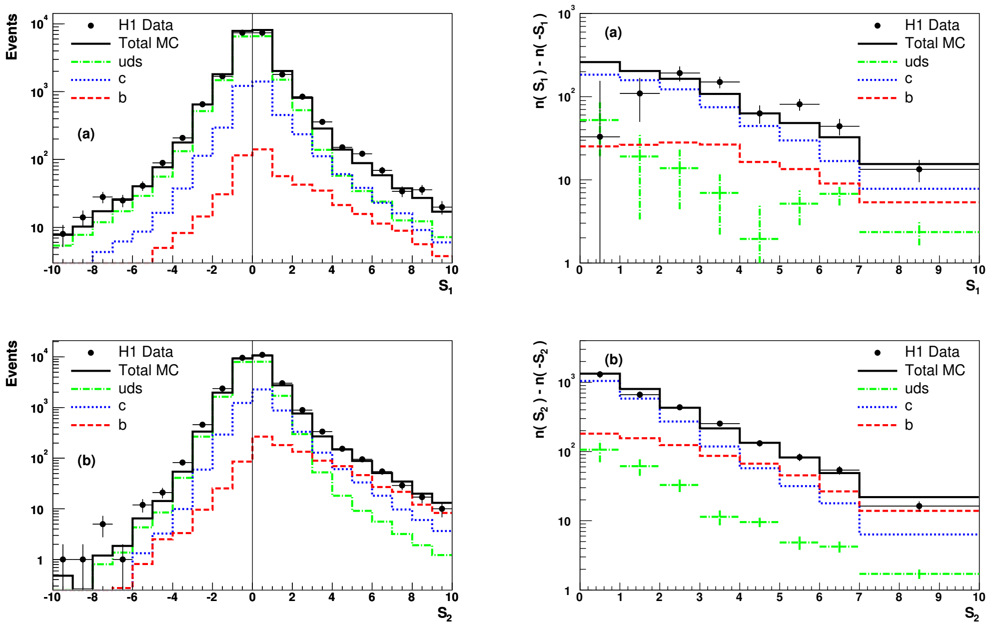


Fig. 7: Upper left: significance $S_1 = \delta/\sigma(\delta)$ distribution per event for events that contain one selected track associated to the jet axis. Lower left: significance $S_2 = \delta/\sigma(\delta)$ distribution per event of the track with the second highest absolute significance for events with ≥ 2 selected tracks associated to the jet. Right: S_1 and S_2 distributions after subtracting the negatvie bins in the S_1 and S_2 distributions from the positive.

Fig. 9 shows a summary of the data/theory comparison for all HERA beauty results as a function of Q^2. For the measurements sensitive to b quarks with $p_T^b \sim m_b$ or lower (black points) there is a trend that the “massive” NLO QCD predictions [7] tend to underestimate the b production rate at very low Q^2. Depending on the chosen set of structure functions and parameters, a “mixed” prediction (VFNS) [16, 17] might describe the data better. For the higher p_T measurements (red/grey points), no clear trend is observed. Note that theoretical errors, which are typically of order 30%, are not shown. Fig. 10 shows a similar compilation for all HERA measurements in photoproduction ($Q^2 < 1$ GeV), now as a function of the b quark p_T. A similar trend is observed towards low p_T (but note that several measurements appear in both figures). Again, more precise measurements are needed to determine whether these trends can be confirmed.

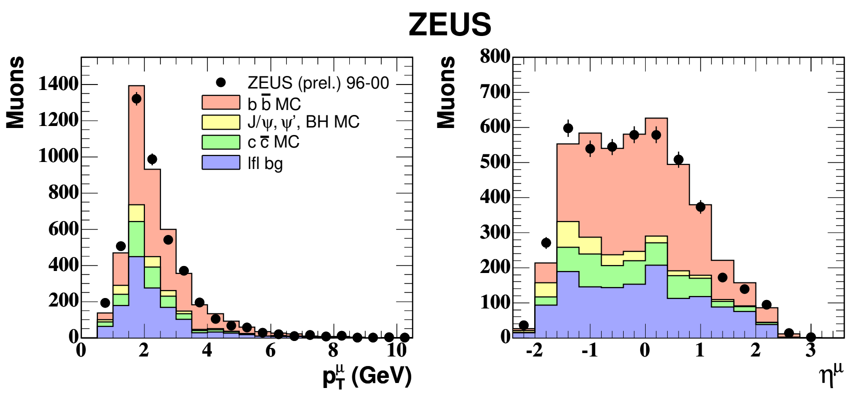
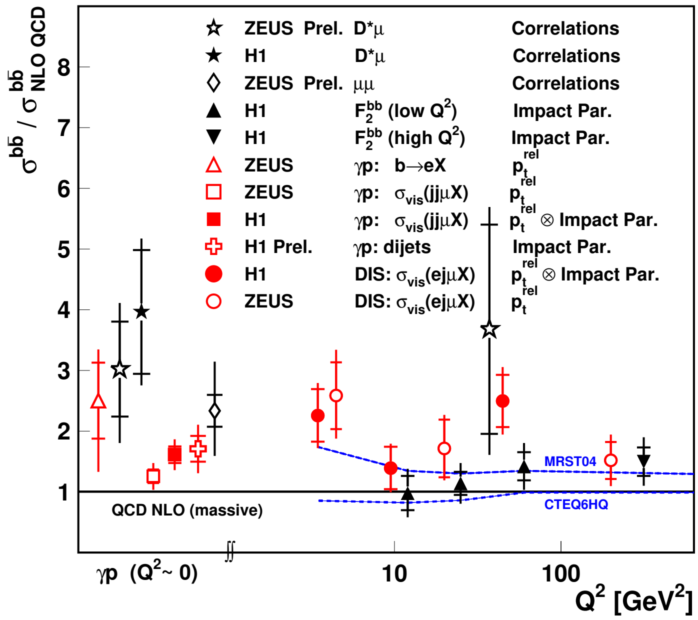


Fig. 8: Muon transverse momentum (a) and pseudorapity (b) distributions for nonisolated low transverse momentum muon pairs in ZEUS (two entries per event). The beauty contribution is dominant.



Key	Ref.	Signature	p_T cuts
	Photoproduction		
◊	[25]	$\mu\mu$	low
☆	[26]	$D^*\mu$	low
★	[27]	$D^*\mu$	low
△	[28]	2 jets+e	medium
□	[29]	2 jets+μ	medium
■	[30]	2 jets+μ	medium
✠	[23]	tracks	high
	DIS		
☆	[26]	$D^*\mu$	low
▲	[31]	tracks	low
▼	[24]	tracks	low
●	[30]	1 jet+μ	medium
○	[32]	1 jet+μ	medium

Fig. 9: Ratio of beauty production cross section measurements at HERA to NLO QCD predictions in the massive scheme as function of the photon virtuality Q^2. Measurements with low p_T cuts are shown in black, while measurements with medium or high p_T cuts are shown in red/grey. For more details see Table. The predictions from the VFNS NLO calculations by MRST and CTEQ for the DIS kinematic regime $Q^2 > 2$ GeV^2 are also shown (valid for comparison with the black low threshold points). Since theoretical errors are different for each point, they are not included in this plot.

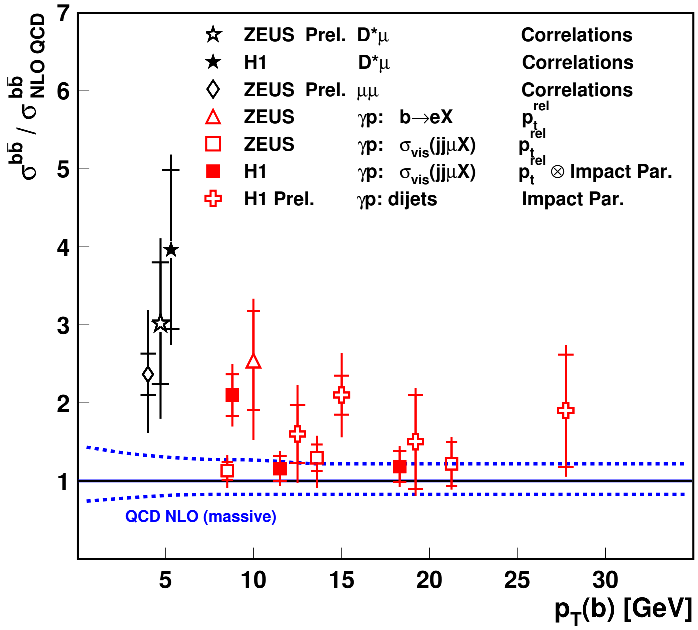


Fig. 10: Ratio of beauty production cross section measurements in photoproduction at HERA to NLO QCD predictions in the massive scheme as function of the transverse momentum of the b quark p_{Tb}. The dashed line gives an indication of the size of the theoretical uncertianties.

4 Quarkonium production

Inelastic heavy quarkonia, like open charm and beauty production, are produced at HERA via the process of photon-gluon fusion. The two charm or beauty quarks hadronize to form a charmonium or bottomonium state.

A number of models have been suggested to describe inelastic quarkonium production in the framework of perturbative QCD, such as the color-singlet model (CSM) [33, 34], the color-evaporation model [35, 36] and soft color interactions [37]. Comprehensive reports on the physics of charmonium production are available [38, 39].

Recently the ansatz of non-relativistic quantum chromodynamics (NRQCD) factorization was introduced. In the NRQCD approach non-perturbative effects associated with the binding of a $q\bar{q}$ pair into a quarkonium are factored into NRQCD matrix elements that scale in a definite manner with the typical relative velocity v of the heavy quark in the quarkonium. This way, colour octet quark anti-quark states, carrying different angular momenta and color charges than the quarkonium, can contribute to the charmonium production cross section. Theoretical calculations based on the NRQCD factorization approach [40, 41, 42] are available in leading order [43, 44, 45, 46, 47, 48]. In the NRQCD factorization approach the size of the color octet contributions, which are described by long distance matrix elements (LDME), are additional free parameters and have been determined in fits to the Tevatron data [49]. The NRQCD factorization approach incorporates the color singlet model i.e. the state $q\bar{q}[1,^3S_1]$ which is recovered in the limit in which the long distance matrix elements for other $q\bar{q}$ states tend to zero.

At HERA, cross sections measurements for photoproduction of J/ψ and $\psi(2S)$ and for electroproduction of J/ψ mesons have been performed [52, 53, 54, 55]. Bottomonium data are not available due to statistical limitations of the data.

For J/ψ and $\psi(2S)$ photoproduction, calculations of the color-singlet contribution are available to next-to-leading order perturbation theory [50, 51]. Calculations which include the color octet contributions as predicted by NRQCD are available in leading order.

Figure 11 shows the measurements of the J/ψ photoproduction cross section by the H1 collaboration [52] and the ZEUS collaboration [53] which are in good agreement with each other. The variable z (left figure) denotes the fraction of the photon energy in the proton rest frame that is transferred to the J/ψ. Reasonable agreement is found between the HERA data and the NRQCD factorization ansatz in leading order (LO, CS+CO). The uncertainty indicated by the open band is due to the uncertainty in the

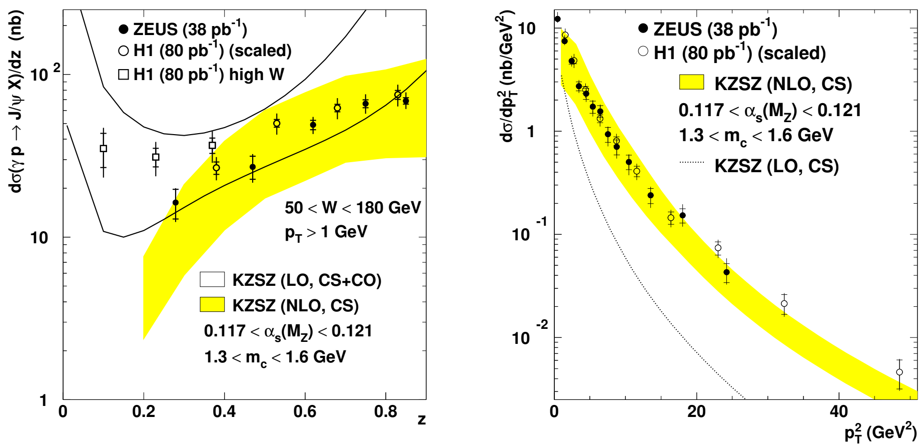


Fig. 11: Differential charmonium photoproduction cross sections as measured by H1 and ZEUS in comparison to calculations from LO NRQCD factorization (open band), NLO Color singlet contribution (shaded band) and LO color-singlet contribution (dotted line).

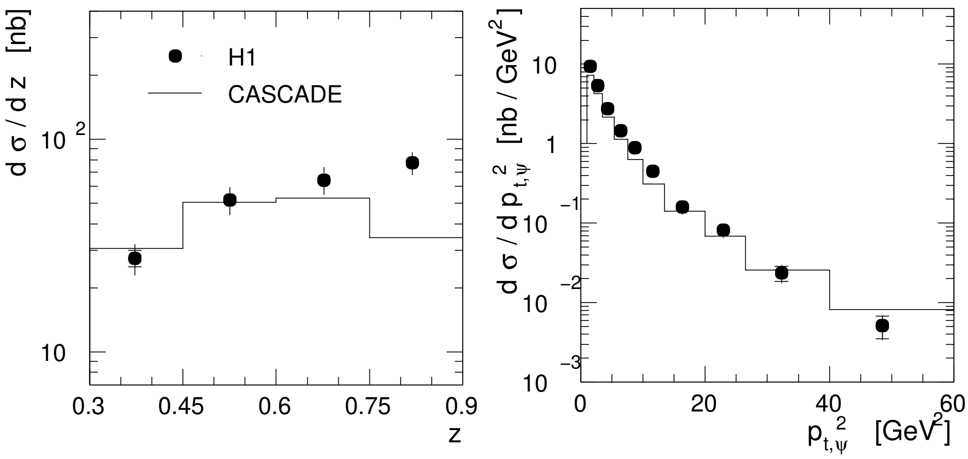


Fig. 12: Differential Charmonium photoproduction in comparison with a prediction using the color singlet-model and k_t factorization as implemented in the Monte Carlo generator CASCADE.

color-octet NRQCD matrix elements. In contrast, the shaded band shows the calculation of the color-singlet contribution (NLO, CS) which is performed to next-to-leading order in α_s [50, 51]. This NLO, CS contribution alone describes the data quite well without inclusion of color-octet contributions. Comparison between the NLO,CS prediction (shaded band) and the LO,CS prediction (dotted line) shows that the NLO corrections are crucial for the description of the HERA photoproduction data.

Charmonium production cross sections have also been calculated in the k_t factorization approach [56, 57, 58]. In these calculations the color-singlet model is used to describe the formation of the charmonium state. Figure 12 shows a comparison of the H1 data with the predictions from the k_t factorization approach as implemented in the Monte Carlo generator CASCADE [59]. Good agreement is observed between data and predictions for $z < 0.8$. At high z values, the CASCADE calculation underestimates the cross section. The CASCADE predictions for the the $p_{t,\psi}^2$ dependence of the cross section fit the data considerably better than the LO,CS calculation in the collinear factorization approach (dotted curve in fig. 11).

In fig 13 the differential cross sections for electroproduction of J/ψ mesons as measured by H1 [54] and ZEUS [55] are shown as a function of z and compared with predictions from the color singlet model (shaded band), with the NRQCD calculation [60] (CS+CO, open band), and also with calculations in the k_t factorization approach (dotted line) as provided by [58] and implemented in the Monte Carlo generator CASCADE (dash-dotted line).

In the left figure the data are seen to agree well with the predictions using the color singlet model (shaded band and lines) while the full NRQCD calculation (open band), including color-octet contributions is wrong in shape and normalization. The agreement deteriorates when the cut $p_{t,\psi}^{*2} > 1$ GeV is applied (right fig. 13). This cut is justified, however, as towards small $p_{t,\psi}^{*2}$ perturbation theory becomes increasingly unreliable due to collinear singularities for the contributions $e + g \to e + c\bar{c}[n] + g$ with n=${}^1S_0^{(8)}$ and ${}^3P_J^{(8)}$ [60].

In conclusion, NRQCD, as presently available in leading order, does not give a satisfactory description of the HERA data. In contrast, the color singlet model shows a reasonable description of the HERA data, when implemented in calculations to next-to-leading order perturbation theory or in calculations in which the k_t-factorization approach is used.

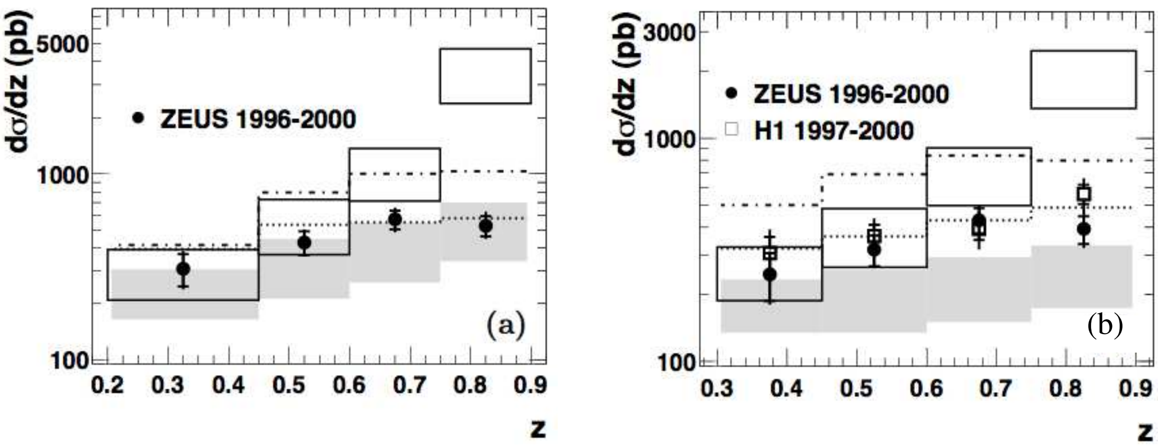


Fig. 13: Differential cross sections $d\sigma/dz$ a) without and b) with a cut on $p^{*2}_{t,\psi} > 1$ GeV. The data are compared to the NRQCD calculation (CS+CO, open band), the color-singlet contribution (CS, shaded band), with a prediction in the k_t factorization approach assuming the CSM (dotted line) and with the Monte Carlo generator CASCADE (dash-dotted line).

5 Charm and Beauty contributions to structure functions

To a good approximation, except at very high Q^2, the cross section for inclusive deep inelastic electron scattering off the proton at HERA can be described in terms of a single proton structure function F_2 (for formula see [1]). This structure function only depends on the photon virtuality, Q^2, and on the Bjorken scaling variable x. Assuming that the electron scatters off a single quark in the proton (0th order QCD, quark-parton model) x can be reinterpreted as the fraction of the proton momentum carried by the struck quark. This is a reasonable approximation for the light quark content of the proton.

For heavy quarks, the situation is a bit more complicated. Due to the heavy quark mass, on-shell heavy quarks can not exist within the proton. Rather, the dominant process for heavy quark production is the 1st order QCD BGF process depicted in Fig. 1. However, this process (and other higher order processes) still contributes to electron scattering, and hence to F_2. This can be interpreted in two ways.

In the massive approximation, heavy quarks are treated as being produced dynamically in the scattering process. The heavy quark contribution to F_2, frequently denoted as $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$, therefore indirectly measures the *gluon* content of the proton. If Q^2 is large enough such that the quark mass can be neglected ($Q^2 \gg (2m_Q)^2$), the splitting of the gluon into a heavy quark pair can be reinterpreted to occur *within* the proton. $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ then measure the occurrence of *virtual* heavy quark pairs in the proton, or the "heavy quark structure" of the proton.

For charm production, the condition $Q^2 \gg (2m_Q)^2$ is valid for a large part of the HERA phase space. For beauty, it is only satisfied at very large Q^2. This is also the region most interesting for physics at the LHC.

Similar arguments hold for the heavy quark structure of the photon.

As an example, Fig. 14 [61, 62, 63, 24, 31] shows $F_2^{c\bar{c}}$ as measured by the ZEUS and H1 collaborations. A different representation of these results is shown in Fig. 6 of [1]. There, also $F_2^{b\bar{b}}$ is shown. Good agreement is observed with QCD predictions. Parametrizations of heavy quark densities of the proton at LHC energies should therefore be valid within their respective errors.

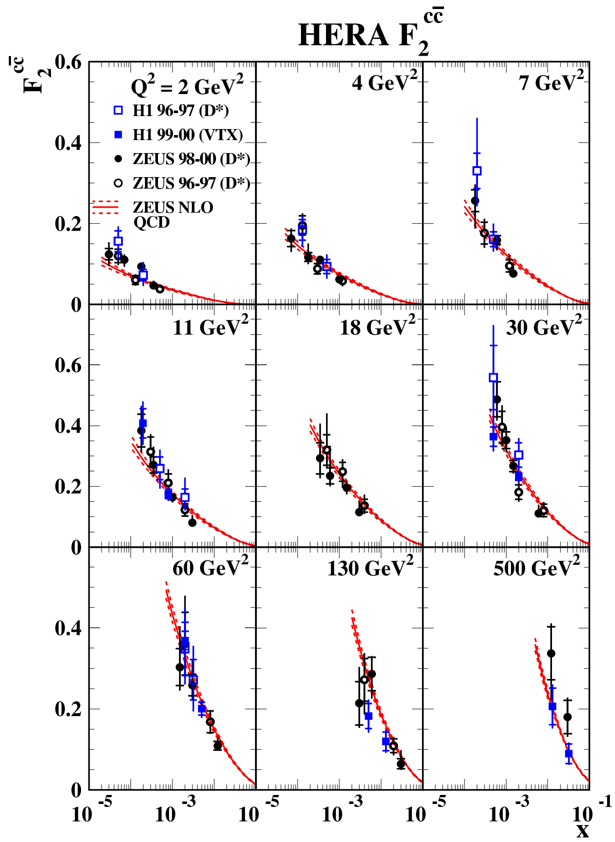


Fig. 14: $F_2^{c\bar{c}}$ results as a function of x in bins of Q^2, from the H1 and ZEUS $D^{*\pm}$ analyses and from the H1 inclusive lifetime tagging measurements The data are compared to a NLO prediction using the ZEUS NLO fit results for the proton parton densities.

6 Charm fragmentation

The large cross section for charm production at HERA allows measurements of charm fragmentation which are very competitive with e^+e^- measurements. As this topic is covered very nicely in [1] and [2] it is not treated further here.

7 Quark-antiquark correlations

Heavy quarks are always produced in pairs. An interesting way to check QCD is thus to verify whether the kinematic correlations between the quark pair are correctly described by QCD.

Fig. 15 shows different interpretations of the same higher order beauty production process. These different interpretations partially manifest themselves in different kinematic regions of beauty production phase space. If the highest virtuality part of the process occurs in the leading order BGF-like subprocess (left), the two b quarks will be almost back-to-back in the detector transverse plane. The two extra gluons can either be reabsorbed into the proton structure, recovering the original BGF graph, or manifest themselves as visible "parton shower" activity in the direction of the proton. Alternatively, if the dominant leading order subprocess is gluon exchange with one of the b quarks (right, "flavour excitation in the photon"), this b quark will recoil against a gluon jet. At sufficiently large momentum transfer (rare at HERA), the second b quark can be treated as a "spectator", and will approximately follow the initial photon direction. At next-to-leading order, contributions to both processes are described by the same Feynman graph, but the two extreme kinematic cases (and all variants in between) are still included. If both heavy quarks are tagged, these different kinematic regions can be distinguished by measuring the momentum and angular correlations between the two quarks.

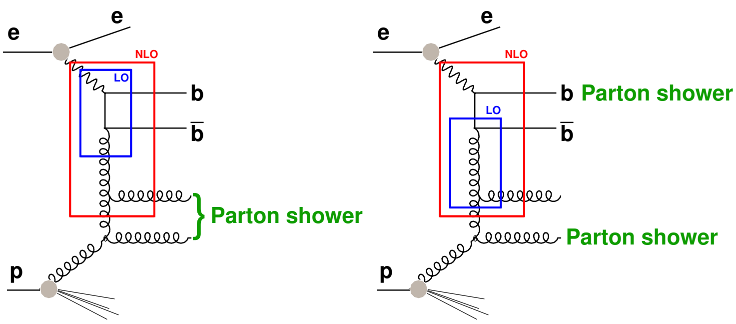


Fig. 15: Example for higher order Feynman graph for beauty production. Different interpretations of the graph in terms of NLO or LO matrix elements plus parton showers are highlighted. Depending on the kinematics and the scheme chosen, part of the gluons could also be reabsorbed into the proton structure function definition, and/or the $\gamma b\bar{b}$ vertex could be interpreted as part of the photon structure.

Fig. 16 [25] shows the angular correlations between the two muons originating from different b quarks of a $b\bar{b}$ pair. Reasonable agreement is observed with QCD predictions. The predominantly back-to-back topology comfirms the dominance of the BGF-like contribution.

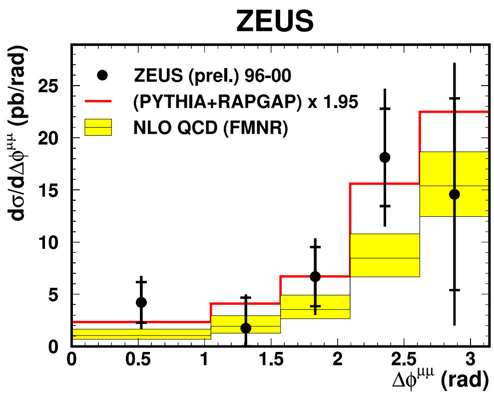


Fig. 16: Differential cross section $d\sigma/d\Delta\phi^{\mu\mu}$ for dimuon events from $b\bar{b}$ decays in which each muon originates from a different $b(\bar{b})$ quark. The data (solid dots) are compared to the leading order + parton shower generators PYTHIA and RAPGAP (histogram) and to massive NLO QCD predictions (shaded band).

8 HERA II prospects

Both the HERA collider and its detectors have been upgraded in 2001/2 to provide more luminosity with polarized electron beams, and improve heavy flavour detection. This program is called HERA II. The luminosity accumulated so far already exceeds the HERA I luminosity. An integrated luminosity up to 700 pb^{-1} is expected at the end of the HERA program in 2007. This enhances the statistics for many studies by almost an order of magnitude with respect to HERA I. The improved detectors offer further handles for improved heavy flavour measurements. H1 has improved the forward coverage of its Micro-Vertex-Detector [68], and added a Fast Track Trigger [67]. ZEUS has implemented a Micro-Vertex-Detector (MVD) [65] for the first time for HERA II, and has added an upgraded forward tracking detector [66]. These improvements allow the application of measurement techniques which could not be used at HERA I, and can be used to improve the data quality, add additional statistics, and/or cover new phase space regions.

New detectors require time to fully understand their systematics, but first preliminary results have already been obtained. Fig. 17 [64] shows the cross section for beauty production obtained using the new

ZEUS MVD with the first 33 pb^{-1} of HERA II data, compared to the HERA I result. Good agreement is observed.

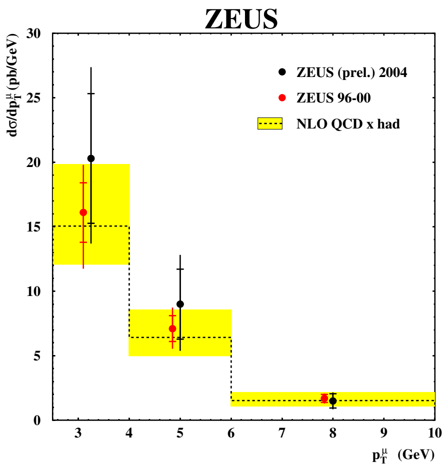


Fig. 17: Differential cross section as function of muon p_T for dimuon + jet events in photoproduction. Preliminary results from the first 33 pb^{-1} of HERA II data are compared to HERA I results and QCD predictions.

The measurements which will profit most from the improved HERA II data sets include double differential measurements such as the beauty and charm contributions to the proton structure function F_2, and multi-tag measurements to explicitly study quark-quark correlations. Statistical improvements of at least one order of magnitude can be expected when the increased luminosity and improved measurement techniques are combined.

9 Conclusions

Heavy flavour production at HERA is a very active field of research yielding multiple insights into the applicability of perturbative QCD. The problem of multiple scales complicates the perturbative expansions and limits the achievable theoretical precision. In general, QCD predictions agree well with the data, although indications for deviations persist in specific regions of phase space. Some of these might be attributable to missing NNLO or even higher order contributions.

The overall reasonable agreement, as well as the self-consistency of the structure functions tested by or derived from heavy flavour production at HERA, enhances confidence in corresponding cross-section predictions at LHC, within their respective theoretical uncertainties.

References

[1] M. Wing, these proceedings, and hep-ex/0508038.

[2] J. Bracinik et al., these proceedings.

[3] J. Dainese et al., these proceedings.

[4] O. Behnke et al., these proceedings.

[5] M. Cacciari et al., these proceedings.

[6] S. Frixione, P. Nason and G. Ridolfi, Nucl. Phys. B **454** (1995) 3 [hep-ph/9506226].

[7] B. W. Harris and J. Smith, Nucl. Phys. B **452** (1995) 109 [hep-ph/9503484].

[8] J. Binnewies, B. A. Kniehl and G. Kramer, Z. Phys. C **76** (1997) 677 [hep-ph/9702408].

[9] B. A. Kniehl, G. Kramer and M. Spira, Z. Phys. C **76** (1997) 689 [hep-ph/9610267].

[10] J. Binnewies, B.A. Kniehl and G. Kramer, Phys. Rev. D **58** (1998) 014014 [hep-ph/9712482].

[11] G. Kramer, Proceedings of the Ringberg Workshop New Trends in HERA Physics 1999, eds. G. Grindhammer, B. A. Kniehl and G. Kramer, Lecture Notes in Physics 546, Springer, 2000, p. 275.

[12] B. A. Kniehl, in 14th Topical Conference on Hadron Collider Physics (eds. M. Erdmann, T. Müller), p. 161 (Springer, Heidelberg, 2003), hep-ph/0211008.

[13] G. Kramer and H. Spiesberger, Eur. Phys. J. C **38** (2004) 309 [hep-ph/0311062].

[14] B. A. Kniehl, G. Kramer, I. Schienbein and H. Spiesberger, arXiv:hep-ph/0502194.

[15] M. Cacciari, S. Frixione and P. Nason, JHEP **0103** (2001) 006 [hep-ph/0102134].

[16] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. C **39** (2005) 155 [hep-ph/0411040].

[17] S. Kretzer, H. L. Lai, F. I. Olness and W. K. Tung, Phys. Rev. D **69** (2004) 114005 [hep-ph/0307022].

[18] S. Frixione and B. R. Webber, JHEP **0206** (2002) 029 [hep-ph/0204244].

[19] S. Frixione, P. Nason and B. R. Webber, JHEP **0308** (2003) 007 [hep-ph/0305252].

[20] S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **38**, 29 (2004) [arXiv:hep-ex/0409033].

[21] ZEUS Collaboration, Submitted to 31st International Conference on High Energy Physics, ICHEP02, 2002, Amsterdam, Abstract 786.

[22] A. Aktas *et al.* [H1 Collaboration], Eur. Phys. J. C **38** (2005) 447 [hep-ex/0408149].

[23] H1 Collaboration, contributed paper 405, XXII International Symposium on Lepton-Photon Interactions at High Energy, Uppsala, Sweden, 2005.

[24] A. Aktas *et al.* [H1 Collaboration], to appear in Eur. Phys. J. C., [hep-ex/0411046].

[25] ZEUS Collaboration, contributed paper 269, XXII International Symposium on Lepton-Photon Interactions at High Energy, Uppsala, Sweden, 2005.

[26] S. Chekanov *et al.* [ZEUS Collaboration], contributed paper 575, International Europhysics Conference on High Energy Physics (EPS 2003), Aachen, Germany, 2003.

[27] A. Aktas *et al.* [H1 Collaboration], submitted to Phys. Lett. B, 03/05, hep-ex/0503038.

[28] J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **18** (2001) 625 [hep-ex/0011081].

[29] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Rev. D **70** (2004) 012008 [hep-ex/0312057].

[30] A. Aktas *et al.* [H1 Collaboration], arXiv:hep-ex/0502010.

[31] A. Aktas *et al.* [H1 Collaboration], [hep-ex/0507081].

[32] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **599** (2004) 173 [hep-ex/0405069].

[33] E. L. Berger and D. L. Jones, Phys. Rev. D **23** (1981) 1521.

[34] R. Baier and R. Rückl, Z. Phys. C **19** (1983) 251.

[35] F. Halzen, Phys. Lett. B **69** (1977) 105.

[36] O. J. P. Eboli, E. M. Gregores and F. Halzen, Phys. Lett. B **451** (1999) 241 [hep-ph/9802421].

[37] A. Edin, G. Ingelman and J. Rathsman, Phys. Rev. D **56** (1997) 7317 [hep-ph/9705311].

[38] M. Krämer, Prog. Part. Nucl. Phys. **47** (2001) 141 [hep-ph/0106120].

[39] N. Brambilla *et al.*, hep-ph/0412158.

[40] W. E. Caswell and G. P. Lepage, Phys. Lett. B **167** (1986) 437.

[41] B. A. Thacker and G. P. Lepage, Phys. Rev. D **43** (1991) 196.

[42] G. T. Bodwin, E. Braaten and G. P. Lepage, Phys. Rev. D **51** (1995) 1125 [Erratum-ibid. D **55** (1997) 5853] [hep-ph/9407339].

[43] M. Cacciari and M. Krämer, Phys. Rev. Lett. **76** (1996) 4128 [hep-ph/9601276].

[44] M. Beneke, M. Krämer, and M. Vanttinen, Phys. Rev. D **57** (1998) 4258 [hep-ph/9709376].

[45] J. Amundson, S. Fleming and I. Maksymyk, Phys. Rev. D **56** (1997) 5844 [hep-ph/9601298].

[46] R. M. Godbole, D. P. Roy, and K. Sridhar, Phys. Lett. B **373** (1996) 328 [hep-ph/9511433].

[47] B. A. Kniehl and G. Kramer, Phys. Lett. B **413** (1997) 416 [hep-ph/9703280].

[48] B. A. Kniehl and G. Kramer, Phys. Rev. D **56** (1997) 5820 [hep-ph/9706369].

[49] E. Braaten, B. A. Kniehl and J. Lee, Phys. Rev. D **62** (2000) 094005 [hep-ph/9911436].

[50] M. Krämer, J. Zunft, J. Steegborn, and P. M. Zerwas, Phys. Lett. B **348** (1995) 657 [hep-ph/9411372].

[51] M. Krämer, Nucl. Phys. B **459** (1996) 3 [hep-ph/9508409].

[52] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **25** (2002) 25 [hep-ex/0205064].

[53] S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **27** (2003) 173 [hep-ex/0211011].

[54] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **25** (2002) 41 [hep-ex/0205065].

[55] S. Chekanov *et al.* [ZEUS Collaboration], hep-ex/0505008.

[56] V. A. Saleev and N. P. Zotov, Mod. Phys. Lett. A **9** (1994) 151 [Erratum-ibid. A **9** (1994) 1517].

[57] S. P. Baranov, Phys. Lett. B **428** (1998) 377.

[58] A. V. Lipatov and N. P. Zotov, Eur. Phys. J. C **27** (2003) 87 [hep-ph/0210310].

[59] H. Jung and G.P. Salam, Eur. Phys. J. C **19** (2001) 351 [hep-ph/0012143];
H. Jung, Comput. Phys. Commun. **143** (2002) 100 [hep-ph/0109102].

[60] B. A. Kniehl and L. Zwirner, Nucl. Phys. **B621** (2002) 337 [hep-ph/0112199].

[61] J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **12** (2000) 35 [hep-ex/9908012].

[62] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Rev. D **69**, 012004 (2004) [hep-ex/0308068].

[63] C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **528** (2002) 199 [hep-ex/0108039].

[64] ZEUS Collaboration, Contributed paper 359, XXII International Symposium on Lepton-Photon Interactions at High Energy, Uppsala, Sweden, 2005.

[65] E.N. Koffeman et al., Nucl. Instrum. Methods **A 309** (2001) 77.

[66] S. Goers [ZEUS-STT Collaboration], *The Straw-Tube Tracker of the ZEUS-Detector at HERA*, Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Lake Como, 2004.

[67] H1 Forward Silicon Tracker proposal, DESY-PRC 99/01 (1999).

[68] A. Baird *et al.*, IEEE Trans. Nucl. Sci. **48** (2001) 1276 [hep-ex/0104010].

Experimental aspects of heavy flavour production at the LHC

J. Baines[a], *A. Dainese*[b], *Th. Lagouri*[c], *A. Morsch*[d], *R. Ranieri*[e], *H. Ruiz*[d], *M. Smizanska*[f], *and C. Weiser*[g]
[a] Rutherford Laboratory, UK
[b] University and INFN, Padova, Italy
[c] Institute of Nuclear and Particle Physics, Charles University, Prague, Czech Republic
[d] CERN, Geneva, Switzerland
[e] University and INFN, Firenze, Italy
[f] Lancaster University, Lancaster, UK
[g] Institut für Experimentelle Kernphysik, Universität Karlsruhe, Karlsruhe, Germany

Abstract
We review selected aspects of the experimental techniques being prepared to study heavy flavour production in the four LHC experiments (ALICE, ATLAS, CMS and LHCb) and we present the expected performance for some of the most significative measurements.

coordinators: A. Dainese, M. Smizanska, and C. Weiser

1 Introduction

author: M. Smizanska

Unprecedently large cross sections are expected for heavy-flavour production in proton–proton collisions at the LHC energy of $\sqrt{s} = 14~\mathrm{TeV}$. Next-to-leading order perturbative QCD calculations predict values of about 10 mb for charm and 0.5 mb for beauty, with a theoretical uncertainty of a factor 2–3. Despite these large cross sections, the LHC experiments, ALICE [1, 2], ATLAS [3], CMS [4], and LHCb [5], will have to deal with rejection of background coming from non-heavy flavour inelastic interactions for which the predicted cross is about 70 mb. The four experiments will work at different luminosity conditions. ATLAS and CMS are designed to work in a wide range of luminosities up to nominal $10^{34}~\mathrm{cm^{-2}s^{-1}}$, while the LHCb optimal luminosity will vary in the range $(2\text{–}5) \times 10^{32}~\mathrm{cm^{-2}s^{-1}}$ and ALICE is designed to work at $3 \times 10^{30}~\mathrm{cm^{-2}s^{-1}}$ in proton–proton collisions. Luminosity conditions in ATLAS, CMS and LHCb allow multiple interactions per bunch crossing, thus leading to requirements of even stronger identification and selection of heavy-flavour events already at trigger level. The first task will be the measurement of integrated and differential charm and beauty production cross sections in the new energy domain covered at the LHC. ALICE will play an important role, having acceptance down to very low transverse momentum, as we discuss in Section 4. These measurements can be performed within a relatively short period of running. Afterwards, the heavy-flavour studies will focus on less inclusive measurements addressing specific production mechanisms that allow to test higher order perturbative QCD predictions, as well as on rare decays of heavy-flavour hadrons, that may carry information on New Physics beyond the Standard Model. In order to meet these requirements dedicated and sophisticated trigger strategies have been prepared by the LHC experiments.

2 Heavy flavour detection in the LHC experiments

author: A. Dainese

The four detectors that will take data at the LHC have different features and design requirements, but all of them are expected to have excellent capabilities for heavy-flavour measurements. Their complementarity will provide a very broad coverage in terms of phase-space, decay channels and observables.

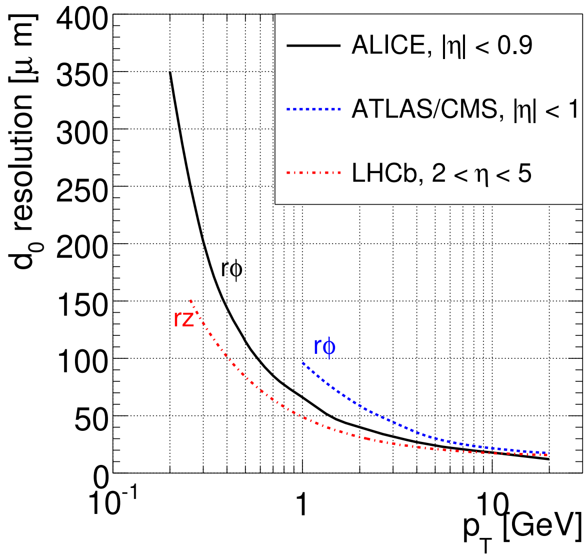


Fig. 1: Track impact parameter resolutions for the four LHC experiments. Note that for ALICE, ATLAS and CMS the impact parameter is defined in the $r\phi$ plane, while for LHCb it is defined in the rz plane.

Experimentally, the key elements for a rich heavy-flavour program are track and vertex reconstruction and particle identification (PID).

Open charm and beauty mesons have typical life-times in the order of a ps ($c\tau$ values are about 125–300 μm for D mesons and 500 μm for B mesons) and the most direct detection strategy is the identification of single tracks or vertices that are displaced from the interaction vertex. The detector capability to perform this task can be characterized by the transverse impact parameter[1] (d_0) resolution. All experiments will be equipped with high position-resolution silicon-detector layers, including pixel detector for the innermost layers, for precise tracking and impact parameter measurement. Tracking is done in the central (pseudo)rapidity region for ALICE ($|\eta| < 0.9$), ATLAS and CMS ($|\eta| \lesssim 2.5$), and in the forward region for LHCb ($2 \lesssim \eta \lesssim 5$). In Fig. 1 we show the d_0 resolution, which is similar for the different experiments, and better than 50 μm for $p_T \gtrsim 1.5$–3 GeV. The inner detector systems of ATLAS, CMS and ALICE will operate in different magnetic fields: The ALICE magnetic field will vary within low values (0.2–0.5 T) leading to a very low p_T cutoff of 0.1–0.2 GeV, while ATLAS (2 T) and CMS (4 T) have higher cutoffs of 0.5 and 1 GeV, respectively, but better p_T resolution at high p_T (e.g., at $p_T = 100$ GeV, $\delta p_T/p_T \approx 1$–2% for ATLAS/CMS at central rapidity and $\approx 9\%$ for ALICE).

Both lepton and hadron identification are important for heavy-flavour detection. D and B mesons have relatively large branching ratios (BR) in the semi-leptonic channels, $\simeq 10\%$ to electrons and $\simeq 10\%$ to muons, and inclusive cross-section measurements can be performed via single leptons or di-leptons. Alternatively, high-p_T leptons can be used as trigger-level tags to select $\mathrm{B} \to \mathrm{J}/\psi + X$ candidate events, that provide more accurate cross section measurements. ALICE can identify electrons with $p_T > 1$ GeV and $|\eta| < 0.9$, via transition radiation and $\mathrm{d}E/\mathrm{d}x$ measurements, and muons in the forward region, $2.5 < \eta < 4$, which allows a very low p_T cutoff of 1 GeV. CMS and ATLAS have a broad pseudorapidity coverage for muons, $|\eta| < 2.4$ and $|\eta| < 2.7$, respectively, but they have a higher p_T cutoff varying between 4 and 6 GeV, depending on η. Both CMS and ATLAS have high-resolution electro-magnetic calorimeters that will be used to identify electrons. Semi-leptonic inclusive measurements do not provide direct information on the D(B)-meson p_T distribution, especially at low p_T, because of the weak corre-

[1] We define as impact parameter the distance of closest approach to the interaction vertex of the track projection in the plane transverse to the beam direction.

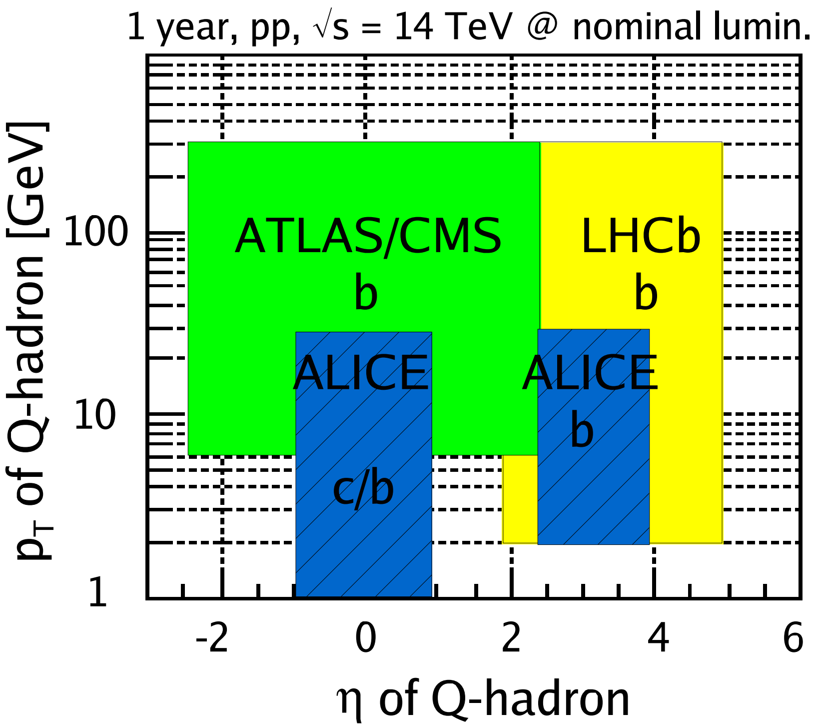


Fig. 2: Schematic acceptances in transverse momentum and pseudorapidity for open heavy flavour hadrons (indicated as 'Q-hadrons') in the four LHC experiments. The high-p_T coverages correspond to one year (i.e. 7 months) of running at nominal luminosity (see beginning of this section).

lation between the lepton and the meson momenta. Therefore, for charm in particular, the reconstruction of exclusive (hadronic) decays is preferable. In this case, hadron identification allows a more effective rejection of the combinatorial background in the low-p_T region. ALICE disposes of $\pi/\mathrm{K}/\mathrm{p}$ separation via $\mathrm{d}E/\mathrm{d}x$ and time-of-flight measurements for $p < 3$–4 GeV and $|\eta| < 0.9$.

Figure 2 shows schematically the p_T vs. η acceptances for charm (c) and beauty (b) hadrons in the four experiments, as expected for one year of running at nominal luminosity (note that the value of the luminosity is different for each experiment, as previously discussed). ATLAS and CMS have similar acceptances for beauty measurements; the minimum accessible p_T is relatively large because of the strong magnetic fields, which in turn, together with the high luminosity, allow to cover transverse momenta up to 200–300 GeV. The acceptance of LHCb, although centred at forward rapidity, has a significant overlap, with those of ATLAS and CMS. The acceptance of ALICE for beauty overlaps with ATLAS and CMS at central rapidity and with LHCb at forward rapidity. The moderate magnetic field allows measurements down to transverse momenta of about 2 GeV for B mesons in the forward muon arm and in the barrel, and down to about 1 GeV for D mesons in the barrel.

3 Beauty triggers at the LHC

3.1 ATLAS beauty trigger

author: J. Baines

The ATLAS trigger consists of three levels [6]. Level-1 is implemented in hardware, whilst the higher level triggers (level-2 and the Event Filter, EF) are based on general-purpose processors. The level-1 triggers are based on information from the calorimeter and muon trigger chambers. At higher trigger levels, information from the Inner Detector (ID) and precision muon detector is included. The size of the level-2 and EF processor farms is limited, which in turn limits the amount of data processing that can be performed in the trigger. The B-trigger must, therefore, have the flexibility to adapt selections both as the luminosity falls during a beam-coast and, over a longer time-scale, as the peak luminosity of the LHC increases. This is achieved by using a di-muon trigger at the start of higher luminosity fills and introducing additional triggers for lower luminosity fills or as the luminosity falls during a beam coast [7].

A di-muon trigger provides a very effective selection for a range of important channels, e.g. $\mathrm{B_d^0} \to \mathrm{J}/\psi(\mu^+\mu^-)\mathrm{K_s^0}$, $\mathrm{B} \to \mathrm{K}^{0\star}\mu\mu$ and $\mathrm{B} \to \rho^0\mu\mu$. The Level-1 muon trigger is efficient down to a p_T of about 5 GeV in the barrel region and about 3 GeV in the end-caps. However the actual thresholds used for the di-muon trigger will be determined by rate limitations. For example, a p_T threshold of 6 GeV would give a di-muon trigger rate of about 600 Hz after level-1 at a luminosity of 2×10^{33} cm^{-2}s^{-1}. These triggers are mostly due to muons from heavy flavour decays plus some single muons which are doubly counted due to overlaps in the end-cap trigger chambers. The later are removed when the muons are confirmed at level-2 using muon precision chambers and ID information from inside the level-1 Region of Interest (RoI). At the EF tracks are refit, inside regions identified by level-2, and specific selections made on the basis of mass and decay length cuts. These consist of semi-inclusive selections, for example to select $\mathrm{J}/\psi(\mu^+\mu^-)$ decays with a displaced vertex, and in some cases exclusive selections such as for $\mathrm{B} \to \mu^+\mu^-$. The final trigger rate, after the EF, is about 20 Hz at a luminosity of 2×10^{33} cm^{-2}s^{-1}.

At lower luminosities, additional triggers are introduced based on a single muon trigger ($p_T \gtrsim 8$ GeV) together with a calorimeter trigger. The calorimeter trigger identifies clusters of energy deposition in the electromagnetic and hadronic calorimeter consistent with an electron or photon (EM RoI) or a jet (Jet RoI). For hadronic final states, such as $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ and $\mathrm{B_s^0} \to \mathrm{D_s^-a_1^+}$ track are reconstructed in the Inner Detector in RoI of about $\Delta\eta \times \Delta\phi = 1.0 \times 1.5$. By limiting track reconstruction to the part of the ID lying within the RoI, about 10% on average, there is potential for up to a factor of ten saving in execution time compared to reconstruction in the full Inner Detector. Preliminary studies of efficiency and jet-cluster multiplicity have been made using a fast simulation which includes a detailed parameterization of the calorimeter. These studies indicate that a threshold on the jet cluster energy of $E_T > 5$ GeV gives a reasonable multiplicity, i.e. a mean of about two RoI per event for events containing a muon trigger. This threshold would give a trigger that is efficient for $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ events with a B-hadron p_T above about 15 GeV.

Track reconstruction inside e/gamma RoI can be used to select channels such as $\mathrm{B_d^0} \to \mathrm{J}/\psi(e^+e^-)\mathrm{K_s^0}$, $\mathrm{B_d} \to \mathrm{K}^{0\star}\gamma$and $\mathrm{B_s} \to \phi\gamma$. Preliminary studies show that a reasonable compromise between RoI multiplicity and electron efficiency might be obtained with a cluster energy threshold of $E_T > 2$ GeV. This gives a mean RoI multiplicity of about one for events containing a muon trigger and is efficient for channels containing an electron with $p_T > 5$ GeV. Following the ID track reconstruction further selections are made for specific channels of interest. These are kept as inclusive as possible at level-2 with some more exclusive selections at the EF.

In LHC running, there will be competing demands for resources in the level-2 and EF trigger farms and for trigger band-width. By adopting a flexible strategy and making the maximum use of RoI information to guide reconstruction at level-2 and the EF, the B-physics coverage of ATLAS can be maximized.

3.2 CMS beauty trigger

author: R. Ranieri

The Large Hadron Collider (LHC) will provide 40 MHz proton-proton collisions at the centre of mass energy of 14 TeV. At the beginning a luminosity of 2×10^{33} cm^{-2}s^{-1} is expected, corresponding to 20 fb^{-1} collected per year. Assuming the $\mathrm{b\overline{b}}$ production cross section to be 0.5 mb, 10^{13} b-physics events per year are foreseen: all kind of b-particles will be produced and studies will be performed not only in $\mathrm{B_d^0}$, but also in $\mathrm{B_s^0}$ meson system. A wide b-physics programme, including CP violation, $\mathrm{B_s^0}-\overline{\mathrm{B}}_\mathrm{s}^0$ mixing and rare decays can therefore be covered by the CMS experiment. The apparatus will be equipped with a very precise tracking system made with silicon microstrip and pixel detectors [8, 9].

The rate at which events can be archived for offline analyses is 100 Hz [10, 11]. The trigger thresholds are optimized for a wide physics discovery program with selection of high transverse momentum (p_T) processes. Low-p_T events, as required for b-physics , are selected mainly by the first level muon trigger, then an exclusive reconstruction of few relevant *benchmark* channels can separate interesting events from the background. The b-physics programme could evolve with time following both the theoretical developments and the results which will be obtained in the next years by b-factories and Tevatron experiments.

The lowest trigger level (Level-1) is based on the fast response of calorimeters and muon stations with coarse granularity. No information on secondary vertices is available, hence the Level-1 selection of b-physics events exploits the leptonic signatures from beauty hadron decays, therefore a single muon or a di-muon pair is required. The Level-1 output at start-up will be 50 kHz. Several studies have been done to optimize the trigger thresholds in order to have the possibility of selecting most of the interesting physics signatures. A total of 3.6 kHz rate is dedicated to the Level-1 muon selection. It is obtained by requiring a single muon with $p_T > 14$ GeV or at least two muons with $p_T > 3$ GeV.

A further selection is made during the High-Level trigger (HLT) by using also the information from the tracking system. The CMS High-Level trigger is entirely based on a CPU farm with some thousand CPUs. Each processor analyses a single event; in principle offline event reconstruction can be performed, but in order to reduce the processing time fast track reconstruction has to be done. Some algorithms will be dedicated to the fast reconstruction and identification of physics processes, thus allowing to start the offline analysis directly from the online selection. They have to fulfill the HLT time constraint, hence they have to be able to analyze and accept (or reject) data within the time limits imposed by the HLT latency. To lower the execution time, which is due mainly to the processing of tracking system signals, track reconstruction is preferably performed only in limited regions of the space (*regional track reconstruction*) and stopped when a certain precision is reached in the measurement of some track parameters, such as transverse momentum and impact parameter (*conditional track finding*). Invariant mass of b-hadrons can thus be measured online with good resolution, allowing to select the searched event topologies.

An additional trigger strategy, which relies on the possibility of lowering the trigger thresholds during the LHC beam coast or lower luminosity fills to collect more b-physics events is under study.

The rare decay $\mathrm{B^0_{s,d}} \to \mu^+\mu^-$ is triggered at Level-1 with 15.2% efficiency. At HLT, the two muons are required to be opposite charged and isolated, to come from a displaced common vertex and have an invariant mass within 150 MeV from the $\mathrm{B_s^0}$ mass. The estimated background rate is below 2 Hz and nearly 50 signal events are expected with 10 fb^{-1}.

The determination of Δm_s and $\Delta\Gamma_s$ will be a valuable input for flavour dynamics in the Standard Model and its possible extensions. The measurement of Δm_s is allowed by the $\mathrm{B_s^0} \to \mathrm{D_s}^-\pi^+$ decay followed by $\mathrm{D_s}^- \to \phi\pi^-$ and $\phi \to \mathrm{K^+K^-}$. The $\mathrm{B_s^0}$ CP state at decay time is tagged by the charge of the pion associated to the $\mathrm{D_s}$ (in this case the π^+). The only way to trigger on these hadronic events is to search for the muon coming from the decay of the other b quark in the event. In addition to the single muon Level-1 trigger, it was studied the possibility of a combined trigger with a low-p_T muon and a soft jet. The CMS High-Level trigger algorithm reconstructs the charged particle tracks with only three points by using the precise pixel detector. Topological and kinematical cuts are applied to reconstruct the three resonances ϕ, $\mathrm{D_s}$ and $\mathrm{B_s^0}$. A 20 Hz output rate is achieved with about 1000 signal events in 20 fb^{-1}. Since the overall possible rate on tape is 100 Hz, the bandwidth allocated to this channel probably could not exceed 5 Hz. If the fraction of events written to tape is scaled accordingly, more than 300 signal events are expected for 20 fb^{-1}. In order to fully cover the range allowed by the Standard Model, about

1000 events are needed.

The decay channel $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\phi$ is very important because it can not be studied with large accuracy before LHC and can reveal hints for physics beyond the Standard Model. Events with a couple of muons are passed to the HLT. The inclusive selection of $\mathrm{J}/\psi \rightarrow \mu^+\mu^-$ decays, obtained with mass requirements on the di-muon system, leads to a total of 15 Hz rate, 90% of which is made of J/ψ from b quarks. With an additional amount of CPU time, perhaps sustainable by the HLT computing power, about 170 000 events are expected in 20 fb^{-1} with less than 2 Hz rate.

3.3 LHCb beauty trigger

author: H. Ruiz

The LHCb detector [5] is optimized for exploiting the B-physics potential of LHC. Together with excellent vertexing and particle identification, an efficient trigger on a wide variety of B decays is one of the main design requirements of the experiment.

The LHCb trigger system [12] is organized in three levels. The first one (L0) runs on custom electronics and operates synchronously at 40 MHz, with a 4 μs latency. The remaining two trigger levels (L1 and HLT) run on a shared farm of 1400 commercial CPUs. A brief description of the three trigger levels and their performance follows.

L0 exploits the relatively high p_T of B decay products. High p_T candidates are identified both in the calorimeter and in the muon system, with p_T thresholds of about 3 and 1 GeV respectively. Complicated events that would consume unreasonable time at higher levels are promptly vetoed in two different ways. First, multiple primary vertex topologies are rejected by using two dedicated silicon layers of the vertex detector. Secondly, events with large multiplicity, measured at a scintillating pad layer, are vetoed. The input rate of events visible in the detector is about 10 MHz, with a $b\bar{b}$ content of 1.5%. L0 reduces this rate by a factor of 10 while increasing the $b\bar{b}$ content to 3%. The typical efficiency of L0 is 90% for channels with dimuons, 70% for radiative decays and 50% for purely hadronic decays.

At the 1 MHz input rate of L1 it becomes feasible to use tracking information, allowing for the search for B vertex displacement signatures. Tracks are first searched at the vertex detector, and then confirmed in two dedicated tracking layers (trigger tracker or TT) which provide a rough estimation of the momentum of the tracks ($\delta p_T/p_T \sim 25\%$). The generic L1 decision is based on the presence of two tracks with an impact parameter higher than 0.15 mm with a sufficiently high value of $log(p_{T1} + p_{T2})$. Alternative selection criteria are applied, based on the presence of tracks matched to L0 neutral calorimeter objects and muon candidates. The output rate of L1 is 40 kHz with a $b\bar{b}$ content of 15%. The efficiencies are at the level of 90, 80 and 70% for channels with di-muons, only hadrons and radiative decays respectively.

The HLT [13] consists of two sequential *layers*. The first one refines the L1 decision with the all the tracking information from the detector, improving the p_T measurement to the level of $\delta p_T/p_T \sim 1\%$. The rate is reduced to 13 kHz and the $b\bar{b}$ content is enriched to 30%. The second layer consists on a series of alternative selections. A first group of them aims for maximal efficiency on the base-line physics channels and the corresponding control samples, by making use of the complete reconstruction of the decay vertex and its kinematical properties. These selections fill 200 Hz of bandwidth, while providing efficiencies typically higher than 90%. The rest of selections aim for more generic signatures that will provide robustness and flexibility to the trigger system. In addition, the samples selected will be useful for calibration and systematic studies. The selections aim for generic J/ψ and D* (600 Hz and 300 Hz respectively), and generic B decays (900 Hz). The latter is based on the detection of single muons with high p_T and impact parameter.

In total, 2 kHz of events will be written on tape, with an expected overall efficiency ranging between 75% for channels with di-muons to 35% for purely hadronic final states.

4 Measurements in preparation at the LHC and expected performance

In the following we present, as examples, the expected performance for the detection of D and B mesons in ALICE[2], and for the study of $\mathrm{b\overline{b}}$ azimuthal correlation in ATLAS. We also include a summary of the capability of ALICE of quarkonia measurements (ψ family and Υ family).

4.1 Charm reconstruction in ALICE

author: A. Dainese

One of the most promising channels for open charm detection is the $\mathrm{D^0} \to \mathrm{K^-}\pi^+$ decay (and charge conjugate) that has a BR of 3.8%. The expected yields ($\mathrm{BR} \times \mathrm{d}N/\mathrm{d}y$ at $y = 0$), in pp collisions at $\sqrt{s} = 14$ TeV and in central Pb–Pb (0–5% σ^{tot}) at $\sqrt{s_{\mathrm{NN}}} = 5.5$ TeV are 7.5×10^{-4} and 5.3×10^{-1} per event, respectively [14].

The main feature of this decay topology is the presence of two tracks with impact parameters $d_0 \sim 100~\mu\mathrm{m}$. The detection strategy [15] to cope with the large combinatorial background from the underlying event is based on the selection of displaced-vertex topologies, i.e. two tracks with large impact parameters and good alignment between the $\mathrm{D^0}$ momentum and flight-line, and on invariant-mass analysis to extract the signal yield. This strategy was optimized separately for pp and Pb–Pb collisions, as a function of the $\mathrm{D^0}$ transverse momentum, and statistical and systematic errors were estimated [16, 17]. The results, in terms of p_T coverage and statistical precision, are found to be similar for the two colliding systems [16, 17].

Figure 3 (left) shows the expected sensitivity of ALICE for the measurement of the $\mathrm{D^0}$ p_T-differential cross section in pp collisions, along with NLO pQCD [18] calculation results corresponding to different choices of the charm quark mass and of renormalization and factorization scales. In the right-hand panel of the figure we present the ratio 'data/theory' ('default parameters/theory parameters') which better allows to compare the different p_T-shapes obtained by changing the input 'theory parameters' and to illustrate the expected sensitivity of the ALICE measurement. The estimated experimental errors are much smaller than the theoretical uncertainty band. We note that the data cover the region at low transverse momentum where the accuracy of the pQCD calculation becomes poorer and where novel effects, determined by the high partonic density of the initial state at LHC energies, may play an important role (see "Small-x effects in heavy quark production" section of this report).

4.2 Beauty production measurements in ALICE

author: A. Dainese

The expected yields ($\mathrm{BR} \times \mathrm{d}N/\mathrm{d}y$ at $y = 0$) for $\mathrm{B} \to e^\pm + X$ plus $\mathrm{B} \to \mathrm{D}\,(\to e^\pm + X) + X'$ in pp collisions at $\sqrt{s} = 14$ TeV and in central Pb–Pb (0–5% σ^{tot}) at $\sqrt{s_{\mathrm{NN}}} = 5.5$ TeV are 2.8×10^{-4} and 1.8×10^{-1} per event, respectively [14].

The main sources of background electrons are: (a) decays of D mesons; (b) neutral pion Dalitz decays $\pi^0 \to \gamma e^+ e^-$ and decays of light mesons (e.g. ρ and ω); (c) conversions of photons in the beam pipe or in the inner detector layers and (d) pions misidentified as electrons. Given that electrons from beauty have average impact parameter $d_0 \simeq 500~\mu\mathrm{m}$ and a hard momentum spectrum, it is possible to obtain a high-purity sample with a strategy that relies on: electron identification with a combined $\mathrm{d}E/\mathrm{d}x$ and transition radiation selection, which allows to reduce the pion contamination by a factor 10^4; impact parameter cut to reject misidentified pions and electrons from sources (b) and (c); transverse momentum cut to reject electrons from charm decays. As an example, with $d_0 > 200~\mu\mathrm{m}$ and $p_T > 2$ GeV, the expected statistics of electrons from B decays is 8×10^4 for 10^7 central Pb–Pb events, allowing the

[2] Given that ALICE is dedicated to the study of nucleus–nucleus collisions at the LHC, some of the presented results are relative to Pb–Pb collisions at $\sqrt{s} = 5.5$ TeV per nucleon–nucleon collisions. These results can be taken as lower limits for the performance in pp collisions, where the background contributions are much lower.

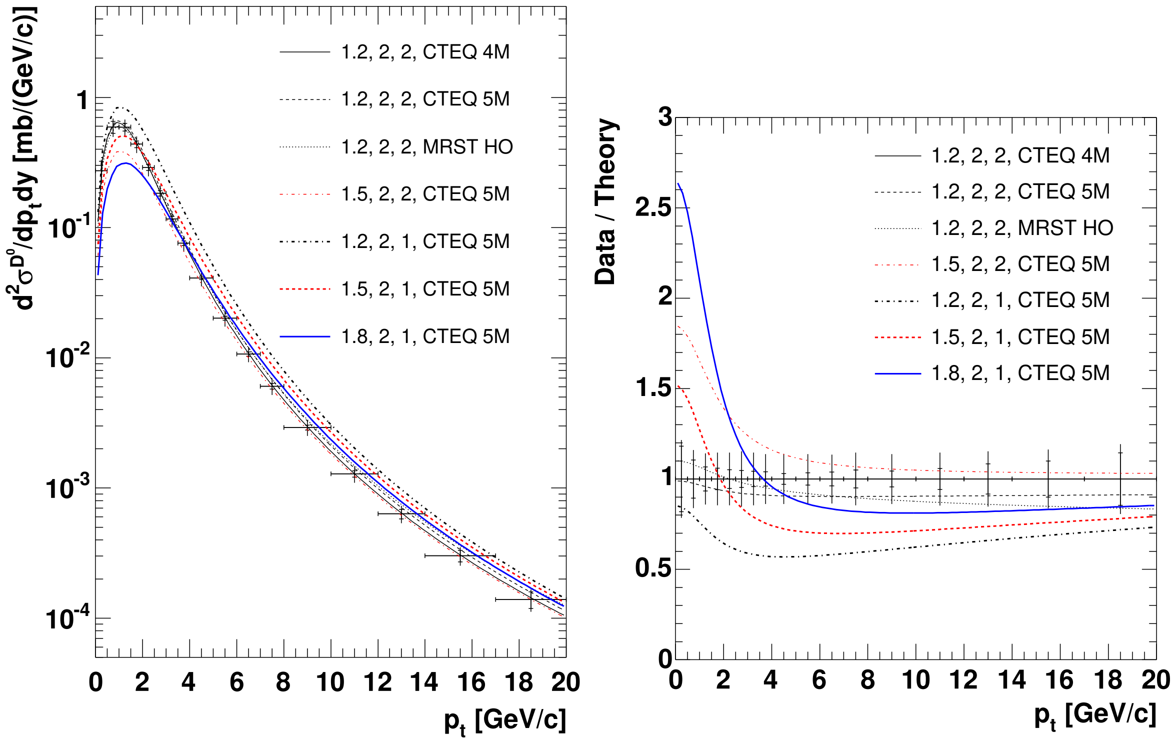


Fig. 3: ALICE sensitivity on $\mathrm{d}^2\sigma^{\mathrm{D}^0}/\mathrm{d}p_T\mathrm{d}y$, in pp at 14 TeV, compared to the pQCD predictions obtained with different sets of the input parameters m_c [GeV], μ_F/μ_0, μ_R/μ_0 and PDF set (μ_0 is defined in the text). The inner bars represent the statistical error, the outer bars the quadratic sum of statistical and p_T-dependent systematic errors. A normalization error of 5% is not shown. The panel on the right shows the corresponding 'data/theory' plot.

measurement of electron-level p_T-differential cross section in the range $2 < p_T < 18$ GeV. The residual contamination of about 10%, acculated in the low-p_T region, of electrons from prompt charm decays and from misidentified charged pions can be evaluated and subtracted using a Monte Carlo simulation tuned to reproduce the measured cross sections for pions and D^0 mesons. A Monte-Carlo-based procedure can then be used to compute, from the electron-level cross section, the B-level cross section $\mathrm{d}\sigma^{\mathrm{B}}(p_T > p_T^{\mathrm{min}})/\mathrm{d}y$ [17]. In the left-hand panel of Fig. 4 we show this cross section for central Pb–Pb collisions with the estimated uncertainties. The covered range is $2 < p_T^{\mathrm{min}} < 30$ GeV.

B production can be measured also in the ALICE forward muon spectrometer, $2.5 < \eta < 4$, analyzing the single-muon p_T distribution and the opposite-sign di-muons invariant mass distribution [17].

The main backgrounds to the 'beauty muon' signal are $\pi^\pm$, $\mathrm{K}^\pm$ and charm decays. The cut $p_T > 1.5$ GeV is applied to all reconstructed muons in order to increase the signal-to-background ratio. For the opposite-sign di-muons, the residual combinatorial background is subtracted using the technique of event-mixing and the resulting distribution is subdivided into two samples: the low-mass region, $M_{\mu^+\mu^-} < 5$ GeV, dominated by muons originating from a single b quark decay through $\mathrm{b} \to \mathrm{c}(\to \mu^+)\mu^-$ ($\mathrm{BD_{same}}$), and the high-mass region, $5 < M_{\mu^+\mu^-} < 20$ GeV, dominated by $\mathrm{b\overline{b}} \to \mu^+\mu^-$, with each muon coming from a different quark in the pair ($\mathrm{BB_{diff}}$). Both samples have a background from $\mathrm{c\overline{c}} \to \mu^+\mu^-$ and a fit is done to extract the charm- and beauty-component yields. The single-muon p_T distribution has three components with different slopes: K and π, charm, and beauty decays. Also in this case a fit technique allows to extract a p_T distribution of muons from B decays. A Monte Carlo procedure, similar to that used for semi-electronic decays, allows to extract B-level cross sections for the data sets (low-mass $\mu^+\mu^-$, high-mass $\mu^+\mu^-$, and p_T-binned single-muon distribution), each set covering

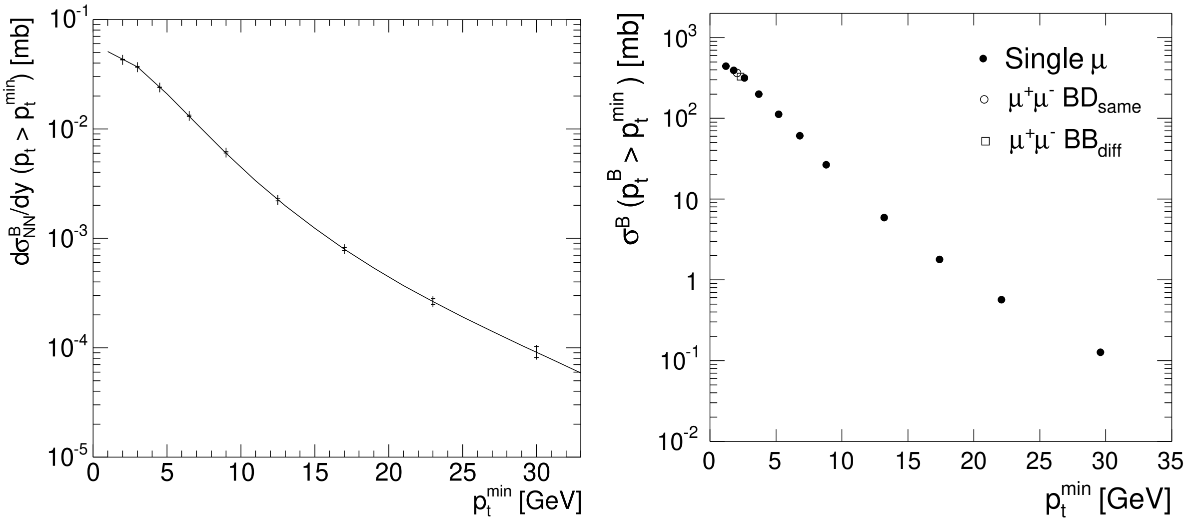


Fig. 4: B production cross section vs. $p_T^{\min}$ reconstructed by ALICE in 10^7 central Pb–Pb events. Left: $\mathrm{d}\sigma^{\mathrm{B}}_{\mathrm{NN}}/\mathrm{d}y$ at $y = 0$ (normalized to one nucleon–nucleon collision) from single electrons in $|\eta| < 0.9$; statistical (inner bars) and quadratic sum of statistical and p_T-dependent systematic errors (outer bars) are shown; a 9% normalization error is not shown. Right: σ^{B} integrated in $2.5 < y^{\mathrm{B}} < 4$ (not normalized to one nucleon–nucleon collision) from single muons and di-muons in $2.5 < \eta < 4$; only (very small) statistical errors shown.

a specific B-meson $p_T > p_T^{\min}$ region, as preliminarly shown in Fig. 4 (right). Since only minimal cuts are applied, the reported statistical errors are very small and high-p_T reach is excellent. Systematic errors are currently under study.

4.3 Study of $\mathrm{b\overline{b}}$ correlations in ATLAS

author: Th. Lagouri

The ATLAS detector [19] is well engineered for studies of b-production, and together with the huge rate of b-quark production that will be seen at LHC, offers great potential for the making of novel precise b production measurements. Correlations between b and $\overline{\mathrm{b}}$ quarks and events with more than one heavy-quark pair, $\mathrm{b\overline{b}b\overline{b}}$, $\mathrm{b\overline{b}c\overline{c}}$, $\mathrm{b\overline{b}s\overline{s}}$, that were difficult to access in previous experiments due to limited statistics, will be investigated in detail. A new technique has been developed in ATLAS for measuring correlations, and this will yield results that will shed new light on our understanding of the QCD cross-section for $\mathrm{b\overline{b}}$-production.

A detailed study investigated a possibility of $\mathrm{b\overline{b}}$ correlations measurement using the $\Delta\phi$(J/ψ-μ) distribution, the azimuthal separation of a J/ψ and a muon [20, 21, 22]. This technique is expected to be superior to earlier methods used at the Tevatron Run-1 based on muon–muon or muon–b-jet correlations. The new method does not require separation cuts between the two objects. Such cuts were necessary to control the background, but they required a model-dependent extrapolation of the results to full azimuthal space [23]. Using a full simulation of the Inner Detector and the Muon Spectrometer of the ATLAS detector [19] it is shown that such a distribution can be extracted from heavy flavour events at LHC.

ATLAS studies were done for two channels selected to measure the azimuthal angle difference $\Delta\phi(\mathrm{b\overline{b}})$ between b and $\overline{\mathrm{b}}$ quarks:

$$\overline{\mathrm{b}} \to \mathrm{B_d} \to \mathrm{J}/\psi(\to \mu\mu)\mathrm{K}^0, \quad \mathrm{b} \to \mu + X$$

$$\overline{\mathrm{b}} \to \mathrm{B_s} \to \mathrm{J}/\psi(\to \mu\mu)\phi, \quad \mathrm{b} \to \mu + X$$

The numbers of events expected for 30 fb^{-1} as might be achieved after 3 years of running at a

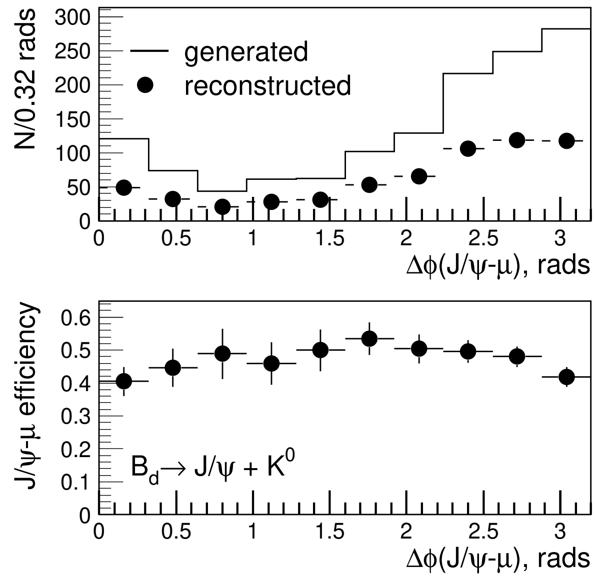


Fig. 5: Distribution of the opening angle $\Delta\phi$(J/ψ-μ) between the J/ψ from the decay $\mathrm{B_d} \to \mathrm{J}/\psi\mathrm{K}^0$ and the muon coming from the associated B hadron decay to muon, both direct b $\to \mu$ and indirect b $\to$ c $\to \mu$.

luminosity of 10^{33} cm^{2}s^{-1} are 4.8$\times$10^4 and 3.2$\times$10^4 respectively for these channels. No isolation cuts are needed to separate exclusively reconstructed B-decays from the muon produced in the semi-leptonic decay of the other B-particle in the event. The reconstruction efficiency remains high in topologies where the azimuthal angle difference $\Delta\phi$(J/ψ-μ) between J/ψ and the muon is small.

Special attention was devoted to background events in which the muon is produced from the decays K$^\pm$, $\pi^\pm \to \mu^\pm + X$ instead of b$\to \mu + X$. The study showed that this background is not problematic in B$_\mathrm{d}$ decays, however it is important in the case of B$_\mathrm{s}^0$ meson.

In summary, the results of the analysis suggest that backgrounds from K/π decays are small, and that backgrounds from events containing 4 b quarks are relatively flat in $\Delta\phi$(J/ψ-μ). The efficiency of the reconstruction of muons with this technique is also relatively flat in $\Delta\phi$(J/ψ-μ) and so we conclude that corrections to the measured $\Delta\phi$(J/ψ-μ) distribution are likely to be small.

4.4 Quarkonia measurements in ALICE

author: A. Morsch

Heavy quarkonia states are hard, penetrating probes which provide an essential tool to study the earliest and hottest stages of heavy-ion collisions [24]. They can probe the strongly interacting matter created in these reactions on short distance scales and are expected to be sensitive to the nature of the medium, i.e. confined or de-confined [25, 26]. The suppression (dissociation) of the heavy-quark resonances is considered as one of the most important observables for the study of the QGP at the LHC (see Ref. [2] for a recent review).

In ALICE, quarkonia will be measured in the di-electron channel using a barrel ($|\eta| < 0.9$) Transition Radiation Detector (TRD) [2] and in the di-muon channel using a forward Muon Spectrometer ($2.5 < \eta < 4$) [2]. The complete spectrum of heavy-quark vector mesons (J/ψ, ψ', Υ, Υ', Υ'') can be measured down to zero p_T. In particular the good mass resolution allows to resolve the Upsilon family.

The Muon Spectrometer uses a low-p_T trigger threshold, $p_T > 1$ GeV, on single muons for charmonia and a high-p_T trigger, $p_T > 2$ GeV, for bottonia detection. The TRD can trigger on single

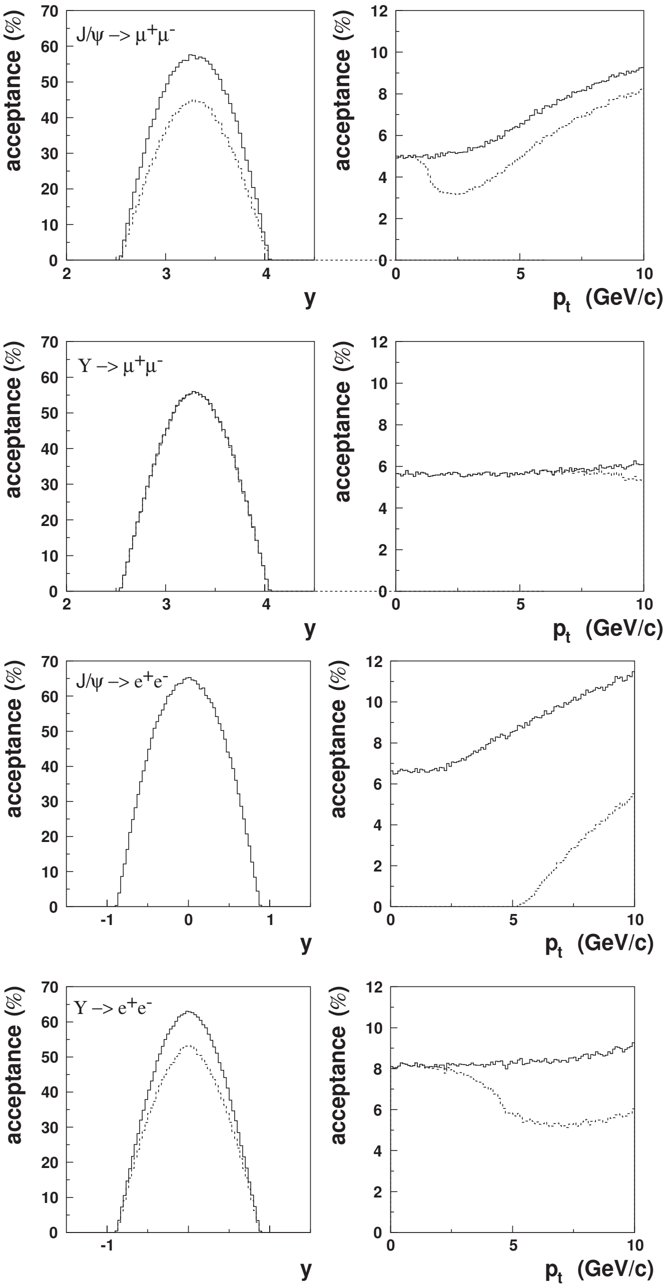


Fig. 6: Acceptance for J/ψ and Υ as a function of y and p_T for measurements in the di-muon channel and di-electron channels. To give an idea of the effect of the trigger, the acceptances are shown without (solid) and with (dashed) a sharp cut on the transverse momentum of single muons of 1 GeV/c (2 GeV/c) for J/ψ (Υ) and for single electrons of $3\ \mathrm{GeV}$.

electrons with $p_T > 3\ \mathrm{GeV}$, which results in a minimum transverse momentum of triggered charmonia of $5.2\ \mathrm{GeV}$. Electron identification combined with the excellent vertexing capabilities of the inner tracking system allows ALICE to distinguish direct charmonium production from secondary charmonium production through B decays.

The energy density dependence will be studied by varying the impact parameters and by studying in addition to the heaviest collision system (Pb–Pb) also intermediate mass and low mass A–A systems. To determine the primary production cross-section of the resonances and the amount of pre-resonance absorption, corresponding measurement have to be performed for pA and pp collisions.

Table 1: Main characteristics of quarkonia detection with the TRD and the Muon Spectrometer in ALICE.

	Muon Spectrometer	TRD
Acceptance	$2.5 < \eta < 4$	$\|\eta\| < 0.9$
Mass Resolution J/ψ	72 MeV	34 MeV
Mass Resolution Υ	99 MeV	93 MeV

Table 1 shows the main quarkonia detection characteristics of the TRD and the Muon Spectrometer and the acceptances for J/ψ and Υ as a function of y and p_T are shown in Fig. 4.4. In one year of pp running at $\langle L \rangle = 3 \times 10^{30}$ cm^{-2}s^{-1} ALICE will detect several 10^6 J/ψs and several 10^4 Υs in the di-muon channel. A Υ statistics of 10^2–10^3 can be obtained in the di-electron channel. For the J/ψ we expect $\approx 10^4$ untriggered low-p_T and $\approx 10^4$ high-p_T triggered events.

References

[1] ALICE Collaboration, ALICE Technical Proposal, CERN/LHCC 95-71 (1995).

[2] ALICE Collaboration, ALICE Physics Performance Report,Vol. I, CERN/LHCC 2003-49 (2003).

[3] ATLAS Collaboration, ATLAS Technical Proposal, CERN/LHCC 94-43 (1994).

[4] CMS Collaboration, The Compact Muon Solenoid Technical Proposal, CERN/LHCC 94-38, LHCC/P1 (1994).

[5] LHCb Collaboration, LHCb Technical Proposal, CERN/LHCC 98-004 (1998).

[6] ATLAS Collaboration, High-Level Trigger, Data Acquisition and Controls TDR, CERN/LHCC 2003-022 (2003).

[7] J. Baines, for the ATLAS Collaboration, ATLAS B-trigger Update, Nucl. Phys. Proc. Suppl. 120 (2003) 139 (2003).

[8] CMS Collaboration, CMS: The Tracker Project Technical Design Report, CERN/LHCC 98-06, CMS TDR 5 (1998).

[9] CMS Collaboration, Addendum to the CMS Tracker TDR, CERN/LHCC 2000-016, CMS TDR 5 Addendum 1 (2000).

[10] CMS Collaboration, CMS: The Trigger and Data Acquisition Project, Volume I: The Level-1 Trigger TDR, CERN/LHCC 2000-038, CMS TDR 6.1 (2000).

[11] CMS Collaboration, CMS: The Trigger and Data Acquisition Project, Volume II: Data Acquisition and High-Level Trigger TDR, CERN/LHCC 2002-26, CMS TDR 6.2 (2002).

[12] LHCb Collaboration, LHCb Trigger TDR, CERN/LHCC 2003-031 (2003).

[13] LHCb Collaboration, LHCb Computing TDR, in preparation.

[14] N. Carrer and A. Dainese, ALICE Internal Note, ALICE-INT-2003-019 [arXiv:hep-ph/0311225] (2003).

[15] N. Carrer, A. Dainese and R. Turrisi, J. Phys. **G 29**, 575 (2003).

[16] A. Dainese, Ph.D. Thesis, Università degli Studi di Padova [arXiv:nucl-ex/0311004] (2003).

[17] ALICE Collaboration, ALICE Physics Performance Report, Vol. II, in preparation.

[18] M. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. **B 373**, 295 (1998).

[19] ATLAS Collaboration, ATLAS Detector and Physics Performance Technical Design Report, CERN/LHCC 99-14 (1999).

[20] S. Robins, ATLAS Internal Note, ATL-PHYS-2000-026 (2000).

[21] Th. Lagouri, Eur. Phys. J. **C 33**, S497 (2004).

[22] Th. Lagouri, ATLAS Internal Note, ATL-COM-PHYS-2004-063 (2004).

[23] S.P. Baranov and M. Smizanska, Phys. Rev. **D 62**, 014012 (2000).

[24] H. Satz, Nucl. Phys. **A 590**, 63c (1995).

[25] T. Matsui and H. Satz, Phys. Lett. **B 178**, 416 (1995).

[26] D. Kharzeev and H. Satz, Phys. Lett. **B 334**, 155 (1994).

Small-x effects in heavy quark production

A. Dainese[a], *K. J. Eskola*[b], *H. Jung*[c], *V. J. Kolhinen*[b], *K. Kutak*[d,e], *A. D. Martin*[f], *L. Motyka*[c], *K. Peters*[c,g], *M. G. Ryskin*[h], *and R. Vogt*[i,j]
[a] University of Padova and INFN, Padova, Italy
[b] Department of Physics, University of Jyväskylä, Jyväskylä, Finland
[c] Deutsches Elektronen-Synchroton DESY, Hamburg, FRG
[d] University of Hamburg, Hamburg, FRG
[e] Henryk Niewodniczanski Institute of Nuclear Physics, Polish Academy of Sciences, Poland
[f] Institute for Particle Physics Phenomenology, University of Durham, Durham, UK
[g] Department of Physics and Astronomy, University of Manchester, UK
[h] Petersburg Nuclear Physics Institute, Gatchina, St. Petersburg, Russia
[i] Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[j] Physics Department, University of California, Davis, CA, USA

Abstract
We study small-x effects on heavy flavor production at the LHC in two approaches including nonlinear, saturation-motivated, terms in the parton evolution. One approach is based on collinear factorization, the other on k_T factorization. The prospects for direct experimental study of these effects in pp collisions at the LHC are discussed.

coordinators: A. Dainese, H. Jung, and R. Vogt

1 Introduction

HERA data are used to constrain the small x, moderate Q^2 parton densities in two approaches. In the first, HERA F_2 data are refit using DGLAP evolution with the first nonlinear recombination terms. Recombination slows the evolution so that, after refitting the data, the gluon distribution is enhanced relative to that obtained by DGLAP alone. The resulting set of parton densities produces charm enhancement in pp collisions at the LHC. On the other hand, assuming k_T factorization, the unintegrated gluon distribution is determined from the HERA F_2^c data, the only inclusive HERA measurements which directly accesses the gluon density. Saturation effects are then included, reducing the small x gluon densities with little distinguishable effect on F_2. This approach leads instead to heavy flavor suppression at the LHC. After a short general introduction, both approaches and their predicted effects on heavy quark production are discussed in detail. Direct experimental study of these effects in pp collisions at the LHC may be able to differentiate between the two approaches.

2 Small-x partons, absorption and the LHC

authors: A.D. Martin and M.G. Ryskin

2.1 Partons densities at low x?

Almost nothing is known about the behaviour of partons at low x. There are essentially no data available for $x < 10^{-4}$ with Q^2 in the perturbative region and there is no reliable theory to extrapolate down in x.

In the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP)-based [1, 2, 3, 4] global analyses, the small-x behaviour is driven by input distributions at a starting scale $Q = Q_0$. Usually these 'input' distributions are written in the form $x^{-\lambda}(1-x)^{\eta}$ where λ and η are free parameters fit to the data.

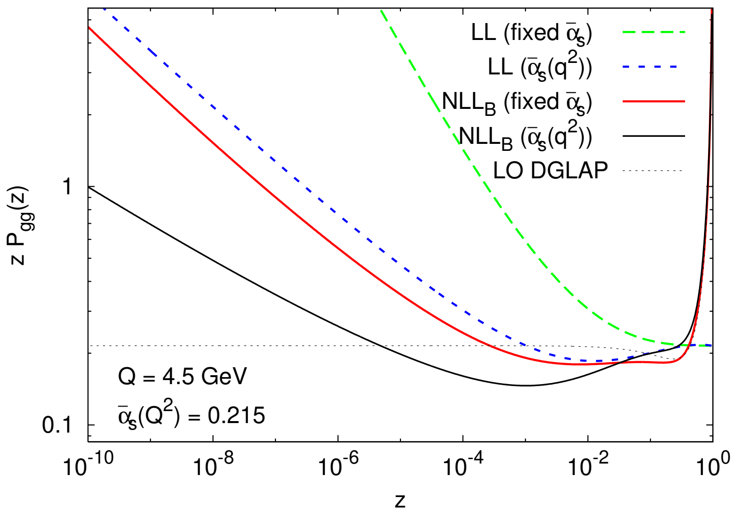


Fig. 1: The gluon-gluon splitting function, P_{gg}, with fixed and running coupling in the LL and resummed NLL BFKL approximations, compared with the LO DGLAP behaviour. The figure is taken from Ciafaloni et al. [9, 10, 11, 12, 13, 14]. The subscript B refers to scheme B which ensures energy-momentum conservation in the splitting.

So one can say nothing without data in the x region of interest. Moreover, there may be large low-x contributions to the gluon of the form $(\alpha_s \ln(1/x))^n$ – the so-called Balitsky-Fadin-Kuraev-Lipatov (BFKL) effects [5, 6, 7, 8], beyond the DGLAP approximation.

Thus it seems better to discuss low-x behaviour in terms of BFKL-evolution. However there are also problems here. The next-to-leading logarithm (NLL) corrections to the leading order (LO) BFKL (CCFM) amplitude are known to be very large and one needs to resum such corrections to obtain a relatively stable result. We cannot justify the perturbative QCD approach at low Q^2 so that the solution of the BFKL equation need to be matched to some non-perturbative amplitude at $Q = Q_0$. This non-perturbative distribution (analogous to the ‘input’ in the DGLAP case) is not known theoretically. Either it has to be fit to low x data or some phenomenological model such as a Regge-based parametrization has to be used.

After a reasonable resummation of the NLL corrections in the region where the starting virtuality Q_0 is not close to the final value of Q, $Q > Q_0$, the resummed BFKL amplitude turns out to be similar to that resulting from DGLAP evolution [9, 10, 11, 12, 13, 14]. For example, the preasymptotic effects on the resummation of the gluon-gluon splitting function are so large that the NLL BFKL power growth only sets in for $z < 10^{-5}$, as can be seen from Fig. 1. Moreover, the behaviour of the convolution $P_{gg} \otimes g/g$, normalized to g, in the NNLO DGLAP and NLL approximations is exactly the same down to $z \sim 10^{-4}$ [15].

Thus, in practice, the DGLAP and BFKL based approaches are rather close to each other in the HERA kinematic regime. In both cases, the main problem is the low-x behaviour of the amplitude at $Q = Q_0$ where we need to phenomenologically determine possible non-perturbative contributions, power corrections and so on.

2.2 The puzzle of the $x^{-\lambda}$ behaviour

Since the BFKL amplitude grows as a power of x, $A \propto x^{-\lambda}$, it will violate unitarity as $x \to 0$. Indeed, even after the NLL resummation, the expected power, $\lambda \simeq 0.3$, is rather large. Thus, we first discuss absorption effects which tame the violation of unitarity. The upper limit of the small x behaviour of the

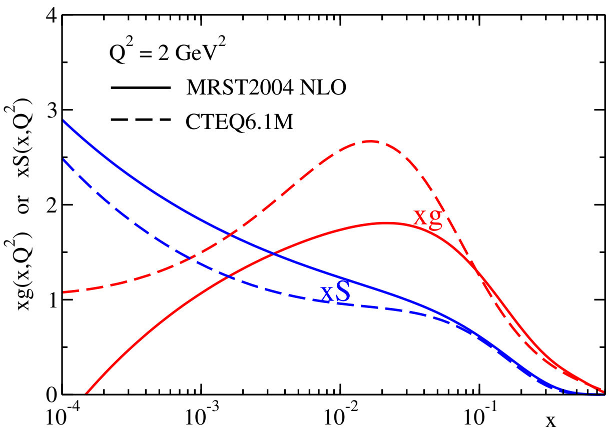


Fig. 2: The behaviour of the gluon and sea quark distributions at $Q^2 = 2$ GeV^2 found in the CTEQ6.1M [17] and MRST2004 NLO [18] global analyses. The valence-like behaviour of the gluon is evident.

parton distributions $a = g, q$ is given by the extrapolation

$$xa(x, q^2) = \left(\frac{x_0}{x}\right)^{0.3} x_0 a(x_0, q^2) \tag{1}$$

below $x_0 = 0.001$. The distributions are reliably determined from global parton analyses at $x > x_0$.

On the other hand, it is reasonable to expect that at $Q \lesssim Q_0 \sim 1$ GeV the behaviour will reflect that of hadron-hadron interactions: $\lambda = 0.08$ [16]. Most likely the lower value of λ is explained by absorptive/screening effects in hadron-hadron collisions. So, for extrapolation down to $x \sim 10^{-7} - 10^{-6}$ we may regard $\lambda = 0.08$ as a lower limit since, in DIS, we expect smaller absorptive effects than those in hadron-hadron interactions.

However, present global analyses, which do not allow for absorption effects, reveal that at $Q \sim 1 - 1.5$ GeV and low x, the sea quarks have a Pomeron-like behaviour, $xq \sim x^{-0.2}$, whereas the gluon distribution has a valence-like behaviour, $xg \sim \sqrt{x}$. This different behaviour is evident from Fig. 2, which shows the behaviour of the gluon and sea quark distributions, $xS = 2x(\bar{u} + \bar{d} + \bar{s})$ for $Q^2 = 2$ GeV^2. Such a result looks strange from the Regge viewpoint where the same vacuum singularity (Pomeron) should drive both the sea quarks and the gluons since the same power is expected for sea quarks and gluons, $\lambda_g = \lambda_q$.

This difference demonstrates that the actual situation is even more complicated. It is worth noting that a simultaneous analysis of inclusive and diffractive DIS data indicates that, after accounting for screening effects and allowing for some power corrections, it may be possible to describe the HERA data with $\lambda_g = \lambda_q = 0$ [19]. The absorptive effects, estimated from fitting the diffractive DIS data, enlarge the input gluon distribution at low x.

It may initially seem strange that accounting for absorptive effects gives a *larger* gluon density[1] at low x and Q^2. The point is that the only way to describe the data, *which are sensitive to absorptive effects*, within the framework of DGLAP evolution without absorption, is to choose a very low 'input' gluon density in order to *mimic* the screening corrections 'hidden' in the data. Indeed, there is a tendency for the gluon distribution to even become negative at low x and Q^2. On the other hand, allowing for absorption during DGLAP evolution (with the help of the Gribov-Levin-Ryskin (GLR) and Mueller-Qiu (MQ), GLRMQ, equations [22, 23]) the same data may be described with a larger and definitely positive input gluon density at $Q = Q_0$.

[1]The same result was obtained in Ref. [20, 21] – note, however, it was based on LO evolution and the large NLO correction to P_{qg} will change the q, g relationship.

2.3 Estimates of absorptive effects: GLRMQ to BK

The saturation of parton densities ($\lambda = 0$) may be obtained using the Balitski-Kovchegov (BK) [24, 25] equation, based on the BFKL equation, as well as the aforementioned GLRMQ equations. The latter equation is based on DGLAP evolution. These equations sum the set of fan diagrams which describe the rescattering of intermediate partons on the target nucleon. The screening caused by these rescatterings prohibits the power growth of the parton densities.

The GLR equation for the gluon may be written symbolically as

$$\frac{\partial xg}{\partial \ln Q^2} \;=\; P_{gg} \otimes g + P_{gq} \otimes q - \frac{81\alpha_s^2}{16R^2Q^2} \int \frac{dy}{y} [y\, g(y, Q^2)]^2 \, . \tag{2}$$

The nonlinear shadowing term, $\propto -[g]^2$, arises from perturbative QCD diagrams which couple four gluons to two gluons so that two gluon ladders recombine into a single gluon ladder. The minus sign occurs because the scattering amplitude corresponding to a gluon ladder is predominantly imaginary. The parameter R is a measure of the transverse area πR^2 where the gluon density is sufficiently dense for recombination to occur.

The BK equation is an improved version of the GLR equation. It accounts for the more precise triple-pomeron vertex (first calculated in Ref. [26, 27, 28]) and can be used for the non-forward amplitude. The GLR equation was in momentum space, whereas the BK equation is written in coordinate space in terms of the dipole scattering amplitude $N(\mathbf{x}, \mathbf{y}, Y) \;\equiv\; N_{\mathbf{xy}}(Y)$. Here $\mathbf{x}$ and $\mathbf{y}$ are the transverse coordinates of the two t-channel gluons which form the colour-singlet dipole and $Y = \ln(1/x)$ is the rapidity. The BK equation reads

$$\frac{\partial N_{\mathbf{xy}}}{\partial Y} \;=\; \frac{\alpha_s N_c}{\pi} \int \frac{d^2\mathbf{z}}{2\pi} \frac{(\mathbf{x}-\mathbf{y})^2}{(\mathbf{x}-\mathbf{z})^2(\mathbf{y}-\mathbf{z})^2} \left\{ N_{\mathbf{xz}} + N_{\mathbf{yz}} - N_{\mathbf{xy}} - N_{\mathbf{xz}}\, N_{\mathbf{yz}} \right\} \, . \tag{3}$$

For small dipole densities, N, the quadratic term in the brackets may be neglected and Eq. (3) reproduces the conventional BFKL equation. However for large N, that is $N \to 1$, the right-hand side of Eq. (3) vanishes and we reach saturation when $N = 1$. The equation sums up the set of fan diagrams where at small Y the target emits any number of pomerons (i.e. linear BFKL amplitudes) while at large Y we have only one BFKL dipole.

Starting from the same initial conditions, the solution of the BK equation gives *fewer* small-x partons than that predicted by its *parent* linear BFKL/CCFM equation[2].

In principle, it would appear more appropriate to use the BFKL-based BK equation to describe the parton densities at low x. Unfortunately, however, the BK equation is only a model. It cannot be used for numerical predictions. We discuss the reasons below.

2.4 Status of the BK equation

The Balitski-Kovchegov (BK) equation [24, 25] is an attempt to describe saturation phenomena. However it is just a ‘toy model’ and cannot, at present, be used to reliably estimate absorptive effects at small x. The reasons are as follows:

- The BK equation is based on the LO BFKL/CCFM equation. We know that the NLL corrections are large. We need to know the NLL corrections not only for the linear part of the evolution, but also for the nonlinear term.
- Even neglecting the NLL corrections, we need to match the solution to a boundary condition at rather low Q^2. This boundary condition is not theoretically known.

[2]Analogously, starting from the same input (and not fitting the input to the data) the GLR equation gives fewer small-x partons than that predicted by the parent linear DGLAP equation.

- It sums a limited set of diagrams. The selection of diagrams (the fan graphs) was justified in the region where absorptive effects are relatively small. When these corrections become important, as in the saturation region, one has to allow for many other graphs whose contributions become comparable to the fan diagram contributions[3].
- To solve the BK equation we need an initial condition at fixed x and all Q^2. These conditions are not well enough known. In particular, the maximum (saturation) value of the gluon density depends on the radius: $xg(x, q^2) \propto R^2q^2$. At the moment, the radius R is a free parameter. It may be small — the so-called 'hot spot' scenario. Moreover, we should account for the possibility of dissociation of the target proton[4]. The contribution coming from the dissociation is expected to have a smaller R.

2.5 Relevance to, and of, the LHC

How do the uncertainties at low x affect the predictions for the LHC? Fortunately for inclusive production of possible massive new particles with $M \gtrsim 100$ GeV, the partons are sampled at x values and scales M reliably determined from NLO and NNLO global analyses. For illustration, we discuss W production which has been studied in detail [29, 30, 31]. Central W production ($y_W = 0$) at the LHC samples partons at $x = M_W/\sqrt{s} = 0.006$. However to predict the total cross section, σ_W, we need to integrate over rapidity, important for $|y_W| \lesssim 4$ so that σ_W has some sensitivity to partons as low as $x \sim 10^{-4}$. The total uncertainty on the NNLO prediction of σ_W has been estimated to be $\pm 4\%$ [29]. Therefore W production at the LHC can serve as a good luminosity monitor. To reduce the uncertainty in the prediction of σ_W will require a better theoretical understanding of low x partons.

Of course, if the new particles are sufficiently massive, $M \gtrsim 1$ TeV, and produced by gluon fusion, then the uncertainties due to the PDFs will be larger. However, there are situations where the scale is considerably lower such as exclusive double-diffractive Higgs production which depends on the unintegrated gluon at $Q^2 \approx 5$ GeV^2 with $x \sim M_H/\sqrt{s} \sim 0.01$. The absorptive effects are also expected to be small here.

Turning the discussion around, is it possible for the LHC experiments to determine the behaviour of partons in the x region below 10^{-4} at low scales? One possibility is $\mu^+\mu^-$ Drell-Yan production in which events are observed with the $\mu^+\mu^-$ invariant mass as low as possible and the rapidity as large as possible. For example, for $M_{\mu\mu} = 4$ GeV and $y_{\mu\mu} = 3$, we sample quarks at $x = 1.4 \times 10^{-5}$. This process predominantly samples the sea quark distributions. To study the small x behaviour of the gluon at low scales we may consider χ_c production, or prompt photon production driven by the subprocess $gq \to \gamma q$.

In practice, rather than χ_c, it may be better to study $pp \to J/\psi\ X$ as a function of $y_{J/\psi}$. This process is also sensitive to the gluon distribution through the subprocesses $gg \to J/\psi\ g,\ \ gg \to \chi \to J/\psi\ \gamma$. There are also contributions from the subprocesses $gg \to b\bar{b}$ with $b \to J/\psi$, and $q\bar{q} \to J/\psi$. The analysis of such data will be considerably helped by the detailed observations of prompt J/ψ and J/ψ from b in central production at the Tevatron [32]. In fact, the first ever NLO global parton analysis [33] used J/ψ data as a function of rapidity to constrain the gluon distribution.

The LHCb detector covers the rapidity region of $2 < \eta < 5$ [34], and may be able to perform some of the above measurements. There is another possibility. Since LHCb will operate at a luminosity of 2×10^{32} cm^{-2}s^{-1}, there will usually be a single collision per bunch crossing and thus practically no 'pile-up' problems. Installing a forward detector at LHCb would offer the possibility of observing asymmetric events with one very large rapidity gap to probe the region of $x_{I\!P} \leq 10^{-5}$.

[3]Unfortunately the problem of summing all relevant diagrams has not been solved, even in the simpler case of Reggeon field theory.

[4]We know that these channels provide more than $30 - 40\%$ of F_2^D measured at HERA.

3 Including nonlinear terms in gluon evolution: the GLRMQ and BK approaches

3.1 GLRMQ approach

authors: K.J. Eskola and V.J. Kolhinen

The DGLAP [1, 2, 3, 4] evolution equations describe the scale evolution of the parton distribution functions (PDFs) well in the region of large interaction scale, $Q^2 \gtrsim 4$ GeV^2 [35, 17, 36]. However, toward small values of x and Q^2, the gluon recombination terms start to play an increasingly important role. The inclusion of correction terms which arise from fusion of two gluon ladders leads to nonlinear power corrections to the DGLAP evolution equations. The first of these nonlinear corrections are the GLRMQ terms.

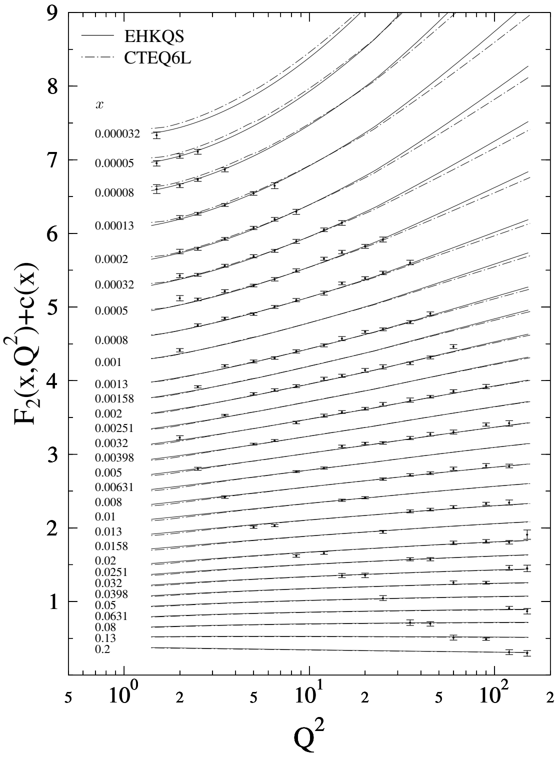


Fig. 3: Calculated $F_2(x, Q^2)$ values compared with the H1 data.

With the GLRMQ corrections, the gluon evolution equation becomes

$$\frac{\partial xg(x,Q^2)}{\partial \ln Q^2} = \left.\frac{\partial xg(x,Q^2)}{\partial \ln Q^2}\right|_{\rm DGLAP} - \frac{9\pi}{2}\frac{\alpha_s^2}{Q^2}\int_x^1 \frac{dy}{y} y^2 G^{(2)}(y,Q^2). \tag{4}$$

We model the two-gluon density in the latter term on the right-hand side as

$$x^2 G^{(2)}(x,Q^2) = \frac{1}{\pi R^2}[xg(x,Q^2)]^2, \tag{5}$$

where $R = 1$ fm is the radius of the proton (we comment further on this later). The corrections to the sea quark distributions are

$$\frac{\partial xq(x,Q^2)}{\partial \ln Q^2} \approx \left.\frac{\partial xq(x,Q^2)}{\partial \ln Q^2}\right|_{\rm DGLAP} - \frac{3\pi}{20}\frac{\alpha_s^2}{Q^2} x^2 G^{(2)}(x,Q^2).$$

We have assumed that the higher-twist gluon term, G_{HT} [23], is negligible.

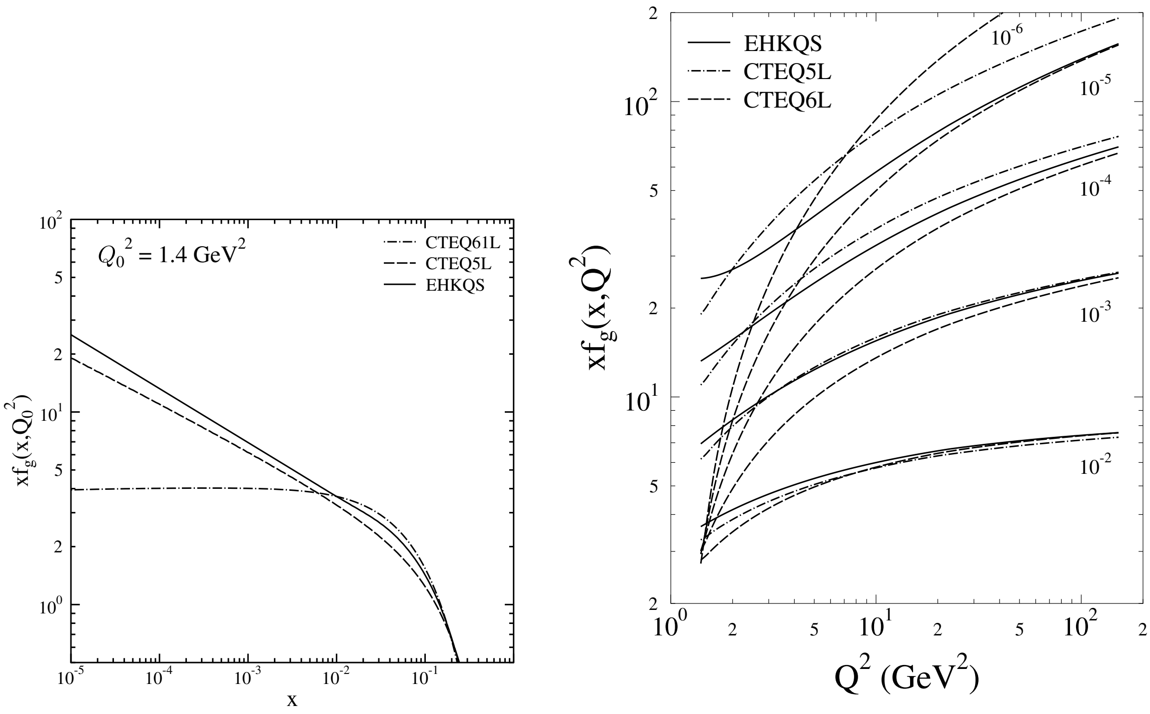


Fig. 4: Left: initial gluon distributions at $Q_0^2 = 1.4$ GeV^2. Right: evolution of gluon distributions for several fixed values of x shows that the effect of the nonlinear terms vanishes as Q^2 increases.

Since these correction terms are negative, they slow down the evolution of the PDFs. Due to the $1/Q^2$ dependence, they also die out in the evolution so that at large scales Eqs. (4) and (6) relax into the linear DGLAP equations.

In order to study the interplay between the nonlinear corrections and the initial PDFs and observe the nonlinear effects in fits to the DIS data, in Ref. [37] we compared the structure function $F_2(x, Q^2) = \sum_q e_q^2[xq(x, Q^2)+x\bar{q}(x, Q^2)]$, calculated with the nonlinearly-evolved PDFs, to the HERA H1 data [38]. As reference distributions we used the CTEQ5L and CTEQ6L PDF sets at large scales. We chose these sets because the CTEQ collaboration uses only the large scale, $Q^2 > 4$ GeV^2, data in their fits, thus avoiding some of the possible nonlinear effects appearing in the small x, $Q^2 < 4$ GeV^2 region in their initial distributions.

At small x, sea quarks dominate F_2 and the gluon distribution dictates its scale evolution. At leading order (LO), the DGLAP contribution can be approximated as $\partial F_2(x, Q^2)/\partial \ln Q^2 \approx (10\alpha_s/27\pi)xg(2x, Q^2)$ [39]. Larger $xg(x, Q^2)$ values correspond to faster $F_2(x, Q^2)$ evolution. The scale evolution of $F_2(x, Q^2)$ at small x computed with CTEQ5L is too fast due to a large small-x small-Q^2 gluon distribution. The newer CTEQ6L set has much smaller gluon distribution in this region (see Fig. 4 (left)), giving a slower evolution and hence a good fit to the H1 data.

Our goal in Ref. [37] was to determine whether the good fit to the data could be maintained using the GLRMQ-corrected DGLAP scale evolution together with initial scale PDFs differing from CTEQ6L. We constructed a new set of PDFs using the CTEQ5L and CTEQ6L distributions piecewise as baselines at scales $Q^2 \sim 3 - 10$ GeV^2 where the linear terms dominate the evolution and evolved them nonlinearly to lower Q^2. We then interpolated between the sets in x and assumed a power-like dependence at small-x

for gluons and sea quarks. These initial PDF candidates were then evolved to higher scales and compared to the data. This iterative procedure was repeated until a sufficiently good fit to the data was found.

As a result, we obtained a new set of initial PDFs[5], called EHKQS, which, when evolved using the nonlinear DGLAP+GLRMQ evolution equations, produced an equally good or even better fit to the H1 data relative to CTEQ6L, shown in Fig. 3. At $Q^2 \sim 1.4$ GeV^2 and $x \sim 10^{-5}$, a good fit to the HERA data requires the nonlinear evolution to be compensated by a larger gluon distribution than obtained with DGLAP alone. The enhancement is a factor of ~ 6 relative to CTEQ6L, as shown in Fig. 4 (left). The Q^2 dependence of EHKQS is compared to CTEQ6L and CTEQ5L in Fig. 4 (right) for several different values of x.

We used $R = 1$ fm as the free proton radius in the two-gluon density term. We did not repeat the calculations with different R but, depending on the transverse matter density of the free proton, some $\sim 20\%$ uncertainty in R can be expected. Since the nonlinear contributions decrease as R increases, a larger R would lead to reduced enhancement of the small x gluons below $Q^2 \sim 10$ GeV^2. Thus, minimizing the χ^2 of the fit with respect to R is a future task.

3.2 BK approach

authors: H. Jung, K. Kutak, K. Peters, L. Motyka

A theoretical framework capable of describing the QCD evolution of parton densities taking gluon rescattering (corresponding to nonlinear effects) into account is the Balitsky-Kovchegov (BK) equation [24, 40, 41, 42, 25]. The equation, based on the BFKL approach [6, 43, 7], may be used to determine the unintegrated gluon density. The BK equation resums the BFKL pomeron fan diagrams with the triple pomeron vertex derived in the high energy limit of QCD. In the doubly logarithmic limit, the BK equation reduces [25] to the collinear Gribov-Levin-Ryskin (GLR) equation [22]. It is the non-collinear limit, however, which gives the dominant contribution to the triple pomeron vertex [44, 45]. We conclude that GLR approach misses an essential part of the nonlinear gluon dynamics.

The solution to the BK equation, constrained by the low-x HERA data will be used to extrapolate the parton densities to the LHC kinematical domain. Extrapolation is necessary as the LHC may probe very low values of x, down to 10^{-7} for $M = 10$ GeV and $\eta \sim 9$, where unitarity corrections may be important even at relatively large scales of a few GeV^2. Last but not least, unitarity corrections may also break k_T factorization. We will also discuss which processes may be affected.

This section is organized as follows. First we give a brief description of the formalism used to determine the gluon evolution. Within this formalism, we fit the HERA charm structure function, F_2^c, data, the most relevant inclusive measurement directly sensitive to the gluon density. Using further assumptions about the sea quarks, F_2 can also be described well. The resulting gluon density is then used to compute heavy quark production and to investigate the nonlinear effects. First we estimate $b\bar{b}$ production at CDF and D0. Then, cross sections for heavy quark production at various LHC experiments are estimated, tracing the impact of the unitarity corrections. Finally, conclusions are given.

The standard framework to determine parton evolution is the collinear DGLAP formalism. It works rather well for inclusive quantities but, for more exclusive processes, the k_T-factorization scheme is more appropriate because both the longitudinal and transverse components of the gluon momenta are considered. In this framework, the process-independent quantity is the unintegrated gluon distribution, connected to the process-dependent hard matrix element via the k_T-factorization theorem. Linear evolution of the unintegrated gluon distribution may be described by one of the small x evolution equations using the k_T-factorization scheme, the BFKL and CCFM [46, 47, 48, 49] equations. These equations are based on resummation of large logarithmic pQCD corrections, $\alpha_s^n \ln^m(1/x)$, and are equivalent at the leading logarithmic level.

[5]In fact, we produced three new sets of initial distributions, differing by the charm quark mass and parton distribution at the initial scale, see Ref. [37] for more details. All sets produced equally good fits to the HERA data.

The very small x kinematic region is also the regime where the growth of the gluon density must be tamed in order to preserve unitarity. Recently, a successful description of unitarity corrections to DIS was derived within the color dipole formulation of QCD. This is the Balitsky-Kovchegov (BK) equation which describes the BFKL evolution of the gluon in a large target, including a nonlinear term corresponding to gluon recombination at high density.

In our analysis, we determine the unintegrated gluon distribution from the BK equation unified with the DGLAP equation following KMS (Kwieciński, Martin and Staśto) [50, 51, 52, 53]. We use the abbreviation KKMS (Kutak, Kwieciński, Martin and Staśto) [52, 53] for the unified nonlinear equation. The linear part of this equation is given by the BFKL kernel with subleading $\ln(1/x)$ corrections, supplemented by the non-singular parts of the DGLAP splitting functions. Thus resummation of both the leading $\ln Q^2$ and $\ln(1/x)$ terms are achieved. The subleading terms in $\ln(1/x)$ are approximated by the so-called consistency constraint and the running coupling constant. The nonlinear part is taken directly from the BK equation, ensuring that the unitarity constraints are preserved. One expects that this framework provides a more reliable description of the gluon evolution at extremely small x, where $\ln(1/x) \gg 1$ and the unitarity corrections are important, than does DGLAP.

We give a short review of the KKMS equation, starting from the impact parameter dependent BK equation. The equation for the unintegrated gluon density, $h(x,k^2,b)$, at impact parameter b from the center of the target, becomes

$$\frac{\partial h(x,k^2,b)}{\partial \ln 1/x} = \frac{\alpha_s N_c}{\pi} k^2 \int_{k_0^2} \frac{dk'^2}{k'^2} \left\{ \frac{h\left(x,k'^2,b\right) - h\left(x,k^2,b\right)}{|k'^2 - k^2|} + \frac{h\left(x,k^2,b\right)}{[4k'^4 + k^4]^{\frac{1}{2}}} \right\}$$
$$-\pi\alpha_s \left(1 - k^2 \frac{d}{dk^2}\right)^2 k^2 \left[\int_{k^2}^{\infty} \frac{dk'^2}{k'^4} \ln\left(\frac{k'^2}{k^2}\right) h(x,k'^2,b)\right]^2 , \qquad (6)$$

the BFKL equation at LLx accuracy, extended by the negative recombination term. The (dimensionless) unintegrated gluon distribution is obtained from $h(x,k^2,b)$ by integration over b,

$$f(x,k^2) = \int d^2b\, h(x,k^2,b). \qquad (7)$$

A comment about the impact parameter treatment is in order. In Eq. (7), we assume that the evolution is local in b. However, the complete BK equation results in some diffusion in the impact parameter plane. This diffusion effect may be neglected if the target is much larger than the inverse of the saturation scale. In this scheme, the impact parameter dependence enters through the initial condition at large x_0, $h(x_0,k^2,b) = f(x_0,k^2)S(b)$ where $f(x_0,k^2)$ is the unintegrated gluon distribution. Note that, due to nonlinearities, the b dependence of $h(x,k^2,b)$ does not factorize from x and k at low x.

The input profile function is assumed to be Gaussian, $S(b) = \exp(-b^2/R^2)/\pi R^2$, where R^2 corresponds to the square of the average nucleon radius. Since the size of the target, R, sets the magnitude of the initial parton density in the impact parameter plane, $h(x_0,k^2,b)$, the unitarity corrections depend on R. At smaller R, gluons are more densely packed in the target and the nonlinear effects are stronger.

References [52, 53] proposed to combine Eq. (6) with the unified BFKL-DGLAP framework developed in Ref. [50]. In this scheme, the (linear) BFKL part is modified by the consistency constraint [54, 55], resulting in the resummation of most of the subleading corrections in $\ln(1/x)$ which arise from imposing energy-momentum conservation on the BFKL kernel [56, 57, 58, 59]. In addition, we assume that the strong coupling constant runs with scale k^2, another source of important NLLx corrections. Finally, the non-singular part of the leading order DGLAP splitting function and quark singlet distribution were included in the x evolution. The final improved nonlinear equation for the unintegrated

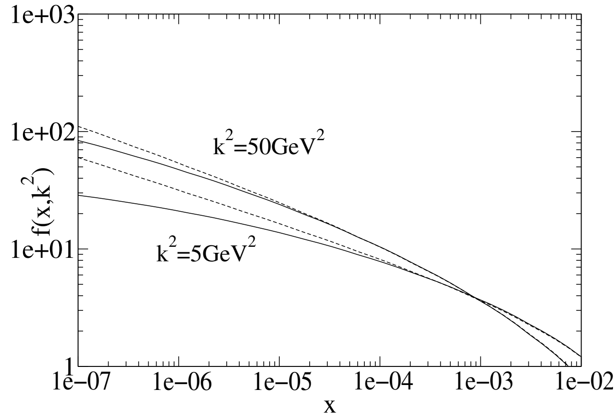


Fig. 5: The unintegrated gluon distribution obtained from Eq. (8) as a function of x for different values of k_T^2. The solid lines correspond to the solution of the nonlinear equation with $R = 2.8$ GeV^{-1} while the dashed lines correspond to the linear part.

gluon density is

$$\begin{aligned} h(x,k^2,b) &= \tilde{h}^{(0)}(x,k^2,b)+ \\ &+ \tfrac{\alpha_s(k^2)N_c}{\pi} k^2 \int_x^1 \tfrac{dz}{z} \int_{k_0^2} \tfrac{dk'^2}{k'^2} \left\{ \tfrac{h(\frac{x}{z},k'^2,b)\,\Theta(\frac{k^2}{z}-k'^2) - h(\frac{x}{z},k^2,b)}{|k'^2-k^2|} + \tfrac{h(\frac{x}{z},k^2,b)}{|4k'^4+k^4|^{\frac{1}{2}}} \right\}+ \\ &+ \tfrac{\alpha_s(k^2)}{2\pi} \int_x^1 dz \left[\left(P_{gg}(z) - \tfrac{2N_c}{z}\right) \int_{k_0^2}^{k^2} \tfrac{dk'^2}{k'^2} h(\tfrac{x}{z},k'^2,b) + P_{gq}(z)\Sigma\left(\tfrac{x}{z},k'^2,b\right) \right]+ \\ &-\pi \left(1 - k^2 \tfrac{d}{dk^2}\right)^2 k^2 \int_x^1 \tfrac{dz}{z} \left[\int_{k^2}^{\infty} \tfrac{dk'^2}{k'^4} \alpha_s(k'^2) \ln\left(\tfrac{k'^2}{k^2}\right) h(z,k'^2,b) \right]^2 . \end{aligned} \tag{8}$$

The second line of the equation corresponds to the BFKL evolution. The theta function, $\Theta(\frac{k^2}{z} - k'^2)$, reflects the consistency constraint that generates the dominant part of the subleading BFKL corrections. The third line corresponds to the DGLAP effects generated by the part of the splitting function, $P_{gg}(z)$, that is not singular in the limit $z \to 0$ and also by the quarks where $\Sigma(x, k^2, b^2)$ corresponds to the impact-parameter dependent singlet quark distribution. The nonlinear screening contribution following from the BK equation is given in the last term. The inhomogeneous contribution, defined in terms of the integrated gluon distribution, carries information about the transverse profile of the proton,

$$\tilde{h}^{(0)}(x,k^2,b) = \frac{\alpha_s(k^2)}{2\pi} S(b) \int_x^1 dz P_{gg}(z) \frac{x}{z} g\left(\frac{x}{z}, k_0^2\right) \ , \tag{9}$$

at $k_0^2 = 1$ GeV^2. The initial integrated density at k_0^2 is parameterized as

$$xg(x,k_0^2) = N(1-x)^\rho \tag{10}$$

where $\rho = 2.5$. The size of the dense gluon system inside the proton is assumed to be $R = 2.8$ GeV^{-1}, in accord with the diffractive slope, $B_d \simeq 4$ GeV^{-2}, of the elastic J/ψ photoproduction cross section at HERA. In this process, the impact parameter profile of the proton defines the t dependence of the elastic cross section, $B_d \simeq R^2/2$, by Fourier transform. In the 'hot-spot' scenario, the radius can be smaller, $R = 1.5$ GeV^{-1}. We also use the hot spot value to compare with measurements and make predictions for the LHC.

Equation (8) was solved numerically both in the linear approximation and in full. The method for solving Eq. (8) was developed in Refs. [50, 52]. In Fig. 5, the effects of linear and nonlinear evolution on the unintegrated gluon distribution are given as a function of x for $k^2 = 5$ and 50 GeV^2. Nonlinear evolution leads to sizeable suppression at the smallest x values. While the nonlinear effects are small in the HERA x range, they may be important at the LHC. In the following sections, we address the importance of these nonlinear effects.

The initial distribution in Eq. (10) was obtained by fitting the HERA F_2^c measurements [60, 61] using the Monte Carlo CASCADE [62, 63] for evolution and convolution with the off-shell matrix elements. We find χ^2 per degree of freedom of 0.46 (1.17) for H1 (ZEUS). The fits were repeated both with the standard KMS evolution without the nonlinear contribution and with extended KMS evolution including the nonlinear part. The predicted F_2^c is equivalent for both linear and nonlinear evolution, independent of R. Thus nonlinear evolution is only a small effect at HERA, even in the hot-spot scenario with $R = 1.5$ GeV^{-1}.

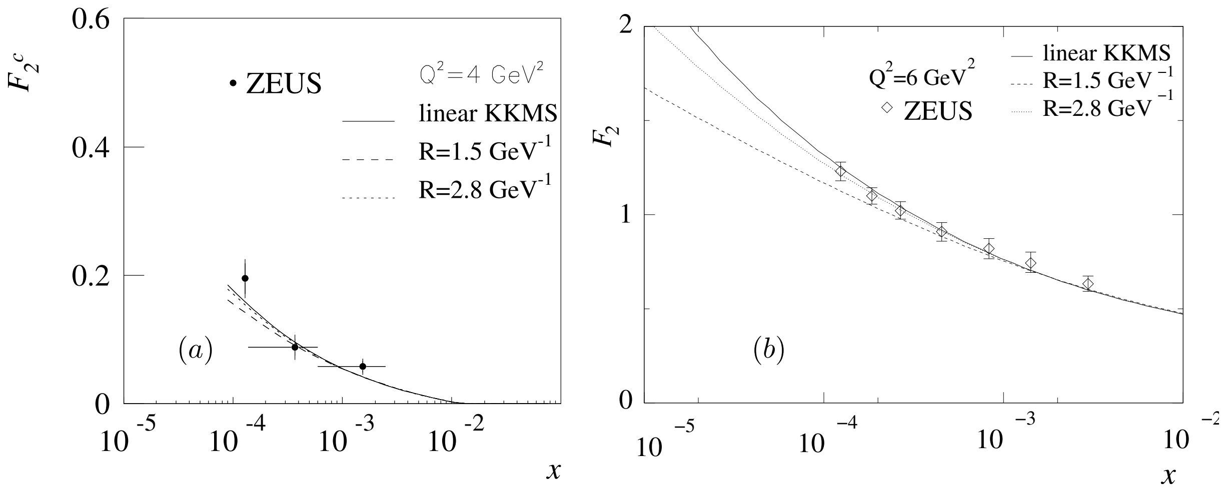


Fig. 6: The charm structure function, F_2^c, [61] at $Q^2 = 4$ GeV^2 (a) and F_2 [64] at $Q^2 = 6$ GeV^2 (b) obtained for KKMS evolution with different values of R.

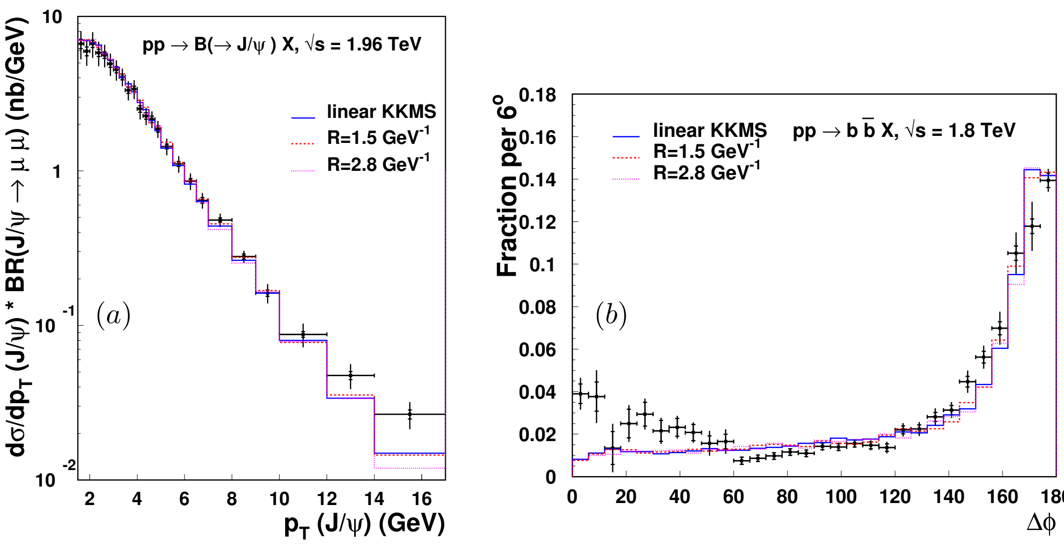


Fig. 7: Bottom production, measured by CDF, is compared to predictions using CASCADE with linear and nonlinear KKMS evolution, including two values of R. (a) The p_T distribution of B meson decays to J/ψ. (b) The azimuthal angle, $\Delta\phi$, distribution of $b\bar{b}$ pair production smeared by the experimental resolution.

In Fig. 6(a) we compare the measured F_2^c [61] to our prediction at $Q^2 = 4$ GeV^2. We have determined our initial distribution from F_2^c since it is the only inclusive measurement at HERA directly sensitive to the gluon distribution. However, we can also describe F_2 [64] by making further assumptions about the sea quark distribution, following the KMS approach. The agreement with the data, shown in Fig. 6(b), is also quite good. There is only a small effect for $Q^2 > 5$ GeV^2, even in the hot-spot scenario with $R = 1.5$ GeV^{-1}.

Next, this constrained gluon density was used to calculate $gg \to b\bar{b}$ production at the Tevatron as a cross check of the fit and the evolution formalism. We use $m_b = 4.75$ GeV and a renormalization scale in α_s of $Q^2 = 4m_b^2 + p_T^2$. The predicted cross section was then compared to both CDF [65, 66] and D0 [67] measurements. The predictions agree well with the data.

In Fig. 7(a) the cross section for B decays to J/ψ is shown as a function the J/ψ p_T [65, 66]. The KKMS gluon density fits the data well in all three scenarios with deviations only appearing for $p_T > 12$ GeV. It is interesting to note that the approach described here gives even better agreement than the NLO collinear approach [68].

In Fig. 7(b), the azimuthal angle distribution between the b and $\bar{b}$ quarks, $\Delta\phi$, is given. The $\Delta\phi$ and $b\bar{b}$ p_T distributions are correlated since $\Delta\phi < 180°$ corresponds to higher pair p_T. Since the k_T-factorization formula allows the incoming gluons to have sizable transverse momenta, the calculated $\Delta\phi$ distribution agrees very well with the data for $\Delta\phi > 60°$ with only smearing due to the experimental resolution. The enhancement of the data relative to the calculations at low $\Delta\phi$ requires further study.

Both plots compare linear (solid histograms) and nonlinear KKMS evolution (dotted and dashed histograms) for $R = 1.5$ GeV^{-1} and 2.8 GeV^{-1} respectively. The nonlinear part of the evolution also has no impact in this kinematic region.

4 Phenomenological applications: heavy quark production at the LHC

4.1 GLRMQ approach

author: R. Vogt

Since the HERA F_2 data can be described by both linear DGLAP and nonlinear DGLAP+GLRMQ evolution, as shown in Fig. 3 of Section 3.1, additional independent probes are needed. Here, we discuss how charm quark production in pp collisions could probe the gluon enhancement predicted in Section 3.1 and described in detail in Ref. [20, 21]. Charm production is an ideal choice since the charm mass is low and its production is dominated by gluons. Assuming factorization, the inclusive differential charm cross section is

$$d\sigma_{pp\to c\bar{c}X}(Q^2, \sqrt{s}) = \sum_{i,j,k=q,\bar{q},g} f_i(x_1, Q^2) \otimes f_j(x_2, Q^2) \otimes d\hat{\sigma}_{ij\to c\bar{c}\{k\}}(Q^2, x_1, x_2) \tag{11}$$

where $\hat{\sigma}_{ij\to c\bar{c}\{k\}}(Q^2, x_1, x_2)$ are the perturbatively calculable partonic cross sections for charm production at scales $Q^2 \sim m_T^2 \gg \Lambda_{\rm QCD}^2$, x_1 and x_2 are the parton momentum fractions and $f_i(x, Q^2)$ are the proton parton densities. We assume that the renormalization and factorization scales are equal. Only the leading order gg and $q\bar{q}$ channels are considered here.

The values of the charm quark mass and scale used in the calculations are chosen to give good agreement with the total cross section data at NLO: $m = 1.2$ GeV and $Q^2 = 4m^2$ for standard DGLAP-evolved NLO PDFs such as CTEQ6M [69] and MRST [70]. Nearly equivalent agreement may be obtained with $m = 1.3$ GeV and $Q^2 = m^2$ [71, 72]. Both choices assure that the PDFs are evaluated above the minimum scales. While scales proportional to m are used in the total cross section, inclusive calculations of distributions also depend on the transverse momentum scale, p_T, so that $m_T = \sqrt{m^2 + p_T^2}$ is used instead [73].

To illustrate the effects of the nonlinear EHKQS distributions [37] of Section 3.1 on charm production at the LHC, we show

$$R(y) \equiv \frac{d\sigma(\text{EHKQS})/dy}{d\sigma(\text{CTEQ61L})/dy} \quad \text{and} \quad R(p_T) \equiv \frac{d\sigma(\text{EHKQS})/dp_T}{d\sigma(\text{CTEQ61L})/dp_T} \tag{12}$$

in Fig. 8 where y is the charm quark rapidity. The results are calculated for the maximum LHC pp, pPb and Pb+Pb energies, $\sqrt{S} = 14$ (solid), 8.8 (dashed) and 5.5 (dot-dashed) TeV respectively. The results for $m = 1.2$ GeV and $Q^2 = 4m_T^2$ are on the left-hand side while those with $m = 1.3$ GeV and $Q^2 = m_T^2$ are on the right-hand side.

The change in the slope of $R(y)$ occurs when one x drops below the minimum value of the EHKQS set where further nonlinearities become important, $x_{\rm min}^{\rm EHKQS} = 10^{-5}$, and enters the unconstrained x region. The minimum x of CTEQ61L is lower, $x_{\rm min}^{\rm CTEQ61L} = 10^{-6}$. While the EHKQS gluon distribution is fixed at its minimum for $x < x_{\rm min}^{\rm EHKQS}$, the CTEQ61L distribution continues to change until $x_{\rm min}^{\rm CTEQ61L}$. In inclusive kinematics with an identified charm quark and fixed $x_T = 2m_T/\sqrt{S}$, the unconstrained x-region contributes to charm production in the region

$$y_l \equiv \ln\left(1/x_T - \sqrt{1/x_T^2 - 1/x_{\rm min}}\right) \leq |y| \leq \ln\left(1/x_T + \sqrt{1/x_T^2 - 1/x_{\rm min}}\right) \,. \tag{13}$$

The upper limit is close to the phase space boundary. Expanding the lower limit, y_l, in powers of $x_T^2/x_{\rm min} \ll 1$, $y_l \approx \ln[m_T/(x_{\rm min}\sqrt{S})] \geq \ln[m/(x_{\rm min}\sqrt{S})]$. If $m = 1.2$ GeV, the small x region contributes to charm production at $|y| \geq y_l = 2.2$, 2.6 and 3.1 for $\sqrt{S} = 14$, 8.8 and 5.5 TeV, respectively. The rather sharp turnover in $R(y)$ indicates where the $x < 10^{-5}$ region begins to contribute. For $|y| > y_l$ and $Q^2 > 4$ GeV^2, as x decreases, the CTEQ61L gluon distribution increases considerably above that of the EHKQS distribution. Thus $R(y) < 1$ at large rapidities when $Q^2 = 4m_T^2$. At midrapidity $R(y)$ is insensitive to the EHKQS extrapolation region, $x < x_{\rm min}^{\rm EHKQS}$. Since $R(y)$ is integrated over p_T, it not only reflects the enhancement at $m_T = m$ because $Q^2 \propto m_T^2$ and the p_T distribution peaks around $p_T \approx 1$ GeV. When $Q^2 = m_T^2$, the ratios are broad because the CTEQ61L gluon distribution is relatively flat as a function of x for $Q^2 \sim 2 - 3$ GeV^2. The enhancement decreases and broadens with decreasing energy.

Since the rapidity distributions are rather flat, there are still important contributions to the p_T distributions from the extrapolation region, up to $\sim 30\%$ at $\sqrt{S} = 14$ TeV for $m = 1.2$ GeV and $Q^2 = 4m^2$. Thus the sensitivity of $R(p_T)$ to the unconstrained region should be kept in mind. At the largest $\sqrt{S}$, the contribution from the $x < 10^{-5}$ region is greatest and if $Q^2 \geq 4m^2$, $xg^{\rm CTEQ61L}(x,Q^2) > xg^{\rm EHKQS}(x,Q^2)$. Because the contribution from the region $x < 10^{-5}$ decreases with $\sqrt{S}$, at low p_T $R(p_T)$ decreases with energy. In contrast, for $Q^2 = m_T^2$, $xg^{\rm EHKQS}(x,Q^2) > xg^{\rm CTEQ61L}(x,Q^2)$ and the enhancement decreases with energy.

Because the DGLAP gluon distributions are already well constrained by HERA data, they cannot absorb additional large effects. Therefore we conclude that, if a low-p_T enhancement in the charm cross section relative to the DGLAP-based result is observed in future experiments, it is a signal of nonlinear effects on the PDF evolution.

4.2 BK approach

authors: H. Jung, K. Kutak, K. Peters, L. Motyka

Since the Tevatron measurements are well described using the unintegrated parton densities constrained by HERA and convoluted with the off-shell matrix elements, the same approach may be used for heavy quark production at the LHC at *e.g.* $\sqrt{s} = 14$ TeV. As discussed previously, see also Fig. 5, heavy quark production at this energy is already in the region where saturation effects may be relevant. In the kinematic regions, such as at the LHC, where nonlinear evolution may become important, the cross section will be suppressed due to the negative sign of the nonlinear term in Eq. (8).

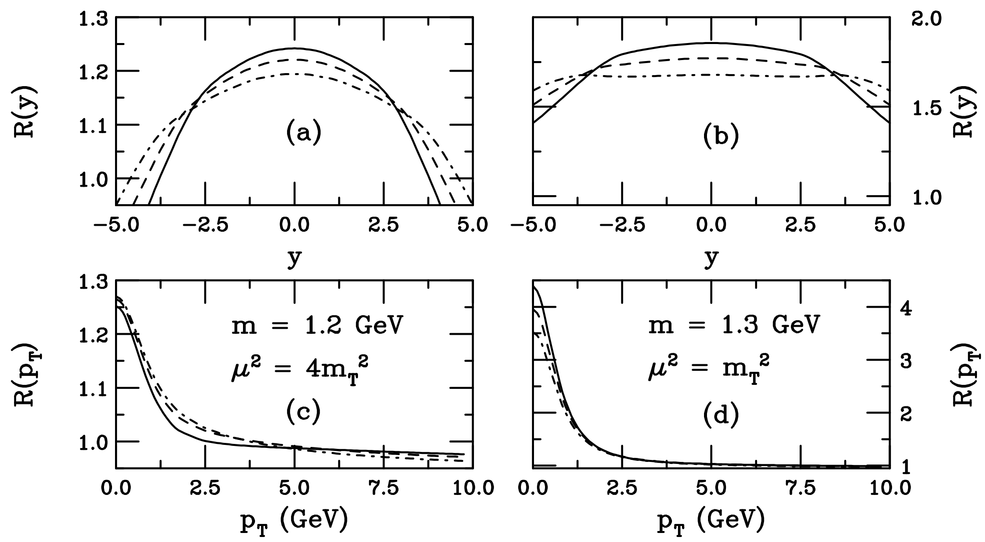


Fig. 8: We present $R(y)$, (a) and (c), and $R(p_T)$, (b) and (d), in pp collisions at $\sqrt{S} = 14$ (solid), 8.8 (dashed) and 5.5 (dot-dashed) TeV. The left-hand side shows $m = 1.2$ GeV and $Q^2 = 4m_T^2$, the right-hand side $m = 1.3$ GeV and $Q^2 = m_T^2$.

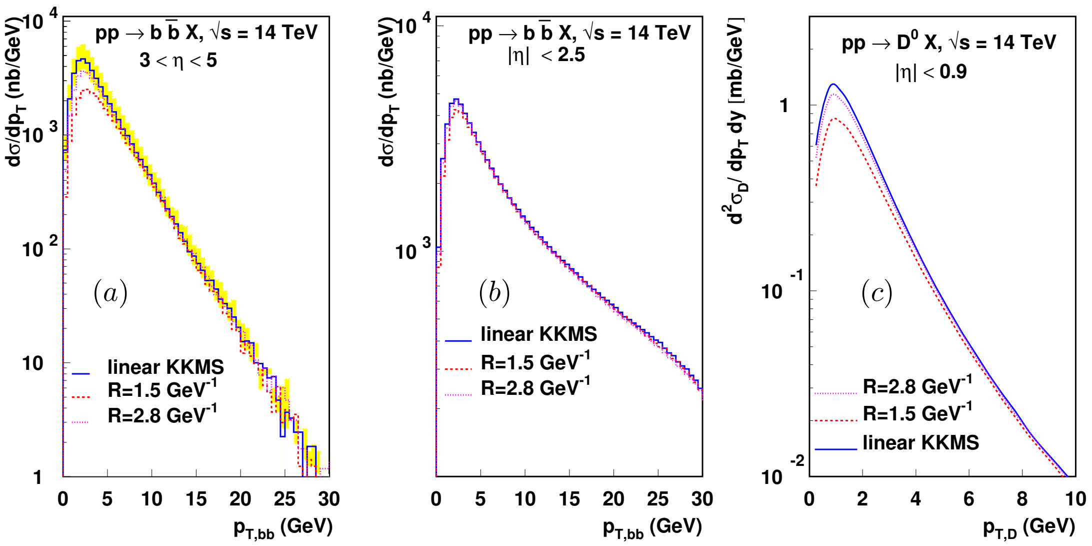


Fig. 9: (a) and (b) show $b\bar{b}$ production as a function of pair p_T without cuts in $3 < |\eta| < 5$ (a) and in the ATLAS/CMS acceptance (b). The D^0 meson p_T distribution in the ALICE acceptance is shown in (c).

First, we compute the $b\bar{b}$ production cross section at 14 TeV without any experimental cuts. In Fig. 9(a) the $b\bar{b}$ differential cross section is shown as a function of pair p_T in the forward region, $3 < |\eta| < 5$. We compare linear evolution (solid histogram), nonlinear evolution with $R = 1.5$ GeV^{-1} (dashed histogram) and $R = 2.8$ GeV^{-1} (dotted histogram). The grey band shows the uncertainty in the linear result due to the b quark mass. We take a central value of 4.75 GeV (the solid histogram) and vary m_b from 4.5 to 5 GeV. The $b\bar{b}$ pair results are shown since the pair p_T is most sensitive to the gluon k_T and thus to the saturation effects. In the hot-spot scenario, saturation effects are visible for $p_{T_{b\bar{b}}} < 5$ GeV. These saturation effects grow with rapidity, increasing the suppression to a factor of $3 - 4$ in the higher rapidity regions. For larger R, the saturation effects are not very significant.

In Fig. 9(b), the $b\bar{b}$ production cross section is computed within the ATLAS and CMS acceptance ($p_T > 10$ GeV and $|\eta| < 2.5$ for both the b and $\bar{b}$ quarks, see Ref. [74]). With these cuts, the observed suppression due to nonlinear effects nearly vanishes. This result suggests that k_T factorization can safely be applied in the central η region. Thus saturation effects due to nonlinear gluon evolution are seen only for $p_{T_{b\bar{b}}} < 10$ GeV and at high η. This regime is accessible with upgraded ATLAS/CMS detectors or in LHCb where the b quark p_T can be measured to 2 GeV for $1.9 < \eta < 4.9$. In this kinematic regime, the hot-spot scenario predicts a factor of two suppression of the $b\bar{b}$ cross section.

Similarly, we investigated $c\bar{c}$ production at ALICE. In ALICE, it will be possible to measure the D^0 down to $p_T \sim 0.5$ GeV in $|\eta| < 0.9$. The result is shown in Fig. 9(c) with $m_c = 1.5$ GeV. In the hot-spot scenario (dashed curve), a factor of two suppression occurs at $p_T \sim 1$ GeV.

5 Perspectives for experimental observation at LHC

authors: A. Dainese and R. Vogt

5.1 Introduction

In Section 4.1, charm production in pp collisions at the LHC was suggested as a promising way to study the effects of nonlinear evolution on the parton densities. Due to gluon dominance of charm production and the small values of x and Q^2 probed, $x \approx 2 \times 10^{-4}$ and $Q^2 \approx 1.69 - 6$ GeV^2 at midrapidity and transverse momentum[6] $p_T \approx 0$, charm production at the LHC is sensitive to the gluon enhancement arising from nonlinear evolution. The resulting charm enhancement was quantified in Ref. [20, 21] by the LO ratios of the differential cross sections computed with the nonlinearly-evolved EHKQS PDFs [37], obtained from DGLAP+GLRMQ evolution, relative to the DGLAP-evolved CTEQ61L PDFs.

The enhancement of the nonlinearly-evolved gluons increases as x and Q^2 decrease. Consequently, the charm enhancement increases with center of mass energy, $\sqrt{S}$. Thus the maximum enhancement at the LHC will be at $\sqrt{S} = 14$ TeV and small charm quark transverse momentum. The sensitivity of the charm enhancement to the value of the charm quark mass, m, as well as to the choice of the factorization, Q_F^2, and renormalization, Q_R^2, scales was studied in Ref. [20, 21] assuming $Q^2 = Q_F^2 = Q_R^2 \propto m_T^2$ where $m_T^2 = p_T^2 + m^2$. The most significant charm enhancement occurs when m and Q^2/m_T^2 are both small. A comparison of the NLO total cross sections with low energy data shows that the data prefer such small m and Q^2 combinations [71, 72]. The largest enhancement is obtained with $m = 1.3$ GeV and $Q^2 = m_T^2$, see Fig. 8 in Section 4.1.

In Section 4.1, only charm enhancement was described. Neither its subsequent hadronization to D mesons nor its decay and detection were considered. In this section, we address these issues to determine whether the charm enhancement survives hadronization and decay to be measured in the ALICE detector [75].

The calculation described in Section 4.1 was to leading order since the EHKQS sets are evolved according to the LO DGLAP+GLRMQ equations using a one-loop evaluation of α_s. Thus these LO

[6] Here we use p_T for the transverse momentum of the charm quark and p_T^D for the transverse momentum of the D meson.

distributions should generally not be mixed with NLO matrix elements and the two-loop α_s. However, the charm quark total cross section is increased and the p_T distribution is broadened at NLO relative to LO [76]. Thus, to determine whether or not the enhancement is experimentally measurable, we assume that the enhancement is the same at NLO and LO and employ a NLO cross section closest to the calculation of the enhancement in Ref. [20, 21].

As described in Ref. [76], the theoretical K factor may be defined in more than one way, depending on how the LO contribution to the cross section is calculated. In all cases, the $\mathcal{O}(\alpha_s^3)$ contribution to cross section is calculated using NLO PDFs and the two-loop evaluation of α_s. If the LO contribution is also calculated using NLO PDFs and a two-loop α_s, this is the "standard NLO" cross section. It is used in most NLO codes, both in the global analyses of the NLO PDFs and in evaluations of cross sections and rates [76]. The K factor formed when taking the ratio of the "standard NLO" cross section to the LO cross section with the NLO PDFs [76], $K_0^{(1)}$, indicates the convergence of terms in a fixed-order calculation [77]. On the other hand, if the LO contribution to the total NLO cross section employs LO PDFs and the one-loop α_s, we have a cross section which we refer to here as "alternative NLO". The K factor calculated taking the ratio of the "alternative NLO" cross section to the LO cross section with LO PDFs [76], $K_2^{(1)}$, indicates the convergence of the hadronic cross section toward a result. If $K_0^{(1)} > K_2^{(1)}$, convergence of the hadronic cross section is more likely [77]. This is indeed the case for charm production [76]. We also note that $K_2^{(1)}$ is a much weaker function of energy than $K_0^{(1)}$. Since, in the absence of nonlinear NLO PDFs, the "alternative NLO" cross section is more consistent with the calculated enhancement, we use this cross section to calculate the NLO D meson rates and p_T spectra. We note also that, in both cases, the p_T distributions have the same slope even though $K_2^{(1)}$, for the alternative NLO cross section, is somewhat smaller. Thus, using a non-standard NLO calculation will not change the slope of the p_T distributions, distorting the result.

The LO and NLO calculations used to obtain the full NLO result in both cases can be defined by modification of Eq. (11) in Section 4.1. We define the full LO charm production cross section as

$$d\sigma_{\rm LO}^{\rm 1L} = \sum_{i,j=q,\overline{q},g} f_i^{\rm LO}(x_1, Q_F^2) \otimes f_j^{\rm LO}(x_2, Q_F^2) \otimes d\hat{\sigma}_{ij\to c\overline{c}}^{\rm LO}(\alpha_s^{\rm 1L}(Q_R^2), x_1, x_2) \tag{14}$$

where the superscript "LO" on $d\hat{\sigma}_{ij\to c\overline{c}}$ indicates the use of the LO matrix elements while the superscript "1L" indicates that the one-loop expression of α_s is used. The LO cross section typically used in NLO codes employs the NLO PDFs and the two-loop (2L) α_s so that

$$d\sigma_{\rm LO}^{\rm 2L} = \sum_{i,j=q,\overline{q},g} f_i^{\rm NLO}(x_1, Q_F^2) \otimes f_j^{\rm NLO}(x_2, Q_F^2) \otimes d\hat{\sigma}_{ij\to c\overline{c}}^{\rm LO}(\alpha_s^{\rm 2L}(Q_R^2), x_1, x_2)\ . \tag{15}$$

In either case, the NLO contribution, $\mathcal{O}(\alpha_s^3)$ for heavy quark production, is

$$d\sigma_{\mathcal{O}(\alpha_s^3)} = \sum_{i,j=q,\overline{q},g} f_i^{\rm NLO}(x_1, Q_F^2) \otimes f_j^{\rm NLO}(x_2, Q_F^2) \otimes \sum_{k=0,q,\overline{q},g} d\hat{\sigma}_{ij\to c\overline{c}k}^{\rm NLO}(\alpha_s^{\rm 2L}(Q_R^2), Q_F^2, x_1, x_2) \tag{16}$$

where the superscript "NLO" on $d\hat{\sigma}_{ij\to c\overline{c}k}$ indicates the use of the NLO matrix elements. The additional sum over k in Eq. (16) includes the virtual ($k = 0$) and real ($k = q$, $\overline{q}$ or g depending on i and j) NLO corrections. In the calculations of $d\sigma_{\rm LO}^{\rm 2L}$ and $d\sigma_{\mathcal{O}(\alpha_s^3)}$, we use the value of $\Lambda_{\rm QCD}^{(4)}$ given for the NLO PDFs and work in the $\overline{\rm MS}$ scheme. The standard NLO cross section is then

$$d\sigma_{\rm NLO}^{\rm std} = d\sigma_{\rm LO}^{\rm 2L} + d\sigma_{\mathcal{O}(\alpha_s^3)} \tag{17}$$

while our "alternative NLO" cross section is defined as

$$d\sigma_{\rm NLO}^{\rm alt} = d\sigma_{\rm LO}^{\rm 1L} + d\sigma_{\mathcal{O}(\alpha_s^3)}\ . \tag{18}$$

Since the enhancement in Ref. [20, 21] was defined using $d\sigma_{\rm LO}^{\rm 1L}$ only, the best we can do is to use the alternative NLO cross section in our analysis, as described below.

We now discuss how the enhancement is taken into account in the context of the NLO computation. We calculate the LO inclusive charm p_T distribution, $d^2\sigma/dp_Tdy$, with the detected charm (anticharm) quark in the rapidity interval Δy with $|y| < 1$, motivated by the pseudorapidity acceptance of the ALICE tracking barrel, $|\eta| < 0.9$. The rapidity, y_2, of the undetected anticharm (charm) quark is integrated over. The charm enhancement factor $R(p_T, \Delta y)$ is then

$$R(p_T, \Delta y) = \frac{\displaystyle\int_{\Delta y} dy \int dy_2 \frac{d^3\sigma(\text{EHKQS})}{dp_T dy dy_2}}{\displaystyle\int_{\Delta y} dy \int dy_2 \frac{d^3\sigma(\text{CTEQ61L})}{dp_T dy dy_2}} \ . \tag{19}$$

Next, we assume that the enhancement calculated at LO is the same when calculated at NLO. This is the only reasonable assumption we can make to test whether the enhancement can be detected with ALICE which will measure the physical p_T^D distribution. The alternative NLO cross section is therefore the closest in spirit to the LO computation in Ref. [20, 21]. Thus, the enhanced NLO charm p_T distribution is

$$R(p_T, \Delta y) \ d\sigma_{\rm NLO}^{\rm alt}(\Delta y)/dp_T \ . \tag{20}$$

Our results are obtained with the same parameters used in Section 4.1, $m = 1.2$ GeV and $Q^2 = 4m_T^2$ as well as $m = 1.3$ GeV and $Q^2 = m_T^2$. These two choices are the baseline results against which other parameter choices will be compared to see if the enhancement is detectable.

5.2 From charm to D enhancement

To make a more realistic D meson distribution, we have modified the charm p_T distribution by the heavy quark string fragmentation in PYTHIA [78]. Charm events in pp collisions at $\sqrt{S} = 14$ TeV are generated using PYTHIA (default settings) with the requirement that one of the quarks is in the interval $|y| < 1$. The charm quarks are hadronized using the default string model. Since c and $\overline{c}$ quarks fragment to D and $\overline{D}$ mesons[7], respectively, in each event related (c, D) and $(\overline{c}, \overline{D})$ pairs can easily be identified[8]. These pairs are reweighted to match an arbitrary NLO charm quark p_T distribution, $dN_{\rm NLO}^c/dp_T$. If $dN_{\rm PYTHIA}^c/dp_T$ is the charm p_T distribution given by PYTHIA, each (c, D) pair is assigned the weight

$$\mathcal{W}(p_T) = \frac{dN_{\rm NLO}^c/dp_T}{dN_{\rm PYTHIA}^c/dp_T} \tag{21}$$

where p_T is the transverse momentum of the charm quark of the pair. Therefore, the reweighted final-state D distribution corresponds to the one that would be obtained by applying string fragmentation to the NLO c-quark distribution. The resulting D distribution is significantly harder than that obtained using the Peterson fragmentation function [79]. The enhancement survives after fragmentation although the D enhancement is somewhat lower than that of the charm because for a given p_T^D, the D spectrum receives contributions from charm quarks with $p_T \gtrsim p_T^D$, where the charm enhancement is smaller.

5.3 Sensitivity to the enhancement

Figure 10 shows the double-differential D^0 cross section, $d^2\sigma_D/dp_T^Ddy$, in $|y| < 1$ as a function of the transverse momentum. The points represent the expected "data" measured by ALICE, obtained from the alternative NLO cross section scaled by the enhancement factor $R(p_T, \Delta y)$ defined in Eq. (19),

[7] Here $D \equiv D^+, D^0$.

[8] Events containing charm baryons were rejected.

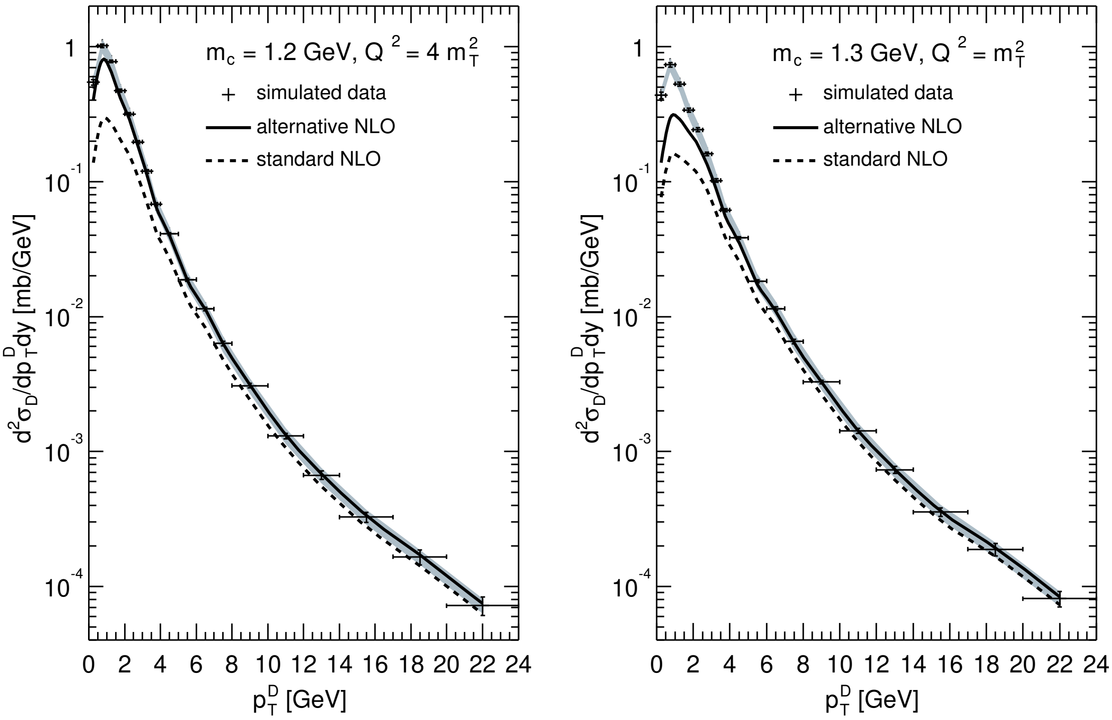


Fig. 10: Comparison of the simulated ALICE data generated from $R(p_T, \Delta y) d\sigma_{\rm NLO}^{\rm alt}$ with the alternative (solid) and standard (dashed) NLO calculations. The effect of string fragmentation is included in the "data" points as well as in the curves. The left-hand side shows the result for $m = 1.2$ GeV and $Q^2 = 4m_T^2$ while the right-hand side is the result for $m = 1.3$ GeV and $Q^2 = m_T^2$. The error bars on the data represent the statistical error and the shaded band represents the p_T-dependent systematic error. The 5% normalization error is not shown.

and modified by string fragmentation. The solid and dashed curves are obtained by applying string fragmentation to the alternative NLO and standard NLO $c\overline{c}$ cross sections, respectively. Thus, the "data" points include the enhancement while the curves do not. The horizontal error bars indicate the bin width, the vertical error bars represent the statistical error and the shaded band gives the p_T-dependent systematic error. The 5% p_T-independent systematic error on the normalization is not shown. (See Ref. [80] for a discussion of the error analysis. The standard NLO cross section, Eq. (17), and the $\mathcal{O}(\alpha_s^3)$ contribution to the alternative NLO cross section, Eq. (16), were calculated using the HVQMNR code [81] with CTEQ6M and $\Lambda_{\rm QCD}^{(4)} = 0.326$ GeV. The LO contribution to the alternative NLO cross section, Eq. (14), was calculated using the CTEQ61L PDFs. Fragmentation was included as described in Section 5.2. The enhancement, the difference between the data and the solid curves for $p_T^D \lesssim 3$ GeV, is more pronounced for the larger mass and lower scale, on the right-hand side of Fig. 10.

There is a significant difference between the alternative and standard NLO distributions. Part of the difference is due to the one- and two-loop evaluations of α_s since $\alpha_s^{\rm 2L} < \alpha_s^{\rm 1L}$. However, the most important contribution is the large differences between the LO and NLO gluon distributions, especially at low scales [80].

In order to address the question of the experimental sensitivity to the effect of nonlinear gluon evolution on low-p_T charm production, we consider, as a function of p_T^D, the ratio of the simulated data, including the enhancement, to alternative NLO calculations using a range of m and Q^2 with PYTHIA string fragmentation. We denote this ratio as "Data/Theory" and try to reproduce this ratio with NLO calculations employing recent linearly-evolved PDFs and tuning m and Q^2.

Since the enhancement has disappeared for $p_T^D \gtrsim 5$ GeV, we refer to this unenhanced region as high p_T^D. The p_T^D region below 5 GeV, where the enhancement is important, is referred to as low p_T^D.

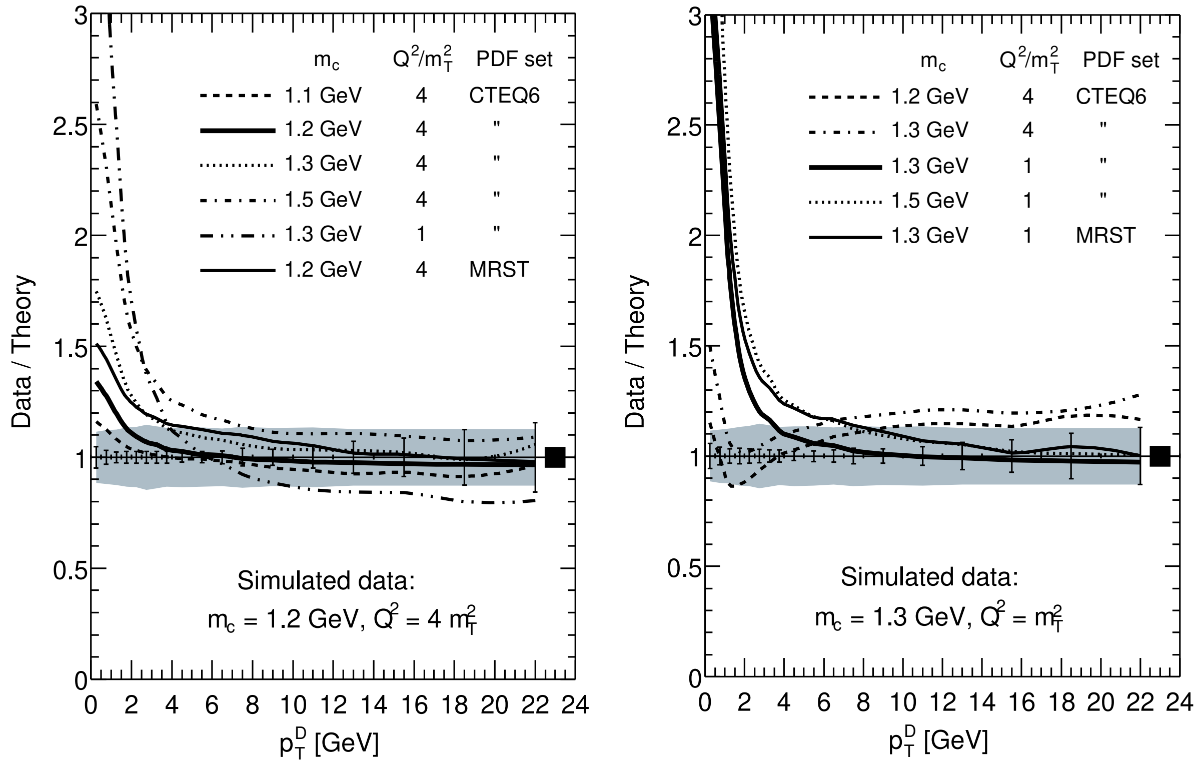


Fig. 11: Ratio of the generated ALICE data relative to calculations of the alternative NLO cross sections with several sets of parameters and PYTHIA string fragmentation. The left-hand side shows the result for $m = 1.2$ GeV and $Q^2 = 4m_T^2$ while the right-hand side is the result for $m = 1.3$ GeV and $Q^2 = m_T^2$.

If no set of parameters can describe both the high- and low-p_T^D components of the distribution equally well, and, if the set that best reproduces the high-p_T^D part underestimates the low-p_T^D part, this would be a strong indication of the presence of nonlinear effects.

The Data/Theory plots are shown in Fig. 11. The points with the statistical (vertical bars) and p_T-dependent systematic (shaded region) error correspond to the data of Fig. 10, including the enhancement, divided by themselves, depicting the sensitivity to the theory calculations. The black squares on the right-hand sides of the lines $\mathrm{Data/Theory} = 1$ represent the 5% p_T-independent error on the ratio coming from the cross section normalization. This error is negligible relative to present estimates of other systematic uncertainties ($\simeq 13\%$).

On the left-hand side, the thick solid curve with $m = 1.2$ GeV and $Q^2 = 4m_T^2$ best agrees with the high-p_T^D ratio by construction since $R \approx 1$ at large p_T^D. It also shows the effect of the enhancement well beyond the error band for $p_T^D \lesssim 2$ GeV. Better agreement with the data over the entire p_T^D range can be achieved only by choosing a charm quark mass lower than 1.2 GeV, below the nominal range of charm masses, as illustrated by the dashed curve for $m = 1.1$ GeV. Higher masses with $Q^2 = 4m_T^2$ produce much larger Data/Theory ratios than the input distribution. The ratio with $m = 1.3$ GeV and $Q^2 = m_T^2$ (dot-dot-dashed curve) gives a much larger ratio at low p_T^D and drops below ≈ 1 for $p_T^D > 8$ GeV.

We also present the ratio using the MRST parton densities (MRST2001 LO [36] in Eq. (14) and MRST2002 NLO [82] in Eq. (16)) with $m = 1.2$ GeV and $Q^2 = 4m_T^2$. We find that this result also agrees reasonably well with the CTEQ6 results for the same m and Q^2. Thus, the enhancement seems to be rather independent of the PDF. The CTEQ61L and the MRST2001 LO distributions are similar at low x, suggesting that non-linearly evolved PDFs based on MRST2001 LO would produce an enhancement like that of Ref. [20, 21].

On the right-hand side of Fig. 11 the thick solid curve, employing the same parameters as the data, gives the best agreement at high p_T^D. We note that even though the results with $Q^2 = 4m_T^2$ and $m \leq 1.3$ GeV lie closer to the data at low p_T^D and within the combined statistical and systematic error at higher p_T^D, the curves with these parameters have the wrong slopes for $p_T^D \lesssim 8$ GeV. The statistical sensitivity is expected to be good enough to distinguish the difference in curvature. The results with the MRST PDFs do not alter the conclusions.

5.4 Conclusions

We have studied whether the EHKQS gluon distributions [37] could generate an observable D meson enhancement in pp collisions at the LHC. Using the EHKQS LO PDFs and LO matrix elements for charm quark production and PYTHIA string fragmentation for D meson hadronization, the enhancement indeed survives to the D mesons.

The D meson enhancement, however, drops rapidly with transverse momentum. Therefore, D measurement capability at small p_T^D is necessary to verify the effect experimentally. The ALICE detector can directly reconstruct $D^0 \to K^-\pi^+$. We have demonstrated that, in the most optimistic case, the enhancement can be detected above the experimental statistical and systematic errors. When the charm mass is somewhat smaller, $m = 1.2$ GeV, but the scale is larger, $Q^2 = 4m_T^2$, it is more difficult to detect the enhancement over the experimental uncertainties.

Acknowledgments

The work of K. Kutak was supported in part by the Graduiertenkolleg Zukünftige Entwicklungen in der Teilchenphysik. M.G. Ryskin would like to thank the IPPP at the University of Durham for hospitality, and A.D. Martin thanks the Leverhulme Trust for an Emeritus Fellowship. Their work was supported by the Royal Society, by grants INTAS 00-00366, RFBR 04-02-16073, and by the Federal Program of the Russian Ministry of Industry, Science and Technology SS-1124.2003.2. The work of RV was supported in part by the Director, Office of Energy Research, Division of Nuclear Physics of the Office of High Energy and Nuclear Physics of the U. S. Department of Energy under Contract Number DE-AC02-05CH11231.

References

[1] Yu. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[2] V. N. Gribov, and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).

[3] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 675 (1972).

[4] G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977).

[5] L. N. Lipatov, Sov. J. Nucl. Phys. **23**, 338 (1976).

[6] E. A. Kuraev, L. N. Lipatov and V. S. Fadin, Sov. Phys. JETP **44**, 443 (1976).

[7] I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[8] L. N. Lipatov, Sov. Phys. JETP **63**, 904 (1986).

[9] M. Ciafaloni, D. Colferai, G. P. Salam and A. M. Stasto, Phys. Lett. **B 576**, 143 (2003).

[10] M. Ciafaloni, D. Colferai, G. P. Salam and A. M. Stasto, Phys. Rev. **D 68**, 114003 (2003).

[11] R. S. Thorne, Phys. Lett. **B 474**, 372 (2000).

[12] R. S. Thorne, Phys. Rev. **D 64**, 074005 (2001).

[13] G. Altarelli, R. D. Ball and S. Forte, Nucl. Phys. **B 674**, 459 (2003).

[14] V. A. Khoze, A. D. Martin, M. G. Ryskin and W. J. Stirling, Phys. Rev. **D 70**, 074013 (2004).

[15] G. P. Salam, arXiv:hep-ph/0501097 (2005).

[16] A. Donnachie and P. V. Landshoff, Phys. Lett. **B 296**, 227 (1992).

[17] J. Pumplin et al., JHEP **07**, 012 (2002).

[18] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Phys. Lett. **B 604**, 61 (2004).

[19] A. D. Martin, M. G. Ryskin and G. Watt, Phys. Rev. **D 70**, 091502 (2004).

[20] K. J. Eskola, V. J. Kolhinen and R. Vogt, Phys. Lett. **B 582**, 157 (2004).

[21] K. J. Eskola, V. J. Kolhinen and R. Vogt, J. Phys. **G 30**, S1171 (2004).

[22] L. V. Gribov, E. M. Levin and M. G. Ryskin, Phys. Rept. **100**, 1 (1983).

[23] A. H. Mueller and J.-W. Qiu, Nucl. Phys. **B 268**, 427 (1986).

[24] I. Balitsky, Nucl. Phys. **B 463**, 99 (1996).

[25] Yu. V. Kovchegov, Phys. Rev. **D 60**, 034008 (1999).

[26] J. Bartels, Phys. Lett. **B 298**, 204 (1993).

[27] J. Bartels, Z. Phys. **C 60**, 471 (1993).

[28] J. Bartels and M. Wusthoff, Z. Phys. **C 66**, 157 (1995).

[29] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C 35**, 325 (2004).

[30] J. Huston, J. Pumplin, D. Stump and W. K. Tung, arXiv:hep-ph/0502080 (2005).

[31] R. S. Thorne, A. D. Martin, R. G. Roberts and W. J. Stirling, arXiv:hep-ph/0507015 (2005).

[32] CDF Coll., D. Acosta et al., Phys. Rev. **D 71**, 032001 (2005).

[33] A. D. Martin, R. G. Roberts and W. J. Stirling, Phys. Rev. **D 37**, 1161 (1988).

[34] LHCb Coll., Z. J. Ajaltouni et al., Nucl. Phys. Proc. Suppl. **115**, 297 (2003).

[35] H. L. Lai et al., Eur. Phys. J. **C 12**, 375 (2000).

[36] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C 23**, 73 (2002).

[37] K. J. Eskola, H. Honkanen, V. J. Kolhinen, J.-W. Qiu and C. A. Salgado, Nucl. Phys. **B 660**, 211 (2003).

[38] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 21**, 33 (2001).

[39] K. Prytz, Phys. Lett. **B 311**, 286 (1993).

[40] I. Balitsky, Phys. Rev. Lett. **81**, 2024 (1998).

[41] I. Balitsky, Phys. Rev. **D 60**, 014020 (1999).

[42] I. Balitsky, Phys. Lett. **B 518**, 235 (2001).

[43] E. A. Kuraev, L. N. Lipatov and V. S. Fadin, Sov. Phys. JETP **45**, 199 (1977).

[44] A. Kovner, J. G. Milhano and H. Weigert, Phys. Rev. **D 62**, 114005 (2000).

[45] J. Bartels and K. Kutak, *Momentum space analysis of the triple pomeron vertex*, 2005. In preparation.

[46] M. Ciafaloni, Nucl. Phys. **B 296**, 49 (1988).

[47] S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. **B 234**, 339 (1990).

[48] S. Catani, F. Fiorani and G. Marchesini, Nucl. Phys. **B 336**, 18 (1990).

[49] G. Marchesini, Nucl. Phys. **B 445**, 49 (1995).

[50] J. Kwiecinski, A. D. Martin and A. M. Stasto, Phys. Rev. **D 56**, 3991 (1997).

[51] M. A. Kimber, J. Kwiecinski and A. D. Martin, Phys. Lett. **B 508**, 58 (2001).

[52] K. Kutak and J. Kwiecinski, Eur. Phys. J. **C 29**, 521 (2003).

[53] K. Kutak and A. M. Stasto, Eur. Phys. J. **C 41**, 343 (2005).

[54] B. Andersson, G. Gustafson, H. Kharraziha and J. Samuelsson, Z. Phys. **C 71**, 613 (1996).

[55] J. Kwiecinski, A. D. Martin and P. J. Sutton, Z. Phys. **C 71**, 585 (1996).

[56] G. P. Salam, JHEP **07**, 019 (1998).

[57] G. P. Salam, Acta Phys. Polon. **B 30**, 3679 (1999).

[58] M. Ciafaloni, D. Colferai and G. P. Salam, Phys. Rev. **D 60**, 114036 (1999).

[59] M. Ciafaloni and D. Colferai, Phys. Lett. **B 452**, 372 (1999).

[60] H1 Coll., C. Adloff et al., Phys. Lett. **B 528**, 199 (2002).

[61] ZEUS Coll., J. Breitweg et al., Eur. Phys. J. **C 12**, 35 (2000).

[62] H. Jung, Comput. Phys. Commun. **143**, 100 (2002).

[63] H. Jung and G. P. Salam, Eur. Phys. J. **C 19**, 351 (2001).

[64] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 21**, 443 (2001).

[65] CDF Coll., D. Acosta et al., Phys. Rev. **D 71**, 032001 (2005).

[66] CDF Coll., D. Acosta et al., Phys. Rev. **D 71**, 092001 (2005).

[67] D0 Coll., B. Abbott et al., Phys. Lett. **B 487**, 264 (2000).

[68] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **07**, 033 (2004).

[69] D. Stump et al., JHEP **10**, 046 (2003).

[70] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C 4**, 463 (1998); A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Phys. Lett. **B 443**, 301 (1998).

[71] R. Vogt, Int. J. Mod. Phys. **E 12**, 211 (2003).

[72] R. Vogt, *In proceedings of the 18*$^{\rm th}$ *Winter Workshop on Nuclear Dynamics, edited by R. Bellwied et al., Nassau, The Bahamas, p. 253*. EP Systema, Debrecen, Hungary, 2002. [arXiv:hep-ph/0203151].

[73] Vogt, R., Z. Phys. **C 71**, 475 (1996).

[74] O. Behnke et al., *Experimental overview*. These proceedings.

[75] ALICE Coll., *Technical proposal*, 1995. CERN/LHCC 95-71.

[76] R. Vogt, Heavy Ion Phys. **17**, 75 (2003).

[77] N. Kidonakis, E. Laenen, S. Moch and R. Vogt, Phys. Rev. **D 67**, 074037 (2003).

[78] T. Sjostrand et al., Comput. Phys. Commun. **135**, 238 (2001).

[79] C. Peterson, D. Schlatter, I. Schmitt and P. M. Zerwas, Phys. Rev. **D 27**, 105 (1983).

[80] A. Dainese, R. Vogt, M. Bondila, K. J. Eskola and V. J. Kolhinen, J. Phys. **G 30**, 1787 (2004).

[81] M. L. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. **B 373**, 295 (1992).

[82] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Phys. Lett. **B 531**, 216 (2002).

Heavy quark fragmentation

J. Braciník[1], *M. Cacciari*[2], *M. Corradi*[3], *G. Grindhammer*[1]
[1]Max-Planck-Institut für Physik, München, Germany
[2]LPTHE - Université P. et M. Curie (Paris 6), Paris, France
[3]INFN Bologna, via Irnerio 46, Bologna, Italy

Abstract
The fragmentation of heavy quarks into hadrons is a key non-perturbative ingredient for the heavy quark production calculations. The formalism is reviewed, and the extraction of non-perturbative parameters from e^+e^- and from ep data is discussed.

coordinator: *M. Corradi*

1 Introduction

author: M. Cacciari

When we try to describe in QCD the production of a hadron we are always faced with the necessity to take into account the non-perturbative hadronization phase, i.e. the processes which transform perturbative objects (quarks and gluons) into real particles. In the case of light hadrons the QCD factorization theorem [1, 2, 3, 4, 5, 6] allows to factorize these non-perturbative effects into universal (but factorization-scheme dependent) *fragmentation functions* (FF):

$$\frac{d\sigma_h}{dp_T}(p_T) = \sum_i \int \frac{dx}{x} \frac{d\sigma_i}{dp_T}\left(\frac{p_T}{x};\mu\right) D_{i\to h}(x;\mu) + \mathcal{O}\left(\frac{\Lambda}{p_T}\right) . \tag{1}$$

In this equation, valid up to higher twist corrections of order Λ/p_T (Λ being a hadronic scale of the order of a few hundred MeV and p_T for instance a transverse momentum), the partonic cross sections $d\sigma_i/dp_T$ for production of the parton i are calculated in perturbative QCD, while the fragmentation functions $D_{i\to h}(x;\mu)$ are usually extracted from fits to experimental data. Thanks to their universality they can be used for predictions in different processes. The artificial factorization scale μ is a reminder of the non-physical character of both the partonic cross sections and the fragmentation functions: it is usually taken of the order of the hard scale p_T of the process, and the fragmentation functions are evolved from a low scale up to μ by means of the DGLAP evolution equations.

This general picture becomes somewhat different when we want to calculate the production of heavy-flavoured mesons. In fact, thanks to the large mass of the charm and the bottom quark, acting as a cutoff for the collinear singularities which appear in higher orders in perturbative calculations, one can calculate the perturbative prediction for heavy *quark* production. Still, of course, the quark $\to$ hadron transition must be described. Mimicking the factorization theorem given above, it has become customary to complement the perturbative calculation for heavy quark production with a non-perturbative fragmentation function accounting for its hadronization into a meson:

$$\frac{d\sigma_H}{dp_T}(p_T) = \int \frac{dx}{x} \frac{d\sigma_Q^{pert}}{dp_T}\left(\frac{p_T}{x},m\right) D^{np}_{Q\to H}(x) . \tag{2}$$

It is worth noting that at this stage this formula is not given by a rigorous theorem, but rather by some sensible assumptions. Moreover, it will in general fail (or at least be subject to large uncertainties) in the region where the mass m of the heavy quark is not much larger than its transverse momentum p_T, since the choice of the scaling variable is not unique any more, and $\mathcal{O}(m/p_T)$ corrections cannot be neglected.

Basic arguments in QCD allow to identify the main characteristics of the non-perturbative fragmentation function $D^{np}_{Q\to H}(x)$. In 1977 J.D. Bjorken [7] and M. Suzuki [8] independently argued that the average fraction of momentum lost by the heavy quark when hadronizing into a heavy-flavoured hadron is given by

$$\langle x\rangle^{np} \simeq 1 - \frac{\Lambda}{m}\,. \tag{3}$$

Since (by definition) the mass of a heavy quark is much larger than a hadronic scale Λ, this amounts to saying that the non-perturbative FF for a heavy quark is very hard, i.e. the quark loses very little momentum when hadronizing. This can also be seen with a very simplistic argument: a fast-traveling massive quark will lose very little speed (and hence momentum) when picking up from the vacuum a light quark of mass Λ to form a heavy meson[1].

This basic behaviour is to be found as a common trait in all the non-perturbative heavy quark FFs derived from various phenomenological models. Among the most commonly used ones we can cite the Kartvelishvili-Likhoded-Petrov [12], Bowler [13], Peterson-Schlatter-Schmitt-Zerwas [14] and Collins-Spiller [15] fragmentation functions. These models all provide some functional form for the $D^{np}_{Q\to H}(x)$ function, and one or more free parameters which control its hardness. Such parameters are usually not predicted by the models (except perhaps on an order-of-magnitude basis), and must be fitted to the experimental data.

During the '80s many such fits were performed, and these and similar functions were also included in many Monte Carlo event generators. Eventually, some 'best' set of parameter values (for instance for the PSSZ form) was determined [16] and subsequently widely used.

These first applications, given the limited accuracy of the available data, tended to overlook two aspects which have become more important in recent years, when the accuracy of the data has vastly improved:

- A non-perturbative FF is designed to describe the heavy quark $\to$ hadron transition, dealing with events mainly populated by soft gluons of energies of a few hundred MeV. However, if a heavy quark is produced in a high energy event it will initially be far off shell: perturbative hard gluons will be emitted to bring it on-shell, reducing the heavy quark momentum and yielding in the process large collinear logarithms (for instance of the form $\alpha_s^n \log^n(p_T/m)$ in a transverse momentum differential cross section). Of course, the amount of gluon radiation is related to the distance between the heavy quark mass scale and the hard scale of the interaction, and is therefore process-dependent. One can (and it was indeed done) either fit different free parameters at different centre-of-mass energies (or transverse momenta), or try to evolve directly the non-perturbative FF by means of the DGLAP equations, hence including into it the perturbative collinear logarithms. However, this is not what non-perturbative fragmentation functions are meant for, and doing so spoils the validity of the relation in eq. (3).
- Since only the final heavy hadron is observed, both the non-perturbative FF and the perturbative cross section for producing the heavy quark must be regarded as non-physical objects. The details of the fitted non-perturbative FF (e.g. the precise value(s) of its free parameter(s)) will depend on those of the perturbative cross sections: different perturbative calculations (leading order, next-to-leading order, Monte Carlo, ...) and different perturbative parameters (heavy quark masses, strong coupling, ...) will lead to different non-perturbative FFs. These in turn will have to be used only with a perturbative description similar to the one they have been determined with. Hence the limited accuracy (and hence usefulness) of a 'standard' determination of the parameters [17].

The first point was addressed by Mele and Nason in a paper [18] which deeply changed the field of heavy quark fragmentation, and essentially propelled it into the modern era. Mele and Nason observed that, in the limit where one neglects heavy quark mass terms suppressed by a large energy scale, a heavy quark cross section can be factored into a massless, $\overline{MS}$-subtracted cross section for producing

[1] More modern and more rigorous derivations of this result can be found in [9, 10, 11].

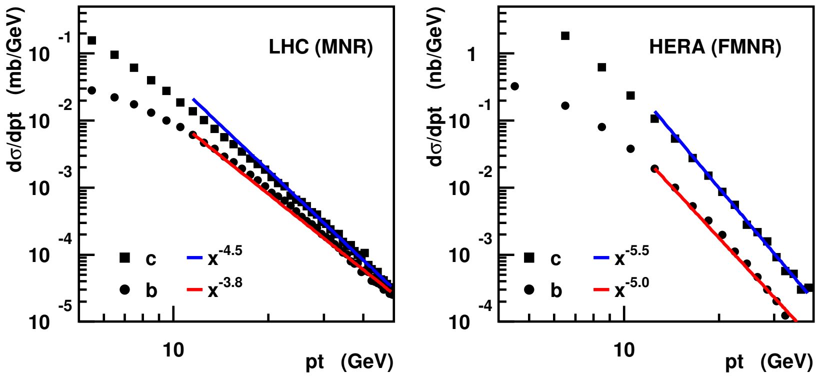


Fig. 1: Power-law fits to the heavy quark p_T distributions at LHC (left) and HERA (right) obtained with the NLO programs MNR and FMNR. The resulting exponents are $N = 4.5/3.8$ for charm/beauty at LHC and $N = 5.5/5.0$ for *c*/*b* at HERA.

a light parton, and a process-independent[2], *perturbative* heavy quark fragmentation function describing the transition of the massless parton into the heavy quark:

$$\frac{d\sigma_Q^{pert,res}}{dp_T}(p_T, m) = \sum_i \int \frac{dx}{x} \frac{d\sigma_i}{dp_T}\left(\frac{p_T}{x};\mu\right) D_{i\to Q}(x;\mu,m) + \mathcal{O}\left(\frac{m}{p_T}\right) . \tag{4}$$

The key feature of this equation is that it is entirely perturbative: every term can be calculated in perturbative QCD. The perturbative fragmentation functions $D_{i\to Q}(x;\mu,m)$ (not to be confused with the non-perturbative one $D^{np}_{Q\to H}(x)$) can be evolved via DGLAP equations from an initial scale of the order of the heavy quark mass up to the large scale of the order of p_T. This resums to all orders in the strong coupling the collinear logarithms generated by the gluon emissions which bring the heavy quark on its mass shell, leading to a more accurate theoretical prediction for $d\sigma_Q/dp_T$.

Once a reliable perturbative cross section for the production of a heavy quark is established, one is simply left with the need to account for its hadronization. For this purpose one of the functional forms listed above can be used for the non-perturbative FF, and implemented as in eq. (2), but using the improved, resummed cross section given by eq. (4). Since most of the the scaling-violation logarithms are accounted for by the evolution of the perturbative FF, the non-perturbative one can now be scale-independent and only contain the physics related to the hadronization of the heavy quark. It will always, however, depend on the details of the perturbative picture used.

2 Extraction of heavy quark fragmentation parameters from e^+e^- and their impact on HERA and LHC

author: M. Corradi

2.1 Importance of $\langle x\rangle^{\rm np}$

According to the factorization of the fragmentation functions (FF), the differential cross section $d\sigma/dp_T$ for the production of a heavy hadron H can be written as the convolution of the perturbative heavy quark differential cross section $d\sigma^{\rm pert}/dp_T$ and the non-perturbative fragmentation function $D^{\rm np}(x)$:

$$\frac{d\sigma}{dp_T}(p_T) = \int \frac{dx}{x} D^{\rm np}(x) \frac{d\sigma^{\rm pert}}{dp_T}\left(\frac{p_T}{x}\right) . \tag{5}$$

This convolution neglects mass terms $\mathcal{O}(m_Q/p_T)$ and non-perturbative terms $\mathcal{O}(\Lambda_{\rm qcd}/m_Q)$.

[2]Mele and Nason extracted this function from the e^+e^- cross section, convincingly conjecturing its process independence, which was successively established on more general grounds in [19]

Function	$D(x)$	parameters	x range
Kartvelishvili	$(1-x)x^{\alpha}$	$\alpha = 2/\delta - 3$	$[0, 1]$
Peterson	$\frac{1}{x}\left(1-\frac{1}{x}-\frac{\epsilon}{(1-x)}\right)^{-2}$	ϵ	$[0, 1]$
Gauss	$\exp(-(\frac{x-\mu}{2\sigma})^2)$	$\mu=1-\delta$ $\sigma=\delta/2$	$[-\infty, \infty]$
Box	const.	-	$[1-2\delta, 1]$
Triangular:	$x - x_0$	$x_0 = 1 - 3\delta$	$[1-3\delta, 1]$

Table 1: Test functions used in Fig. 2. The functions assume a value different from zero in the range given by the third column.

The heavy quark p_T distribution behaves at large p_T like a power law $d\sigma^{\rm pert}/dp_T = Cp_T^{-N}$. Figure 1 shows power-law fits to the p_T distributions of heavy quarks at LHC and in photoproduction at HERA as obtained from the NLO programs MNR [20] and FMNR [21]. For $p_T > 10$ GeV N was found to range from 3.8 (b at LHC) to 5.5 (c at HERA). Combining this power-law behavior with eq. (5), the hadron p_T distribution is given by

$$\frac{d\sigma}{dp_T}(p_T) = \int dx\, x^{N-1}\, D^{\rm np}(x)\, Cp_T^{-N} = \frac{d\sigma^{\rm pert}}{dp_T}\hat{D}_N^{\rm np}, \tag{6}$$

where $\hat{D}_N^{\rm np} = \int dx\, x^{N-1}\, D^{\rm np}(x)$ is the N^{th} Mellin moment of the non-perturbative FF.

The hadron distribution is therefore governed by the 4^{th} - 5^{th} Mellin moments of $D^{\rm np}(x)$. It is interesting to translate the Mellin moments into more intuitive central moments

$$\mu_n = \int dx\, (x - \langle x\rangle)^n\, D^{\rm np}(x) \quad \text{for} \quad n \geq 2 \tag{7}$$

where $\langle x\rangle = \int dx\, x D^{\rm np}(x)$ is the mean value. The first Mellin moments, written in terms of $\langle x\rangle$ and μ_n, are: $\hat{D}_1 = 1$, $\hat{D}_2 = \langle x\rangle$, $\hat{D}_3 = \langle x\rangle^2 + \mu_2$, $\hat{D}_4 = \langle x\rangle^3 + 3\mu_2\langle x\rangle + \mu_3$.

In heavy quark fragmentation, the mean value of $D^{\rm np}(x)$ can be written as $\langle x\rangle = 1 - \delta$ where $\delta = \mathcal{O}(\Lambda_{\rm qcd}/m_Q)$ is small [11]. For any positive function with $\langle x\rangle = 1 - \delta$, defined in the interval $[0, 1]$, the central moments are limited by δ, $|\mu_n| \leq \delta$. In practice, reasonable heavy quark fragmentation functions are concentrated in a small region around $1 - \delta$ and therefore the higher central moments are small. To be specific, if the function is different from zero in a region of size $\pm K\delta$ (with $K = \mathcal{O}(1)$) around $1 - \delta$ then $|\mu_n| \leq (K\delta)^n$. This means that the Mellin moments of reasonable FFs are given, to a good approximation, by the mean value to the $N - 1$ power:

$$\hat{D}_N = \langle x\rangle^{N-1} + \mathcal{O}(\delta^2). \tag{8}$$

The expansion to δ^2 involves the second central moment μ_2: $\hat{D}_N = \langle x\rangle^{N-1} + \frac{(N-1)!}{2(N-3)!}\mu_2\langle x\rangle^{N-3} + \mathcal{O}(\delta^3)$. For a reasonable FF and a perturbative distribution falling with the power $-N$, Eq. 6 and 8 give

$$\frac{d\sigma}{dp_T}(p_T) = \frac{d\sigma^{\rm pert}}{dp_T^Q}(p_T)\ \left(\langle x\rangle^{\rm np}\right)^{N-1} + \mathcal{O}(\delta^2). \tag{9}$$

Therefore the effect of the non-perturbative FF is to introduce a shift in the normalisation that depends on the average x, while the details of the shape of $D(x)$ have negligible effect. To check that this reasoning works with realistic fragmentation functions and realistic perturbative p_T distributions, various functions with the same $\langle x\rangle$ but different shapes have been tested in convolution with the perturbative p_T spectrum for b production at LHC obtained with the NLO program MNR. The test functions considered are the Peterson [14] and Kartvelishvili [12] fragmentation functions, a Gaussian distribution with $\sigma = 1 - \mu$, a flat and a triangular distribution. Table 1 gives more detail about these functions. Three average values were chosen: $\langle x\rangle = 0.9, 0.8, 0.666$. Figure 2 shows the result of this test. For each average value, the convolutions are very similar, even if the test functions are very different. For $\langle x\rangle = 0.9, 0.8$, which are typical beauty or charm values, the hadron spectra agree within few %. For the extreme value of

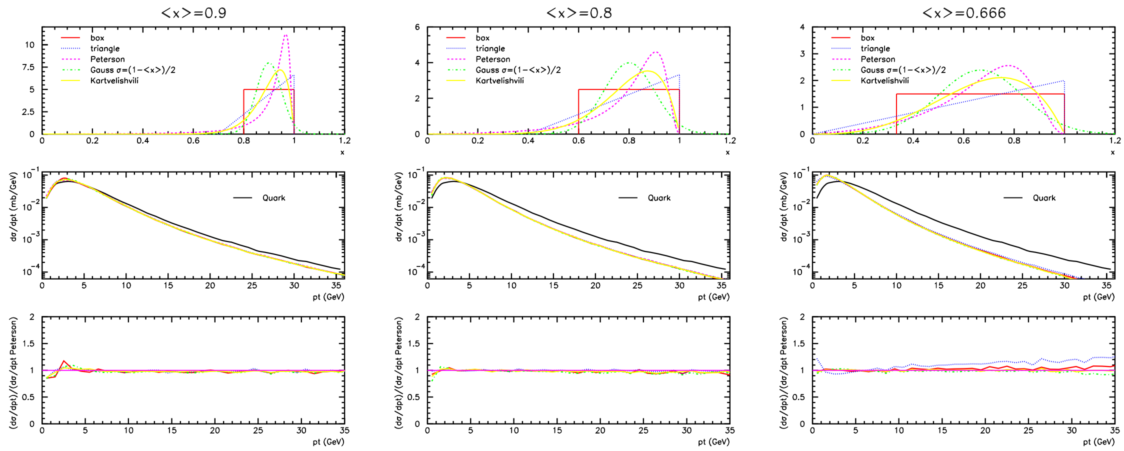


Fig. 2: Effect of the convolution of the heavy quark transverse momentum distribution with different test functions for different values of the non-perturbative FF $\langle x\rangle = 0.9$ (left), 0.8 (center), 0.666 (right). For each $\langle x\rangle$, the upper plot shows the test functions, the middle plot shows the perturbative p_T distribution obtained with the MNR program for beauty at LHC and the hadron p_T distributions after the convolution with the test functions. The lower plot shows the ratio of the different hadron p_T distributions to the result obtained with the Peterson one.

$\langle x\rangle = 0.666$, the results for the Peterson and Kartvelishvili functions give very similar hadron spectra while the less realistic Gaussian and Box shapes differ at most 10% from Peterson at large p_T and the extreme Triangular function shows deviations up to $\sim 20\%$.

In conclusion the relevant fragmentation parameter for the inclusive hadron spectra at pp and ep colliders is the mean value $\langle x\rangle^{\rm np}$ of the non-perturbative FF. The next part will discuss, on the basis of e^+e^- data, what values of $\langle x\rangle^{\rm np}$ are relevant for different calculations.

2.2 Extraction of $\langle x\rangle^{\rm np}$ from e^+e^- data

In e^+e^- interactions it is convenient to express the factorization ansatz, given for the heavy-hadron p_T in eq. (5), in terms of the heavy-hadron momentum normalized to the maximum available momentum: $x_p = p^H/p^H_{\rm max}$, where $p^H_{\rm max} = \sqrt{(\frac{1}{2}E_{\rm cms})^2 - m_H^2}$:

$$\frac{d\sigma}{dx_p}(x_p) \;=\; \int \frac{dx}{x}\, D^{\rm np}(x)\, \frac{d\sigma^{\rm pert}}{dx_p}(\frac{x_p}{x})$$

which corresponds to the following relation for the mean values: $\langle x_p\rangle = \langle x\rangle^{\rm np}\langle x\rangle^{\rm pert}$ where $\langle x_p\rangle$ is the mean hadron x_p, $\langle x\rangle^{\rm np}$ is the mean value of the non-perturbative FF and $\langle x\rangle^{\rm pert} = \int\, dx\; x\; \frac{d\sigma^{\rm pert}}{dx_p}$ is the mean value of the perturbative distribution. Then, taking $\langle x_p\rangle$ from experimental data and $\langle x\rangle^{\rm pert}$ from a particular perturbative calculation, it is possible to extract the value of $\langle x\rangle^{\rm np}$ valid for that calculation as

$$\langle x\rangle^{\rm np} = \langle x_p\rangle/\langle x\rangle^{\rm pert}. \tag{10}$$

Two perturbative calculations will be considered to extract $\langle x\rangle^{\rm pert}$: a fixed-order (FO) next-to-leading order (i.e. $\mathcal{O}(\alpha_S)$) calculation and a calculation that includes also the resummation of next-to-leading logarithms (NLL) and Sudakov resummation, both obtained with the HVQF program [19]. From the point of view of fragmentation, the FO calculation only considers the emission of a gluon from one of the two heavy quarks generated in the e^+e^- collision while the NLL calculation includes the evolution of the FF from the hard interaction scale down to the scale given by the heavy quark mass. The parameters used for the FO and NLL models are $m_c = 1.5$ GeV, $m_b = 4.75$ GeV, $\Lambda_{\rm QCD} = 0.226$ GeV and the

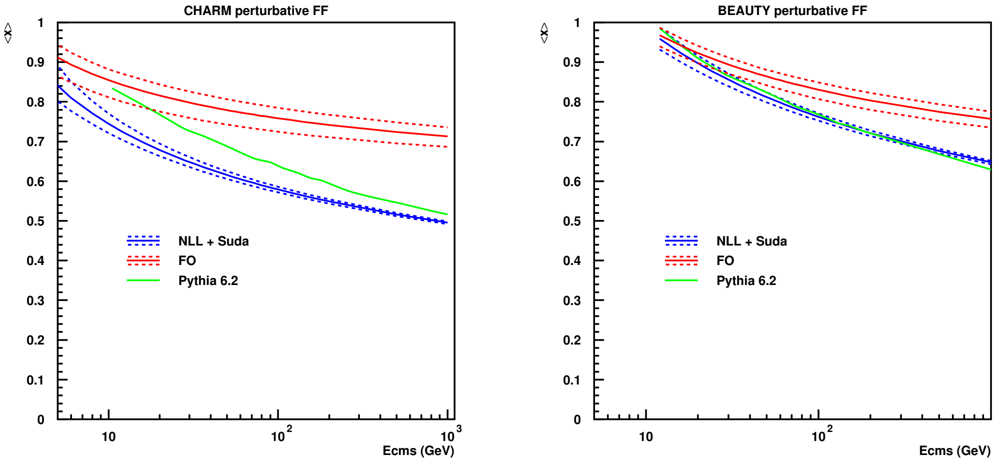


Fig. 3: Average fragmentation function from the perturbative calculations for charm (left) and beauty (right) as a function of the e^+e^- center of mass energy.

Charm (D^*) measurement	$E_{\rm cms}$ (GeV)	Measured variable	Value	ISR corr. (%)	$x_E \to x_p$ (%)	$\langle x_p \rangle^{\rm corr}$
OPAL [22]	92	$\langle x_E \rangle$	$0.516^{+0.008}_{-0.005} \pm 0.010$	$+0.4$	-0.4	0.516 ± 0.012
ALEPH [23]	92	$\langle x_E \rangle$	$0.4878 \pm 0.0046 \pm 0.0061$	$+0.4$	-0.4	0.488 ± 0.008
DELPHI [24]	92	$\langle x_E \rangle$	$0.487 \pm 0.015 \pm 0.005$	$+0.4$	-0.4	0.487 ± 0.016
TASSO [25]	36.2	$\langle x_E \rangle$	0.58 ± 0.02	$+6.7$	-1.8	0.61 ± 0.02
DELCO [26]	29	$\epsilon^*_{\rm Pet.}$	$0.31^{+0.10}_{-0.08}$	$+6.3$	-	0.55 ± 0.03
ARGUS [27]	10.5	$\langle x_p \rangle$	0.64 ± 0.03	$+4.2$	-	0.67 ± 0.03
CLEO [28]	10.5	$\langle x_p \rangle$	$0.611 \pm 0.007 \pm 0.004$	$+4.2$	-	0.637 ± 0.008
BELLE [29]	10.58	$\langle x_p \rangle$	$0.61217 \pm 0.00036 \pm 0.00143$	$+4.2$	-	0.6379 ± 0.0016

Beauty (B^{wd}) measurement	$E_{\rm cms}$ (GeV)	Measured variable	Value	ISR corr. (%)	$x_E \to x_p$ (%)	$\langle x_p \rangle^{\rm corr}$
OPAL [30]	92	$\langle x_E \rangle$	$0.7193 \pm 0.0016 \pm 0.0038$	$+0.3$	-0.9	0.715 ± 0.004
SLD [31]	92	$\langle x_E \rangle$	$0.709 \pm 0.003 \pm 0.005$	$+0.3$	-0.9	0.705 ± 0.006
ALEPH [32]	92	$\langle x_E \rangle$	$0.716 \pm 0.006 \pm 0.006$	$+0.3$	-0.9	0.712 ± 0.008
DELPHI [33]	92	$\langle x_E \rangle$	$0.7153 \pm 0.0007 \pm 0.0050$	$+0.3$	-0.9	0.711 ± 0.005
JADE [34]	36.2	$\langle x_E \rangle$	$0.76 \pm 0.03 \pm 0.04$	$+5.4$	-3.5	0.77 ± 0.06
DELCO [35]	29	$\langle x_E \rangle$	0.72 ± 0.05	$+4.8$	-4.7	0.72 ± 0.05
PEP4-TPC [36]	29	$\langle x_E \rangle$	$0.77 \pm 0.04 \pm 0.03$	$+4.8$	-4.7	0.77 ± 0.07

Table 2: Experimental results on the average fragmentation function in e^+e^- collisions for D^* mesons and weakly decaying beauty hadrons. The table reports, for each experiment, the published variable and the corrections applied to obtain $\langle x_p \rangle^{\rm corr}$. All the measurement have been corrected for initial state radiation (ISR). Measurements reported in terms of $\langle x_E \rangle$ have been corrected to $\langle x_p \rangle$. In the case of ARGUS the average has been calculated from the full distribution. In the case of TASSO the error on the average was re-evaluated using the full distribution since the published error seems incompatible with the data. DELCO reports a fit with a Peterson distribution that has been translated into $\langle x_p \rangle^{\rm corr}$. Systematical and statistical uncertainties, where reported separately, have been added in quadrature. The ALEPH beauty measurement refers to B^+ and B^0 mesons only (i.e. excluding B_s and Λ_b), a MC study shows that this correspond to underestimating $\langle x_p \rangle^{\rm corr}$ by $\sim 0.1\%$ only, which is negligible.

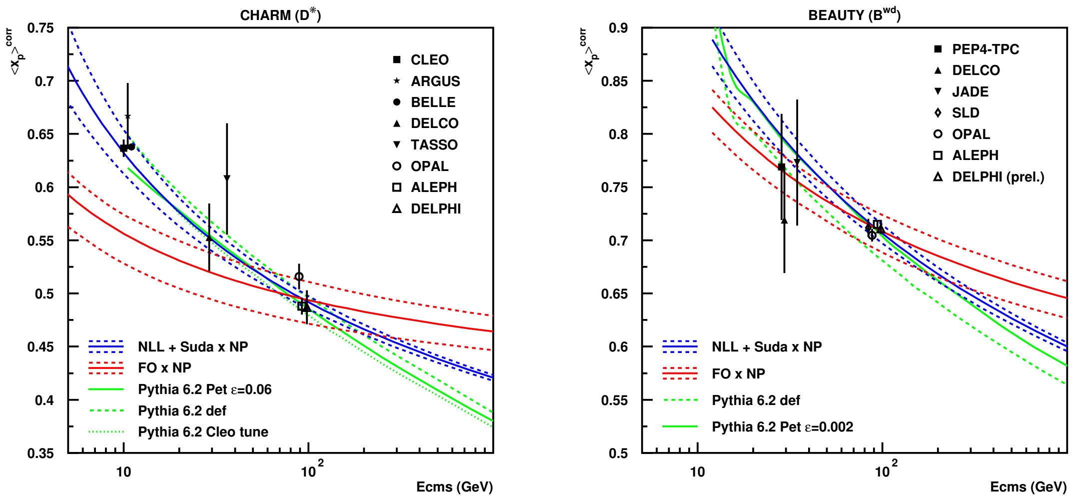


Fig. 4: Average fragmentation function as a function of the center of mass energy for charm (left) and beauty (right). The plots show the experimental results and the curves from NLL and FO theory with a non-perturbative fragmentation obtained using the data at the Z^0 energy. The curves from PYTHIA 6.2 with different fragmentation choices are also shown. The experimental points at the $\Upsilon(4s)$ and Z^0 resonances are shown slightly displaced in the horizontal axis for better legibility.

renormalisation and factorization scales $\mu_R = \mu_F = E_{\rm cms}$. The starting scale for FF evolution in the NLL model was chosen to be m_Q. The theoretical uncertainty was obtained by varying independently the normalisation and factorization scales by a factor 2 and $1/2$ and taking the largest positive and negative variations as the uncertainty.

The experimental data are also compared to the PYTHIA 6.2 Monte Carlo program [37] which contains an effective resummation of leading-logarithms based on a parton-shower algorithm and which is interfaced to the Lund fragmentation model. In this case the MC model gives directly $\langle x_p \rangle$, while $\langle x \rangle^{\rm pert}$ has been obtained taking the heavy quark at the end of the parton shower phase. The quark masses have been set to $m_c = 1.5$ GeV and $m_b = 4.75$ GeV, and all the parameters were set to the default values except for specific fragmentation parameters explained below. Three sets of fragmentation parameters were chosen for charm: the default fragmentation (Lund-Bowler), a longitudinal string fragmentation of the Peterson form with $\epsilon = 0.06$ (`MSTJ(11)=3, PARJ(54)=-0.06`) and the Lund-Bowler fragmentation with parameters re-tuned by the CLEO collaboration [28] (`PARJ(41)=0.178, PARJ(42)=0.393, PARJ(13)=0.627`). The two sets chosen for beauty are the default Lund-Bowler fragmentation and the Peterson fragmentation with $\epsilon = 0.002$ (`MSTJ(11)=3, PARJ(55)=-0.002`). Figure 3 shows $\langle x \rangle^{\rm pert}$ from the perturbative calculations as a function of the centre of mass energy $E_{\rm cms}$ for charm and bottom.

2.3 Charm

Charm fragmentation data are available from various e^+e^- experiments. The most precise are those at the Z^0 pole at LEP (ALEPH [23], OPAL [22], DELPHI [24]) and near the $\Upsilon(4s)$ (ARGUS [27], CLEO [28], BELLE [29]). Less precise data are available in the intermediate continuum region from DELCO [26] at PEP and TASSO [25] at PETRA. Measurements in which the beauty component was not subtracted have been discarded [38, 39, 40]. The experimental data are reported in Table 2. Only measurements relative to the $D^{*\pm}(2010)$ meson are considered, to avoid the complications due to cascade decays that are present for ground state mesons. Charm quarks originating from gluon splitting rather than from the virtual boson from e^+e^- annihilation may be relevant at LEP energies. This contribution is anyway already subtracted in the published data considered here, and it is consistently not considered in the perturbative calculations. Most of the experiments published the mean value of the x distribution. The only exception is ARGUS, for which the mean value was computed from the published distribution.

Some of the experiments give the results directly in terms of x_p, others in terms of the energy fraction $x_E = 2E_H/E_{\text{cms}}$. The latter has been corrected to x_p using the PYTHIA MC. The difference between $\langle x_p \rangle$ and $\langle x_E \rangle$ can be as large as 12% at $E_{\text{cms}} = 10.5$ GeV and reduces to less than 1% at $E_{\text{cms}} = 92$ GeV. Since the low-mass measurements are already given in terms of x_p, the applied corrections from x_E to x_p was always small. QED corrections are also needed to compare the experimental data to the QCD predictions. The initial state radiation (ISR) from the electrons has the effect of reducing the energy available for the e^+e^- annihilation and therefore to reduce the observed value of $\langle x_p \rangle$. A correction, obtained by comparing the PYTHIA MC with and without ISR, was applied to the data to obtain $\langle x_p \rangle^{\text{corr}}$. The correction is $\sim 4\%$ at $E_{\text{cms}} = 10.5$ GeV, is largest in the intermediate region and is negligible at $E_{\text{cms}} = 92$ GeV.

Only LEP data at $E_{\text{cms}} = 92$ GeV were used to extract $\langle x \rangle^{\text{np}}$ since the factorization of the non-perturbative FF could be spoiled by large $\mathcal{O}(m_Q/E_{\text{cms}})$ terms at lower energies. Table 3 reports the LEP average $\langle x_p \rangle^{\text{corr}}$, the perturbative results at 92 GeV and the resulting $\langle x \rangle^{\text{np}}$ for NLL and FO calculations as well as $\langle x \rangle$ and $\langle x \rangle^{\text{pert}}$ from PYTHIA. Figure 4 (left) shows $\langle x_p \rangle$ obtained by multiplying the perturbative calculations with the corresponding $\langle x \rangle^{\text{np}}$, compared to the experimental data and to the PYTHIA MC with different fragmentation parameters.

With the non-perturbative $\langle x \rangle^{\text{np}} = 0.849 \pm 0.018$ obtained at LEP energies, the NLL calculation can reproduce all the data within a quite small theoretical uncertainty. The FO calculation is instead too flat to reproduce the data even considering its large theoretical uncertainty band. The non-perturbative fragmentation $\langle x \rangle^{\text{np}}$ obtained at LEP energy for the FO calculation is quite small (0.65 ± 0.04) since it compensates the effect of the FF evolution that is missing in the perturbative part. Therefore FO calculations with $\langle x \rangle^{\text{np}}$ extracted at LEP energy undershoot drastically the data at the $\Upsilon(4s)$.

The PYTHIA MC with the Lund-Bowler fragmentation reproduces the data reasonably well. The result with default parameters is slightly above the data while the result with the parameters tuned by the CLEO collaboration is slightly below. Both are compatible within the experimental uncertainty with all the experimental values with the exception of the very precise measurement from Belle from which they differ anyway by less than 2%. PYTHIA with the Peterson fragmentation with $\epsilon = 0.06$ reproduces well the LEP data but is too low at lower energies.

2.4 Beauty

In the case of beauty we consider fragmentation measurements for the mix of weakly decaying hadrons B^{wd}. Precise measurements are available only at the Z^0 peak (SLD [31], ALEPH [32], OPAL [30], DELPHI [33]). Lower energy measurements from PEP (PEP4-TPC [36], DELCO [35]) and PETRA (JADE [34]) have larger uncertainties. As for charm, corrections have been applied for ISR and to convert $\langle x_E \rangle$ to $\langle x_p \rangle$. The data are shown in Table 2 and the results in Table 3 and Figure 4 (right). Since precise data are available only at a single energy, it is impossible to test the energy beaviour of the theoretical predictions. As in the charm case, the energy dependence of PYTHIA and NLL theory are similar, while the FO prediction is much more flat, suggesting that also for beauty the non-perturbative fragmentation obtained for FO at the Z^0 could not be applied at lower energy. PYTHIA with Peterson fragmentation with $\epsilon = 0.002$ reproduces the data, while the default Lund-Bowler fragmentation is too soft.

2.5 Effect on predictions for heavy quark production at HERA and LHC

Going back to the heavy-hadron production in ep and pp collisions, eq. 9 shows that the uncertainty on the differential heavy-hadron cross section $d\sigma/dp_T$ is related to the uncertainty on the average non-perturbative fragmentation by

$$\Delta(d\sigma/dp_T) = N\Delta(\langle x \rangle^{\text{np}}),$$

where $-N$ is the exponent of the differential cross section.

Charm (D^*) @ 92 GeV	$\langle x_p \rangle^{\rm corr}$	$\langle x_p \rangle^{\rm pert}$	$\langle x_p \rangle^{\rm np}$
Data	0.495 ± 0.006	-	-
NLL	-	0.583 ± 0.007	0.849 ± 0.018
FO	-	0.76 ± 0.03	0.65 ± 0.04
PYTHIA Def.	0.500	0.640	-
PYTHIA CLEO	0.484	0.640	-
PYTHIA Pet. $\epsilon = 0.06$	0.490	0.640	-

Beauty (B^{wd}) @ 92 GeV	$\langle x_p \rangle^{\rm corr}$	$\langle x_p \rangle^{\rm pert}$	$\langle x_p \rangle^{\rm np}$
Data	0.7114 ± 0.0026	-	-
NLL	-	0.768 ± 0.010	0.927 ± 0.013
FO	-	0.83 ± 0.02	0.85 ± 0.02
PYTHIA Def.	0.686	0.773	-
PYTHIA Pet. $\epsilon = 0.002$	0.710	0.773	-

Table 3: Average fragmentation functions at the Z^0 resonance for charm (top) and beauty. The table shows the average of the experimental data, the results from the NLL and FO calculations and from the PYTHIA MC with different fragmentation parameters. For the NLL and FO calculations $\langle x_p \rangle^{\rm np}$ is obtained by dividing the average from the experimental data by the perturbative result $\langle x_p \rangle^{\rm np} = \langle x_p \rangle^{\rm corr} / \langle x_p \rangle^{\rm pert}$.

The state of the art calculations for photo- and hadro-production (FONLL [41, 42]) include NLO matrix elements and the resummations of next-to-leading logarithms. The appropriate non-perturbative fragmentation for FONLL is therefore obtained with the NLL theory which has the same kind of perturbative accuracy [43]. Since the NLL calculation gives a good description of e^+e^- data, it seems appropriate to use the value and the uncertainty of $\langle x \rangle^{\rm np}$ as obtained from e^+e^- data at the Z^0 peak. The relative error for the D^* fragmentation is $\Delta\langle x \rangle^{\rm np} / \langle x \rangle^{\rm np} = 2\%$ which translates into an uncertainty of 9% on charm production at large p_T at LHC ($N = 4.5$) of 9%. For beauty, the relative uncertainty $\Delta\langle x \rangle^{\rm np} / \langle x \rangle^{\rm np} = 1.4\%$ translates into an uncertainty on large-p_T B-hadron production at LHC ($N = 3.8$) of 5.3%. These uncertainty are smaller or of the order of the perturbative uncertainties of the calculation. Nevertheless, it should be noted that this approach is only valid for large transverse momenta. At small transverse momenta the factorization ansatz breaks down and large corrections of order m_Q/p_T may appear. Therefore, for the low-p_T region, the uncertainty on the p_T distribution is large and difficult to evaluate.

For processes such as DIS and for particular observables FONLL calculations are not available. The best theory available in this case is the fixed order NLO theory. In this case the situation is complex since the equivalent FO calculation for e^+e^- does not reproduce the experimental data. The proposed solution is to vary $\langle x \rangle^{\rm np}$ from the same value obtained in the NLL case (that would be correct at low p_T, where the FF evolution is irrelevant) to the value obtained at the Z^0 energy (that would be valid at $p_T \sim 100$ GeV). Therefore we consider for charm $\langle x \rangle^{\rm np} = 0.075 \pm 0.010$ and for beauty $\langle x \rangle^{\rm np} = 0.089 \pm 0.004$. When these values are transported to heavy-hadron production at LHC, the corresponding uncertainties on $d\sigma/dp_T$ at large p_T are 60% for charm and 20% for beauty. Therefore the NLO fixed order calculations cannot be used for precise predictions of the charm (and to a lesser extent beauty) production at pp and ep colliders.

In the FO-NLO and FONLL programs the hadron distributions are obtained by reducing the quark momenta according to a given fragmentation functions. Typical fragmentation functions used in these programs are the Peterson and Kartvelishvili forms. Table 4 summarises the proposed values and uncertainties for $\langle x \rangle^{\rm np}$ to be used with FO-NLO and FONLL calculations and reports the corresponding values and ranges for the Peterson and Kartvelishvili parameters. Similar ranges are used in the calculations presented in the section on "Benchmark cross sections" in these proceedings.

	$\langle x^{\rm np}\rangle$	ϵ(min : max)	α(min : max)	$\Delta\langle x^{\rm np}\rangle/\langle x^{\rm np}\rangle$	$\Delta\sigma/\sigma$ (LHC)	$\Delta\sigma/\sigma$ (HERA)
FONLL D^*	0.849 ± 0.018	0.0040(0.0027 : 0.0057)	10(9 : 12)	2.1%	9%	12%
FONLL B^{wd}	0.927 ± 0.013	0.00045(0.00026 : 0.00072)	24(20 : 30)	1.4%	5%	7%
FO-NLO D^*	0.75 ± 0.10	0.02(0.004 : 0.08)	5(3 : 10)	13%	60%	70%
FO-NLO B^{wd}	0.89 ± 0.04	0.0015(0.0004 : 0.004)	15(10 : 25)	4.5%	20%	22%

Table 4: Proposed value and uncertainty on $\langle x\rangle^{\rm np}$ to be used with FO-NLO and FONLL programs for photo- and hadro-production of D^* mesons and weakly decaying B hadrons. The corresponding value and range for the Peterson ϵ and for the Kartvelishvili α parameters are also reported. The last columns show the corresponding relative uncertainty on $d\sigma/dp_T$ at LHC (assuming a negative power $N = 4.5/3.8$ for charm/beauty) and HERA ($N = 5.5/5.0$ for c/b).

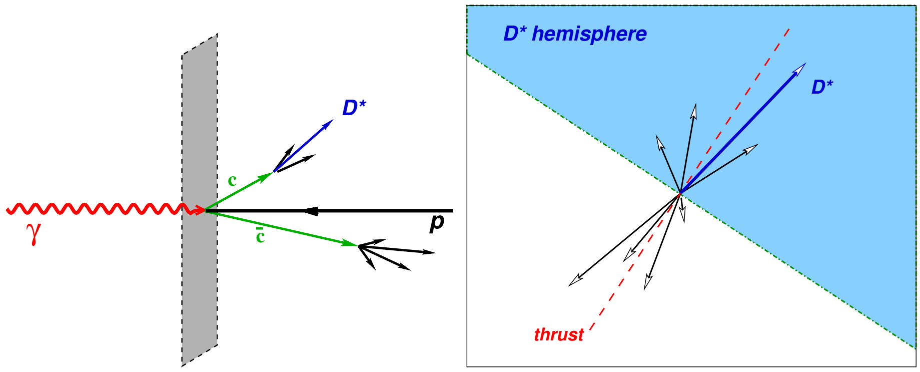


Fig. 5: Hemisphere method

3 Measurements of the charm quark fragmentation function at HERA

authors: J. Bracinı́k and G. Grindhammer

The differential cross section for the inclusive production of a heavy hadron H from a heavy quark h can be computed in perturbative QCD (pQCD) as a convolution of a short-distance cross section $\hat{\sigma}(p_h)$ with a fragmentation function $D^h_H(z)$:

$$d\sigma(p_H) = \int dz dp_h d\hat{\sigma}(p_h) D^h_H(z)\delta(p - zp_h) \tag{11}$$

The quantity z is the fractional momentum of the heavy quark h which is transferred to the heavy hadron H, and the normalized fragmentation function $D^h_H(z)$ gives the probability to observe the hadron H with a momentum fraction z.

The precise definition of $D^{(h)}_{\rm H}(z)$ is in some sense arbitrary. Due to the short and long-distance processes involved, the fragmentation function contains a perturbative and a non-perturbative component. Since the former can be calculated only up to some order in the strong coupling, the non-perturbative component in practice will have to absorb some of the missing higher order corrections. The calculable perturbative part can be absorbed into the definition of $\hat{\sigma}(p_h)$. Since for heavy quarks perturbative gluon emission do not lead to collinear divergencies, the perturbative evolution is well defined, and it is possible to absorb them into $\hat{\sigma}(p_h)$ and to perform perturbative evolution down to a scale of the heavy quark mass m_h. In this case the non-perturbative fragmentation function $D^h_H(z)$ accounts for the transition of an almost on-shell quark h into a heavy hadron H.

According to the QCD factorization theorem, the non-perturbative fragmentation functions (FF) depend neither on the type of the hard process nor on the scale at which the heavy quark h is originally produced. This implies universality of FF and allows - if valid - to extract fragmentation functions from data for one particular reaction (usually e^+e^- annihilation) and to use them to predict cross sections for other reactions (e.g. in pp and ep-collisions). In order to be able to check the reliability of pQCD predictions, it is necessary to check the universality of FF.

In practice, different theoretically motivated functional forms for $D_H^h(z)$ are used, depending on one more free parameters which are fitted to data. Among frequently used expressions are those by Peterson et al. [14] and by Kartvelishvili et al. [12].

From Equation 11 it is clear that $D_H^h(z)$ cannot be measured directly, since all observables are convoluted with the perturbative cross section. In case of ep and pp scattering there are additional convolutions with the parton density functions of one or two interacting hadrons. However, there are some observables which are more sensitive to $D_H^h(z)$ then others.

In e^+e^-, a convenient way to study fragmentation is to study the differential cross section of a heavy meson as a function of a scaled momentum or energy z. A customary experimental definition[3] of z is $\mathrm{z} = E_H/E_{\mathrm{beam}}$, where E_{beam} is the energy of the beams in the center-of-mass system. In leading order, i.e. without gluon emissions, it is also the energy of the charm and anti–charm quark and is equal to $D_H^h(z)$. In contrast to e^+e^- annihilation the choice of a fragmentation observable in ep collisions is more difficult. Two different observables have been used so far, both of them having the feature that in leading order QCD, the z-distributions are equal to $D_H^h(z)$.

In the case of what is called here the jet method, the energy of the charm quark is approximated by the energy of the charm-jet, tagged by a D^*-meson, which is considered to be part of the jet. The scaling variable is then defined as $\mathrm{z}_{\mathrm{jet}} = (E + p_{\mathrm{L}})_{D^*}/(E + p)_{\mathrm{jet}}$.

The idea of the so called hemisphere method (see Figure 5) is to exploit the special kinematics of charm events in ep collisions. The dominant charm production process has been shown to be boson-gluon fusion. If such an event is viewed in the photon-proton center-of-mass frame, the photon puts its full energy into the hard subprocess, while the proton interacts via a gluon, which typically carries only a small fraction of the proton momentum. As the result, both quarks produced, c and $\bar{c}$, move in the direction of the photon. Assuming no initial gluon k_{T} and no gluon radiation, their transverse momenta are balanced (see Fig. 5, left).

This can be seen best by projecting the quark momenta onto the plane perpendicular to the γ-p axis. In this plane it is possible to distinguish rather efficiently between the products of the fragmentation of the charm quark and its antiquark. The momenta of all particles are projected onto the plane and the thrust axis in this plane is found (see Fig. 5, right). The plane is then divided into two hemispheres by the line perpendicular to the thrust axis. All particles, lying in the hemisphere containing the D^*-meson are marked and their three-momenta and energy are summed-up to give the hemisphere's momentum and energy, which is used to approximate the momentum and energy of the respective charm/anti-charm quark. The scaling variable $\mathrm{z}_{\mathrm{hem}}$ is then defined as $\mathrm{z}_{\mathrm{hem}} = (E + p_{\mathrm{L}})_{D^*}/\sum_{\mathrm{hem}}(E + p)$.

The ZEUS collaboration has provided preliminary results [44] on a measurement of normalized differential cross sections of D^*-mesons as a function of $\mathrm{z}_{\mathrm{jet}}$. The measurement was done in photoproduction, in the kinematic range $Q^2 < 1$ GeV^2 and $130 < W < 280$ GeV. The D^*-mesons were reconstructed using the 'golden channel' $D^* \to D^0\pi_{\mathrm{s}} \to K\pi\pi_{\mathrm{s}}$ and were required to be in the central rapidity region $|\eta| < 1.5$ and to have $p_{\mathrm{T}} > 2$ GeV. Jets were reconstructed using the inclusive $k_\perp$ algorithm. They fulfill the conditions $|\eta_{\mathrm{jet}}| < 2.4$ and $E_{\mathrm{T,jet}} > 9$ GeV. The jets were reconstructed as massless jets. The beauty contribution to the D^*-meson cross section, which amounts to about 9%, was subtracted using the prediction of PYTHIA. The scaling variable was calculated as $\mathrm{z}_{\mathrm{jet}} = (E+p_{\mathrm{L}})_{D^*}/(2E_{\mathrm{jet}})$. The cross section as a function of $\mathrm{z}_{\mathrm{jet}}$ is shown in Fig. 6. The uncertainties due to choice of the model used to correct for detector effects, and the subtraction of the beauty component were the largest contributions to the total uncertainty.

The H1 collaboration has recently presented preliminary results [45] on the normalized differential cross section also of D^*-mesons as a function of both $\mathrm{z}_{\mathrm{hem}}$ and $\mathrm{z}_{\mathrm{jet}}$. Their measurement was performed

[3] Sometime there are slightly different definitions of z [28] in case of heavy meson production close to threshold.

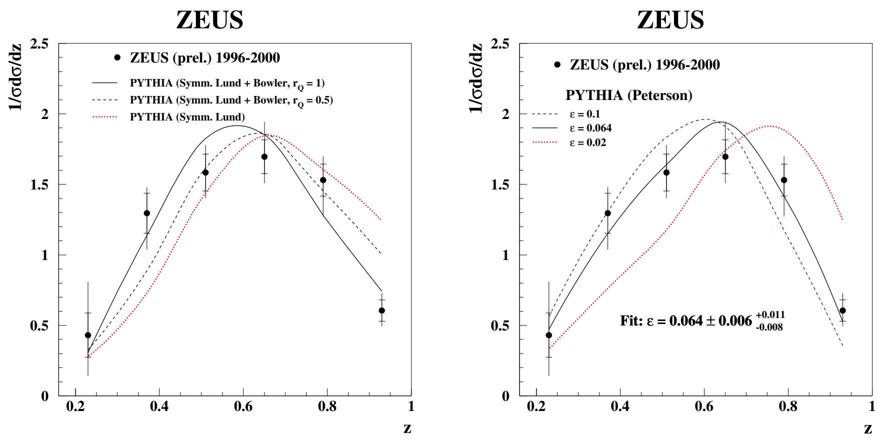


Fig. 6: Normalized differential cross section as a function of $z_{\rm jet}$ as measured by ZEUS in photoproduction for jets with an associated D^*-meson with $|\eta_{\rm jet}| < 2.4$ and $E_{\rm T,jet} > 9$ GeV.

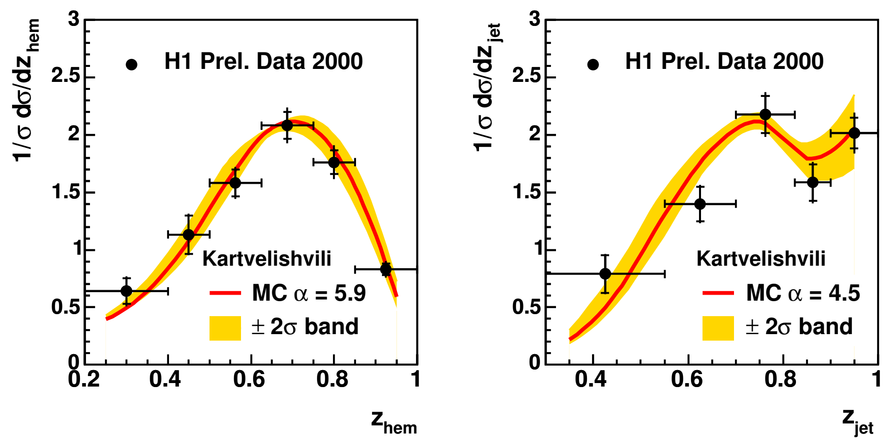


Fig. 7: Normalized differential cross section of D^*-meson as a function of $z_{\rm jet}$ and $z_{\rm hem}$ in DIS as measured by H1.

in the kinematic range $2 < Q^2 < 100$ GeV^2 and $0.05 < y < 0.7$. The D^*-mesons were reconstructed using the 'golden channel' with $|\eta| < 1.5$ in the central rapidity region and $p_T > 1.5$ GeV. The jets were reconstructed using the inclusive $k_\perp$ algorithm in the photon-proton center of mass frame, using the massive recombination scheme. The jets were required to have $E_{T,\mathrm{jet}} > 3$ GeV. The scaling variables were calculated as $z_{\mathrm{jet}} = (E + p_L)_{D^*}/(E + p)_{\mathrm{jet}}$ and $z_{\mathrm{hem}} = (E + p_L)_{D^*}/\sum_{\mathrm{hem}}(E + p)$ and are shown in Fig. 7. The resolved contribution was varied between 10 and 50% and the beauty contribution as predicted by the model was varied by a factor of two. The resulting uncertainties are part of the systematic error of the data points. For these distributions, the contribution of D^*-mesons coming from the fragmentation of beauty, as predicted by RAPGAP, was subtracted. It amounts to about 1.3% for the hemisphere method and 1.8% for the jet method. The dominant systematic errors are due to the model uncertainty and the signal extraction procedure.

Parametrization		Hemisphere Method	Jet Method	Suggested range
Peterson	ε	$0.018^{+0.004}_{-0.004}$	$0.030^{+0.006}_{-0.005}$	$0.014 < \varepsilon < 0.036$
Kartvelishvili	α	$5.9^{+0.9}_{-0.6}$	$4.5^{+0.5}_{-0.5}$	$4 < \alpha < 6.8$

Table 5: Extracted fragmentation parameters for z_{jet} and z_{hem} from H1.

Both collaborations used the normalized z-distributions to extract the best fragmentation parameters for a given QCD model.

In case of ZEUS, PYTHIA was used together with the Peterson fragmentation function. The MC was fit to the data using a χ^2-minimization procedure to determine the best value of ϵ. The result of the fit is $\epsilon = 0.064 \pm 0.006^{+0.011}_{-0.008}$.

The H1 collaboration used RAPGAP 3.1 interfaced with PYTHIA 6.2. The contribution due to D^*-mesons produced in resolved photon processes (in DIS), which amounts to 33% as predicted by the model, has been included in addition to the dominant direct photon contribution. The Peterson and Kartvelishvili parametrizations were both fitted to the data. The results are shown in Table 5.

The parameter of the Peterson fragmentation function as measured by ZEUS and H1 do not agree with each other. This may be due to the different phase-space regions covered by the two measurements (photoproduction versus DIS, $E_{T,\mathrm{jet}} > 9$ GeV versus $E_{T,\mathrm{jet}} > 3$ GeV) and most importantly, the parameters were extracted for two different models[4]. More detailed investigations are needed to resolve this question.

The fragmentation function parameters extracted by H1 with the hemisphere and the jet method differ by less than 3 σ. At the present level of statistical and systematic errors it is not possible to exclude a statistical fluctuation. On the other hand, the potential discrepancy may be a sign of deficiencies in the modelling of the hadronic final state in RAPGAP.

The measured z_{hem} distribution of H1 is compared to data from the ALEPH [23], OPAL [22] and CLEO [28] collaborations in Fig. 8 (left) and to ZEUS [44] and Belle [29] in Figure 8 (right)[5]. The results of H1 are in rough agreement with recent data from CLEO and Belle, taken at at 10.5 and 10.6 GeV, corresponding roughly to the average energy of the system at H1. Differences beyond the measurement errors can be observed. However, this may be due to the somewhat different definitions used for the fragmentation observable z, different kinematics, different processes, or it may be a sign of the violation of universality.

While the z distributions don't need to agree, the fragmentation parameters, which are extracted

[4]While ZEUS has used the default parameters for PYTHIA, H1 has taken the tuned parameter values of the ALEPH collaboration [46]

[5]Data points were taken from the figure in [29] and [44].

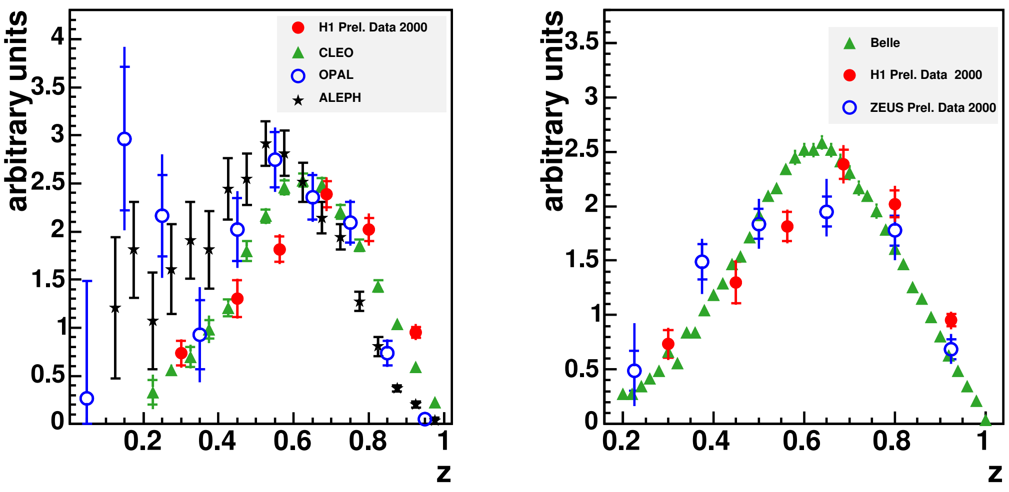


Fig. 8: Comparison of the z-distributions from CLEO, OPAL and ALEPH (left) and Belle and ZEUS (right) with the one from the hemisphere method from H1. All distributions are normalized to unit area from $z = 0.4$ to $z = 1$.

PARAMETRIZATION		ALEPH	OPAL	BELLE	ZEUS	H1: z_{hem}	H1: z_{jet}
Peterson	ε	0.034 ± 0.0037	0.034 ± 0.009	0.054	$0.064^{+0.013}_{-0.010}$	$0.018^{+0.004}_{-0.004}$	$0.030^{+0.006}_{-0.005}$
Kartvelishvili	α	—	4.2 ± 0.6	5.6	—	$5.9^{+0.9}_{-0.6}$	$4.5^{+0.5}_{-0.5}$

Table 6: Extracted fragmentation parameters from e^+e^- annihilation data by ALEPH [23], OPAL [22] and BELLE [29] and from ep data by ZEUS [44] and H1 [45].

from them, should agree. This can be expected only, if a model with consistent parameter settings is used which provides an equally good description of the different processes at their respective scales. The values of the Peterson fragmentation parameter, as extracted by different experiments within the PYTHIA/JETSET models, are summarized in Table 6. Contrary to expectations, discrepancies between various experiments can be seen. A consistent phenomenological analysis of these data is therefore needed in order to resolve the reasons for the discrepancies.

The measurement of the charm fragmentation function at HERA provides an important test of our understanding of heavy quark production. We may hope that HERA II data and a phenomenological analysis of existing data will bring new insights in this area.

References

[1] A. H. Mueller, Phys. Rev. **D18**, 3705 (1978).

[2] J. C. Collins and G. Sterman, Nucl. Phys. **B185**, 172 (1981).

[3] J. C. Collins, D. E. Soper, and G. Sterman, Adv. Ser. Direct. High Energy Phys. **5**, 1 (1988). `hep-ph/0409313`.

[4] J. C. Collins, D. E. Soper, and G. Sterman, Nucl. Phys. **B261**, 104 (1985).

[5] J. C. Collins, D. E. Soper, and G. Sterman, Nucl. Phys. **B308**, 833 (1988).

[6] G. T. Bodwin, Phys. Rev. **D31**, 2616 (1985). Erratum-ibid.D34:3932,1986.

[7] J. D. Bjorken, Phys. Rev. **D17**, 171 (1978).

[8] M. Suzuki, Phys. Lett. **B71**, 139 (1977).

[9] R. L. Jaffe and L. Randall, Nucl. Phys. **B412**, 79 (1994). hep-ph/9306201.

[10] P. Nason and B. R. Webber, Phys. Lett. **B395**, 355 (1997). hep-ph/9612353.

[11] M. Cacciari and E. Gardi, Nucl. Phys. **B664**, 299 (2003). hep-ph/0301047.

[12] V. G. Kartvelishvili, A. K. Likhoded, and V. A. Petrov, Phys. Lett. **B78**, 615 (1978).

[13] M. G. Bowler, Zeit. Phys. **C11**, 169 (1981).

[14] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, Phys. Rev. **D27**, 105 (1983).

[15] P. D. B. Collins and T. P. Spiller, J. Phys. **G11**, 1289 (1985).

[16] J. Chrin, Z. Phys. **C36**, 163 (1987).

[17] M. Cacciari and M. Greco, Phys. Rev. **D55**, 7134 (1997). hep-ph/9702389.

[18] B. Mele and P. Nason, Nucl. Phys. **B361**, 626 (1991).

[19] M. Cacciari and S. Catani, Nucl. Phys. **B617**, 253 (2001). hep-ph/0107138.

[20] M. L. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B373**, 295 (1992).

[21] S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B412**, 225 (1994). hep-ph/9306337.

[22] OPAL Collaboration, R. Akers *et al.*, Z. Phys. **C67**, 27 (1995).

[23] ALEPH Collaboration, R. Barate *et al.*, Eur. Phys. J. **C16**, 597 (2000). hep-ex/9909032.

[24] DELPHI Collaboration, P. Abreu *et al.*, Z. Phys. **C59**, 533 (1993).

[25] TASSO Collaboration, W. Braunschweig *et al.*, Z. Phys. **C44**, 365 (1989).

[26] H. Yamamoto *et al.*, Phys. Rev. Lett. **54**, 522 (1985).

[27] ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. **C52**, 353 (1991).

[28] CLEO Collaboration, M. Artuso *et al.*, Phys. Rev. **D70**, 112001 (2004). hep-ex/0402040.

[29] Belle Collaboration, R. Seuster *et al.* (2005). hep-ex/0506068.

[30] OPAL Collaboration, G. Abbiendi *et al.*, Eur. Phys. J. **C29**, 463 (2003). hep-ex/0210031.

[31] SLD Collaboration, K. Abe *et al.*, Phys. Rev. **D65**, 092006 (2002). hep-ex/0202031; SLD Collaboration, K. Abe *et al.*, Phys. Rev. **D66**, 079905 (2002).

[32] ALEPH Collaboration, A. Heister *et al.*, Phys. Lett. **B512**, 30 (2001). hep-ex/0106051.

[33] G. Barker *et al.*, *A study of the b-Quark Fragmentation Function with the DELPHI Detector at LEP I, Contributed Paper for ICHEP 2002, Abtract 583* (unpublished).

[34] JADE Collaboration, W. Bartel *et al.*, Z. Phys. **C33**, 339 (1987).

[35] DELCO Collaboration, T. Pal *et al.*, Phys. Rev. **D33**, 2708 (1986).

[36] TPC/Two Gamma Collaboration, H. Aihara *et al.*, Phys. Rev. **D31**, 2719 (1985).

[37] T. Sjostrand, L. Lonnblad, and S. Mrenna (2001). `hep-ph/0108264`.

[38] JADE Collaboration, W. Bartel *et al.*, Phys. Lett. **B146**, 121 (1984).

[39] P. Baringer *et al.*, Phys. Lett. **B206**, 551 (1988).

[40] TPC/Two Gamma Collaboration, H. Aihara *et al.*, Phys. Rev. **D34**, 1945 (1986).

[41] M. Cacciari, S. Frixione, and P. Nason, JHEP **03**, 006 (2001). `hep-ph/0102134`.

[42] M. Cacciari, M. Greco, and P. Nason, JHEP **05**, 007 (1998). `hep-ph/9803400`.

[43] M. Cacciari, P. Nason, and C. Oleari, *A study of heavy flavoured meson fragmentation functions in e+ e- annihilation*, 2005. `hep-ph/0510032`.

[44] Z. collaboration, *Measurement of charm fragmentation function in* D^* *photoproduction at HERA, contributed paper to XXXIst International Conference on High Energy Physics, July 2002, Amsterdam, The Nederlands.* (unpublished). 2002.

[45] H. Collaboration, *Determination of the Charm Fragmentation Function in Deep-Inelastic Scattering at HERA, contributed paper to XXIInd International Symposium on Lepton-Photon Interactions at High Energy, June 2005, Uppsala, Sweden.* (unpublished). 2005.

[46] G. Rudolph, private communication.

Benchmark cross sections for heavy-flavour production

O. Behnke[1]*, M. Cacciari*[2]*, M. Corradi*[3]*, A. Dainese*[4]*, H. Jung*[5]*, E. Laenen*[6]*, I. Schienbein*[7]*, H. Spiesberger*[8]

[1]Universität Heidelberg, Philosophenweg 12 69120 Heidelberg, FRG;
[2]LPTHE - Université P. et M. Curie (Paris 6), Paris, France;
[3]INFN Bologna, via Irnerio 46, Bologna, Italy;
[4]Università di Padova and INFN, Padova, Italy;
[5]Deutsches Elektronen-Synchroton DESY, Hamburg, FRG;
[6]NIKHEF,Theory Group, Kruislaan 409, 1098 SJ Amsterdam, The Netherlands;
[7]Southern Methodist University Dallas, 3215 Daniel Avenue, Dallas, TX 75275-0175, USA;
[8]Johannes-Gutenberg-Universität Mainz, D-55099 Mainz, FRG

Abstract

Reference heavy-flavour cross sections at HERA and LHC have been computed following different theoretical approaches and the results have been compared.

coordinators: *M. Corradi, A. Dainese*

1 Introduction

This section presents a comparison of cross sections for HERA and LHC calculated according to different theoretical approaches. Different programs were used to calculate the same reference cross sections, using, as far as possible, the same input parameters and a consistent method to evaluate uncertainties. In this way it is possible to identify processes and kinematical regions in which different approaches give the same answer and regions where they differ. Unified criteria to evaluate the theoretical uncertainty should also allow to understand what approach is expected to be more precise. Moreover these calculations, which incorporate up-to-date parameters and PDF parametrisations, can be used as a reference for experiments and for further theoretical predictions. The cross sections presented here, are available in computer-readable format from the web page `http://www-zeus.desy.de/~corradi/benchmarks`, where figures in color can also be found.

2 Programs

A list of the programs used for the cross section calculations is given below. For further details see the references and the theoretical review on heavy quark production in these proceedings.

- MNR [1] is a fixed-order (FO) NLO program for heavy-flavour hadro-production, it was used for LHC cross sections;
- FMNR [2, 3] is an extension of the previous program to photoproduction, it was used for photoproduction at HERA;
- HVQDIS [4, 5] is a FO-NLO program for heavy-flavour production in deep-inelastic scattering (DIS), it has been used for DIS at HERA;
- FONLL [6, 7] provides matched massive-massless calculations with NLO accuracy and resummation of large p_T logarithms. It is available for hadro- and photo-production and was used for HERA photoproduction and LHC cross sections;
- GM-VFNS [8, 9, 10, 11] is a calculation in the generalised massive variable flavour number scheme. It has been used for charmed hadron p_T spectra at LHC and in photoproduction at HERA;

Parameter	program	central value	lower/upper
$\Lambda^5_{\rm QCD}$	all	0.226 GeV	fix
m_c	all	1.5 GeV	1.3/1.7 GeV
m_b	all	4.75 GeV	4.5/5.0 GeV
p-PDF	all-CASCADE	CTEQ6.1 [15]	MRST2002 [16]/Alekhin [17]
	CASCADE	CCFM A0	-
γ-PDF	FMNR, FONLL	AGF [18]	GRV [19]
$f(c \to D^*)$	all	0.235	fix
c fragmentation:	(F)MNR,HVQDIS	Pet. [20] $\epsilon_c = 0.021$	$0.002/0.11$
	FONLL	BCFY $r = 0.1$	$0.06/0.135$
	GM-VFNS	[9]	-
	CASCADE, RAPGAP	Pet. $\epsilon_c = 0.075$	Def./$\epsilon_c = 0.05$
b fragmentation:	(F)MNR,HVQDIS	Pet. $\epsilon_b = 0.001$	$0.0002/0.004$
	FONLL	Kart. $\alpha = 29.1$	$25.6/34.0$
	CASCADE, RAPGAP	Pet. $\epsilon_b = 0.002$	Def./$\epsilon_b = 0.005$

Table 1: The table shows input parameter to the different programs with the corresponding lower and upper values used for the uncertainty: $\Lambda_{\rm QCD}$, the quark masses, the proton and photon parton densities, the fraction of c quarks decaying into a D^* meson, and the parameters used for fragmentation. The fragmentation form are abbreviated to Pet. for Peterson, Kart. for Kartvelishvili, Def. for the default PYTHIA fragmentation

- CASCADE 1.2009 [12] is a full Monte Carlo program based on unintegrated parton densities and K_T factorisation. It has been used to calculate cross sections for Photoproduction and DIS at HERA and for LHC;
- RAPGAP 3 [13] is a multi-purpose MC program for ep collisions, it implements heavy-flavour production through the boson-gluon-fusion process $\gamma^* g \to Q\bar{Q}$ at leading order. It has been used for DIS at HERA. Both CASCADE and RAPGAP use PYTHIA [14] routines for fragmentation.

3 Parameters and uncertainties

The different calculations were compared using the same input parameters and, where possible, with total uncertainty bands computed in a consistent way. The total uncertainty band includes the effect of the uncertainty on the input parameters and on the missing higher orders in the perturbative expansion.

3.1 Perturbative uncertainty

The perturbative uncertainty was obtained by varying the renormalisation and factorisation scales independently in the range $0.5\mu_0 < \mu_F, \mu_R < 2\mu_0$, while keeping $1/2 < \mu_R/\mu_F < 2$, were μ_0 is the nominal value, typically set to the transverse mass $p_T^2 + m_Q^2$ or to $4m^2 + Q^2$ in the DIS case. The largest positive and negative variations were taken as the perturbative uncertainty band.

3.2 Input parameters

The uncertainty from the input parameters was obtained by varying each parameter around the central value. An effort was made within the working group to find the best central value and uncertainty for the input parameters. The values used for the perturbative parameters $\Lambda^5_{\rm QCD}$, m_c, m_b as well as the parton distribution functions (PDF) for the proton and for the photon are reported in Table 1.

For practical reasons, rather than using the full treatment of the PDF uncertainty, few different parametrisations were tried and it was checked that the choice of the PDF set always gives a small contribution to the total uncertainty band. In the case of CASCADE, the CCFM A0 parametrisation was used as the central value while the PDF parametrisations A0+ and A0-, obtained from fits to DIS data

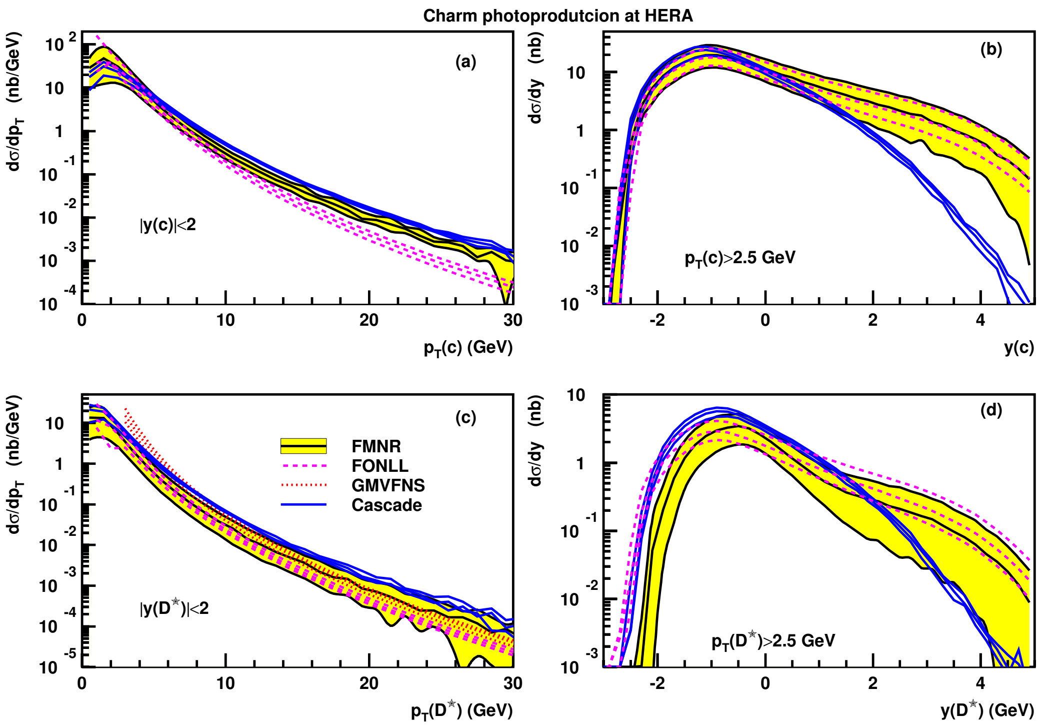


Fig. 1: Cross sections for charm photoproduction at HERA ($Q^2 < 1$ GeV^2, $0.2 < y < 0.8$). The differential cross sections as a function of the p_T of the c quark for rapidity $|Y| < 2$ and as a function of the rapidity of the c quark for $p_T > 2.5$ GeV are shown in (a), (b). Plots (c) and (d) show similar cross sections for the production of a D^* meson. The cross sections are shown for FMNR (shaded band), FONLL (empty band with dashed lines), GM-VFNS (empty band with dotted lines) and CASCADE (empty band with full lines).

with different renormalisation scales, were used in conjunction with the variation of the renormalisation scale.

Since the different programs have different perturbative contents, different parameters for the non-perturbative fragmentation function were used. The values were chosen in order to correspond to the same average fragmentation in e^+e^- collisions as explained in the section on heavy quark fragmentation in these proceedings. Table 1 reports the fragmentation form and the corresponding parameters used in the different programs.

In the FONLL calculation for charm, the BCFY [21] fragmentation parameter r was varied in conjunction with the variation of the charm mass since different values of r are obtained from e^+e^- data for different m_c [22]. Similarly for beauty, the Kartvelishvili [23] parameter α was varied in conjunction with the variation of the b mass [23]. For GM-VFNS, the fragmentation functions and fractions were taken from [9].

The total uncertainty band was obtained from the sum of the uncertainties added in quadrature coming from the parameter variations and the perturbative uncertainty.

4 Results

4.1 HERA Photoproduction

The results for HERA Photoproduction are given as ep cross-sections for $0.2 < y < 0.8$ (y is the Bjorken variable while Y is the rapidity in the laboratory frame) and $Q^2 < 1$ GeV^2. The beam energies have been set to $E_e = 27.52$ GeV, $E_p = 920$ GeV with the proton beam going in the positive rapidity direction.

Figure 1 shows the differential cross sections as a function of the charm quark transverse momen-

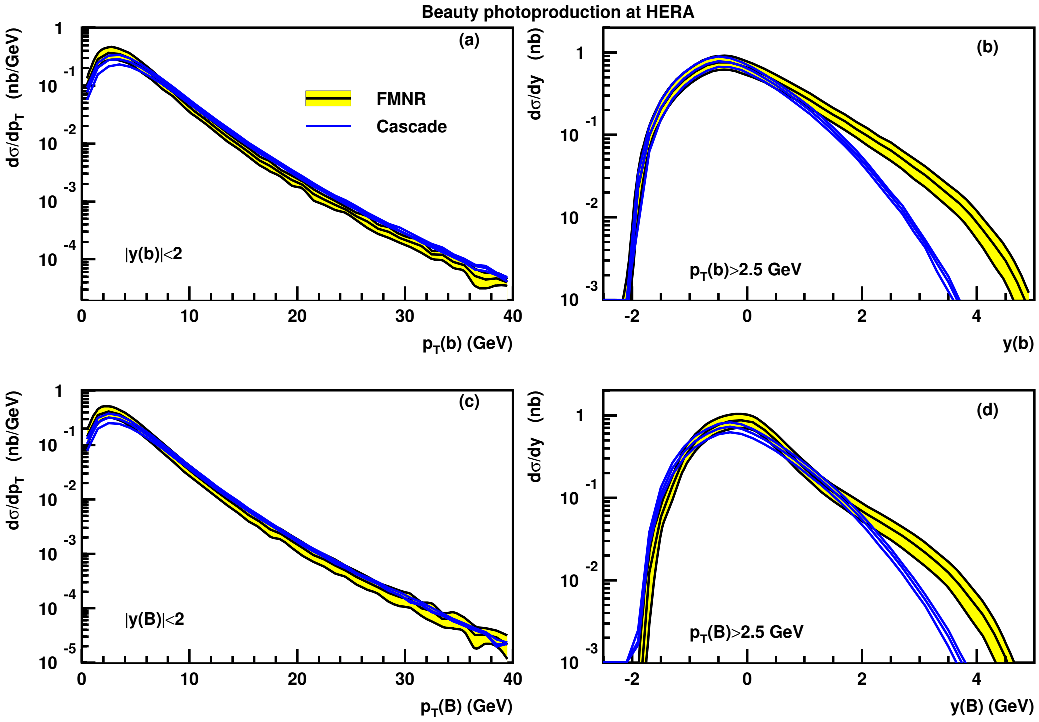


Fig. 2: Cross sections for beauty photoproduction at HERA ($Q^2 < 1$ GeV^2, $0.2 < y < 0.8$). The differential cross sections in p_T and rapidity of the b quark are shown in (a), (b). Plots (c) and (d) show the cross sections for the production of a weakly-decaying B hadron as a function of $p_T(B)$ and $Y(B)$. The cross sections are shown for FMNR (shaded band) and CASCADE (empty band with full lines).

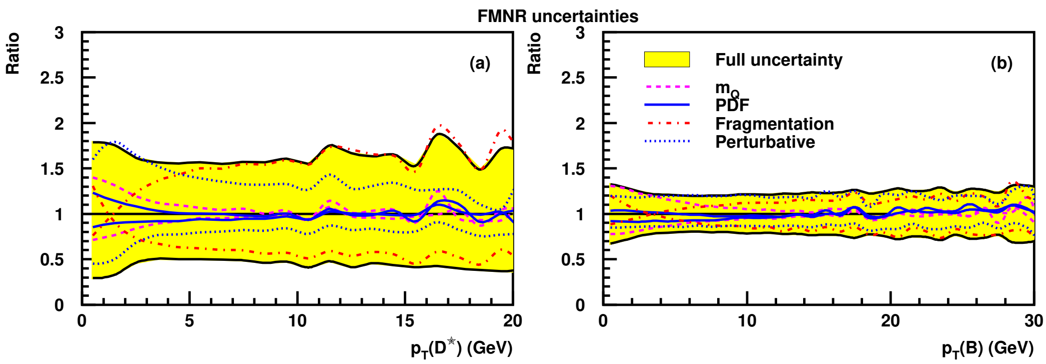


Fig. 3: Breakdown of the different components of the FMNR uncertainty for $d\sigma/dp_T$ for charmed (a) and beauty (b) hadrons in photoproduction at HERA. The plots show the ratio of the upper/lower side of each uncertainty to the nominal value. The following sources of uncertainty are shown: quark mass (m_Q), parton density parametrisation (PDF), fragmentation parameter and the perturbative uncertainty from scale variations.

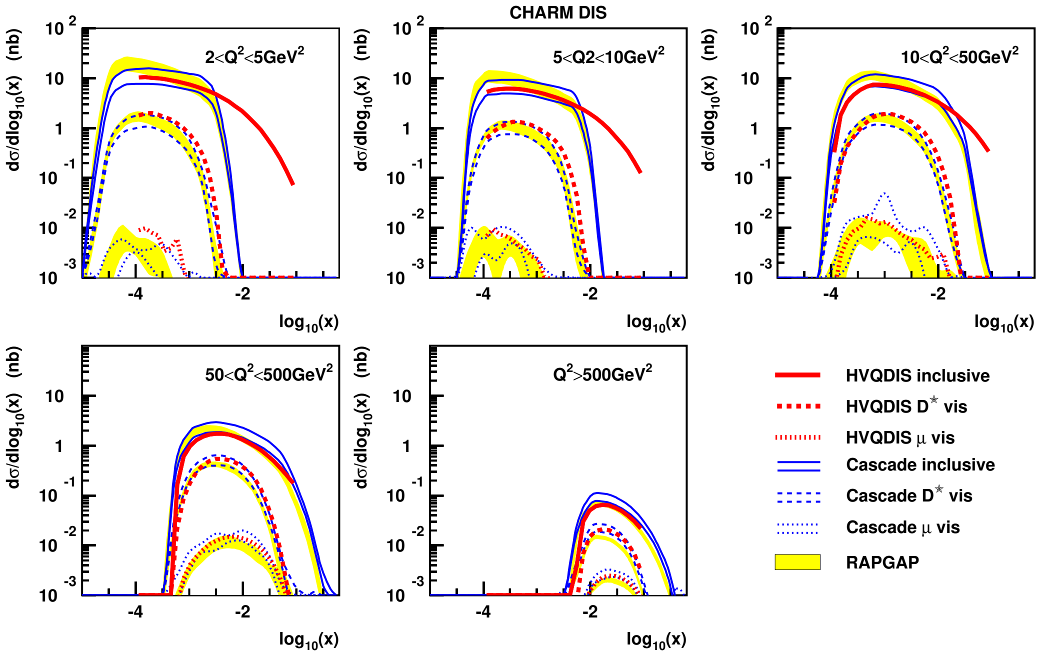


Fig. 4: Charm cross sections in DIS at HERA. Each plot shows the distribution of $\log_{10}(x)$ in a different Q^2 range for the inclusive cross-section, the cross-section for a D^* meson in the "visible" range $p_T(D^*) > 1.5$ GeV, $|Y(D^*)| < 1.5$ and the cross-section for a muon from charm decay in the range $p_T(\mu) > 3$ GeV, $|Y(\mu)| < 2$. The thick curves show the central value from HVQDIS, the thin curves represent the uncertainty band from CASCADE and the shaded area shows the uncertainty band from RAPGAP. The fluctuations in the muon cross sections are due to the limited statistics.

tum (a) and pseudorapidity (b). In (c) and (d) the same cross sections are given for the charmed D^* meson. A meaningful comparison can be performed only for the hadron variables, which are the real physical observables, since the quark level may be defined differently in different approaches. The FO calculation (FMNR) shows a large uncertainty ($\sim 60\%$) at the hadron level due to the related uncertainty on the fragmentation parameters. The resummed programs FONLL and GM-VFNS have much smaller uncertainty and are within the FMNR uncertainty band. The central values from FMNR and FONLL coincide at low transverse momenta. GM-VFNS, instead, tends to grow unphysically at low $p_T(D^*)$. As can be seen in (c), the quark-level disagreement between FO (FMNR) and FONLL calculations is consistently removed at the hadron-level. The unintegrated-PDF Monte Carlo CASCADE tends to be above the other calculations, in particular at large p_T. In the case of beauty (Fig. 2) the uncertainty bands are smaller ($\sim 20\%$ for FMNR), CASCADE and FMNR are in good agreement. Due to the large b mass, the resummed calculation FONLL (not shown) is expected to be similar to the fixed-order one (FMNR). For both beauty and charm, FMNR and FONLL show a shoulder at positive rapidities (b, d) due to the "hadron-like" component of the photon that is not present in CASCADE.

Figure 3 shows the different components of the FMNR uncertainty band for charm and beauty. The uncertainties for quark production are typically dominated by the perturbative scale uncertainty with the exception of the low transverse momentum region ($p_T \sim m_Q$) where the uncertainty from the quark-mass can dominate. For hadron production, the fragmentation dominates the FMNR uncertainty at large p_T. The PDF uncertainty was found to be small. Resummed calculation have smaller uncertainty bands due to the smaller perturbative and fragmentation contributions at large p_T.

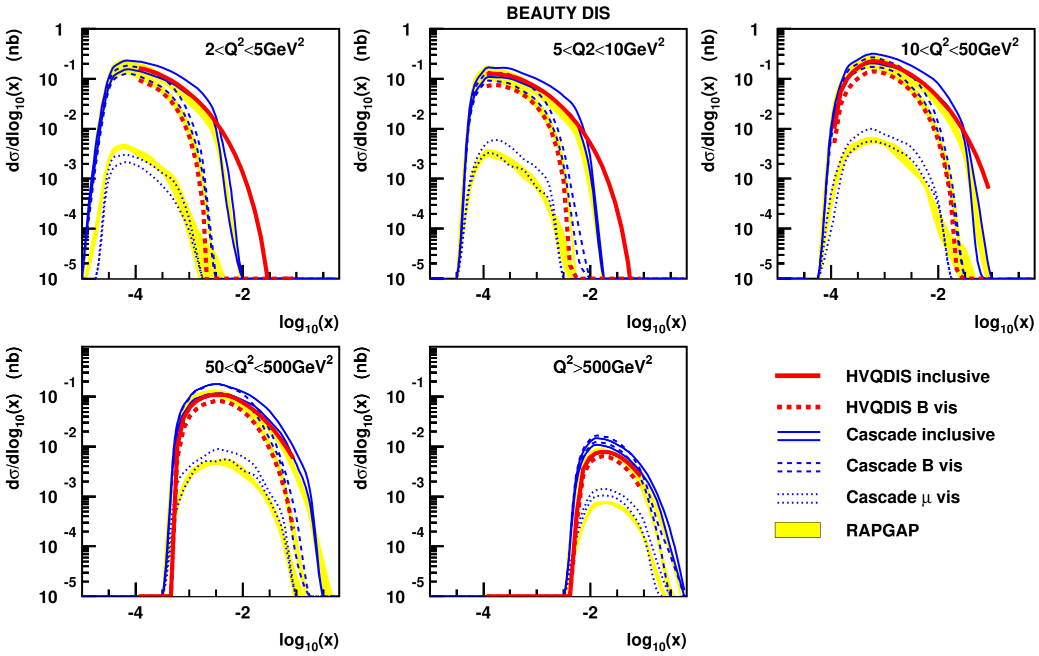


Fig. 5: Beauty cross sections in DIS at HERA. Each plot shows the distribution of $\log_{10}(x)$ in a different Q^2 range for the inclusive cross-section, the cross section for a hadron containing a b quark in the "visible" range $p_T(B) > 3$ GeV, $|Y(B)| < 2$ and the cross section for a muon from beauty decay in the range $p_T(\mu) > 3$ GeV, $|Y(\mu)| < 2$. The thick curves show the central value from HVQDIS, the thin curves represent the uncertainty band from CASCADE and the shaded area shows the uncertainty band from RAPGAP. The muon distributions are not given for HVQDIS.

4.2 HERA DIS

Heavy quark production in DIS is not available in the matched massive-massless approach (except for total cross sections). Therefore the DIS comparison was limited to the FO-NLO program HVQDIS, the unintegrated-PDF MC CASCADE and the RAPGAP Monte Carlo. The DIS cross sections at HERA are reported as $d\sigma/d\log_{10}(x)$ for different bins of Q^2 and are intended at the Born level, without electroweak corrections. Figure 4 shows, for each Q^2 bin, the inclusive charm cross-section, the cross section for observing a D^* meson in the "visible" range $p_T(D^*) > 1.5$ GeV, $|Y(D^*)| < 1.5$ and for observing a muon in the range $p_T(\mu) > 3$ GeV, $|Y(\mu)| < 2$. The three calculations are compatible at intermediate values of x ($\sim 10^{-3}$). At large x and low Q^2, CASCADE and RAPGAP drop to zero much faster than HVQDIS. At low x RAPGAP is significantly larger than HVQDIS while both are within the uncertainty band given by CASCADE. A similar behavior is seen for beauty (Fig. 5). The uncertainty on HVQDIS, not given here, is expected to be small ($\sim 10-20\%$ for beauty [24]). The high-x discrepancy between HVQDIS and the other two calculations seems therfore to be beyond the program uncertainties and deserves further investigations.

4.3 LHC

For LHC, we computed the cross sections in pp collisions at $\sqrt{s} = 14$ TeV.

Figures 6 and 7 show the single inclusive cross sections as a function of p_T, at quark (upper panels) and hadron (lower panels) level, for charm and beauty, respectively. Two rapidity intervals are considered: $|Y| < 2.5$, approximately covering the acceptance of the barrel detectors of ATLAS ($|\eta| < 2.5$), CMS ($|\eta| < 2.5$), and ALICE ($|\eta| < 0.9$); $2.5 < |Y| < 4$, approximately covering the acceptance of LHCb ($2 < \eta < 5$) and of the ALICE muon spectrometer ($2.5 < \eta < 4$).

For charm, we compare the fixed-order NLO results from MNR to the results from the CASCADE event generator, from the GM-VFNS calculation and from the FONLL calculation. The agreement is in general good, in particular in the low-p_T region; at high-p_T CASCADE predicts a larger cross section than the other calculations, especially at forward rapidities. The FONLL central prediction is in agreement with that of the FO NLO calculation at low p_T, while deviating from it at high p_T, where it gives a smaller cross section.

For beauty, we compare FO NLO (MNR), FONLL and CASCADE. Again, there is agreement at low p_T, where, as expected, the FONLL result coincides with the MNR result. At high p_T, both CASCADE and FONLL predict a larger cross section than the MNR central values, but all models remain compatible within the theoretical uncertainties. At forward rapidities, for beauty as for charm, CASCADE gives a significantly larger cross section than MNR.

Figure 8 shows the breakdown of the uncertainties for hadron production as obtained with MNR. The perturbative component dominates at LHC. Only the fragmentation component for charm hadron production becomes comparable in size to the perturbative one at large p_T.

4.4 Q-$\bar{Q}$ correlations

The azimuthal separation between the two heavy quarks $\Delta\phi(Q\bar{Q})$ and the transverse momentum of the quark-antiquark system $p_T(Q\bar{Q})$ are particularly sensitive to higher-order effects since at leading order their distributions are delta functions peaked at $\Delta\phi(Q\bar{Q}) = \pi$ and $p_T(Q\bar{Q}) = 0$. The distribution of these variables is therefore a direct probe of QCD radiation and is well suited for comparing different calculations.

Figures 9 and 10 show the heavy-quark pair p_T distribution and the quark-antiquark relative azimuthal angle distribution for charm and beauty at LHC, respectively. For both distributions, the two quarks of the pair are required to have $|Y| < 2.5$; also minimum p_T selections are applied to mimic the effect of realistic experimental cuts ($p_T^Q > 3$ GeV and $p_T^{\bar{Q}} > 6$ GeV). In the region near $\Delta\phi(Q\bar{Q}) = \pi$

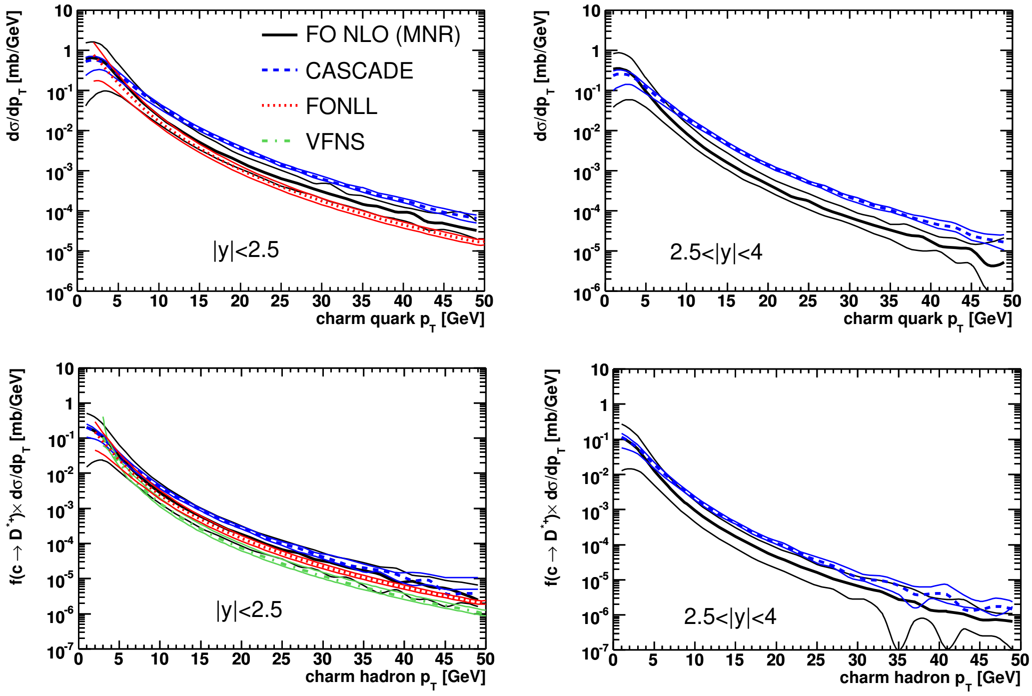


Fig. 6: Cross sections for charm production in pp collisions at the LHC with $\sqrt{s} = 14$ TeV. The differential cross sections in p_T for c quark in the two rapidity ranges $|Y| < 2.5$ and $2.5 < |Y| < 4$ are shown in the upper panels. The lower panels show the cross sections for the production of a D^* meson as a function of $p_T(D^*)$ in the same rapidity ranges.

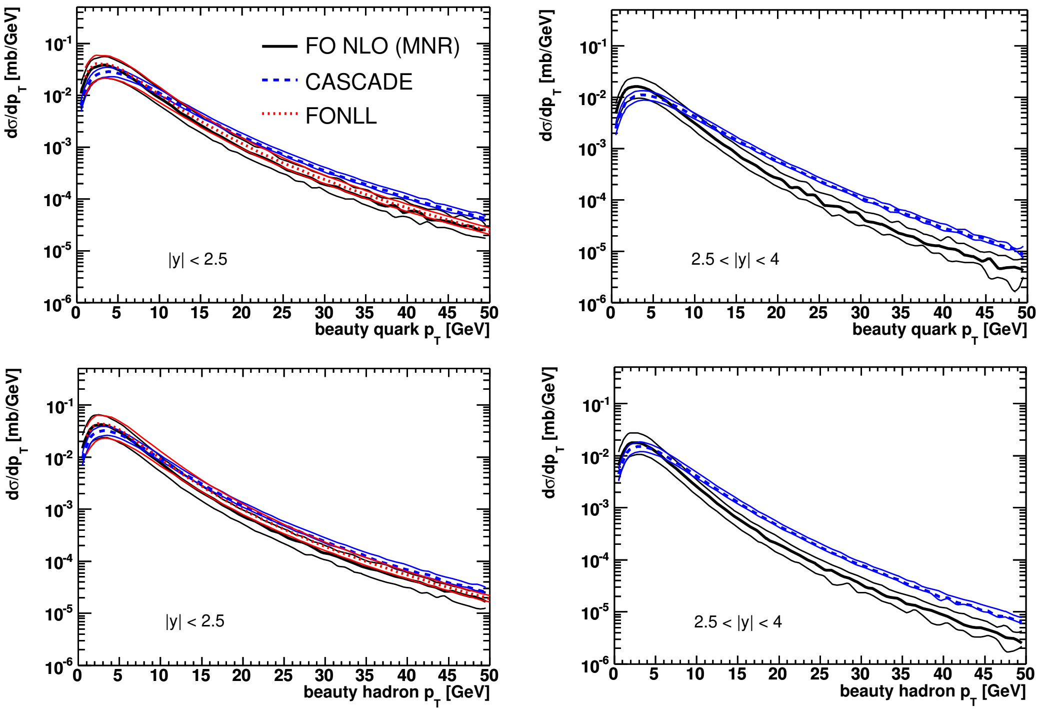


Fig. 7: Cross sections for beauty production in pp collisions at the LHC with $\sqrt{s} = 14$ TeV. The differential cross sections in p_T for b quark in the two rapidity ranges $|Y| < 2.5$ and $2.5 < |Y| < 4$ are shown in the upper panels. The lower panels show the cross sections for the production of a beauty hadron as a function of p_T in the same rapidity ranges.

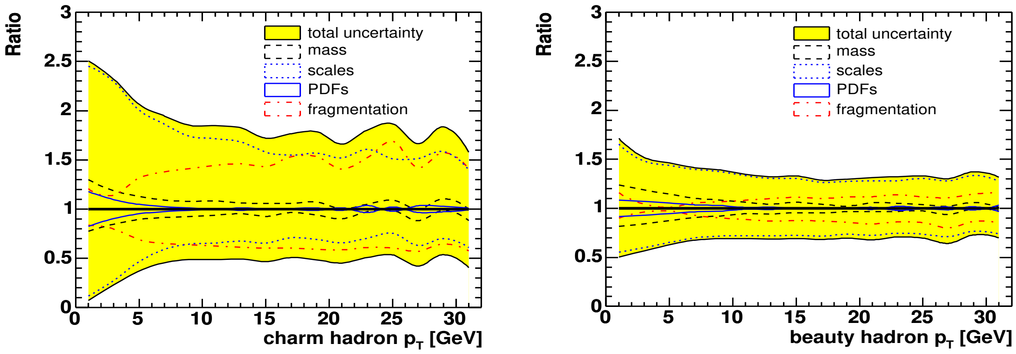


Fig. 8: Breakdown of the different components of the uncertainty on $d\sigma/dp_T$ for charmed (a) and beauty (b) hadrons at LHC as obtained from MNR. The plots show the ratio of the upper/lower side of each uncertainty to the nominal value. The following sources of uncertainty are shown: quark mass (m_Q), parton density parametrisation (PDF), fragmentation parameter and the perturbative uncertainty from scale variations.

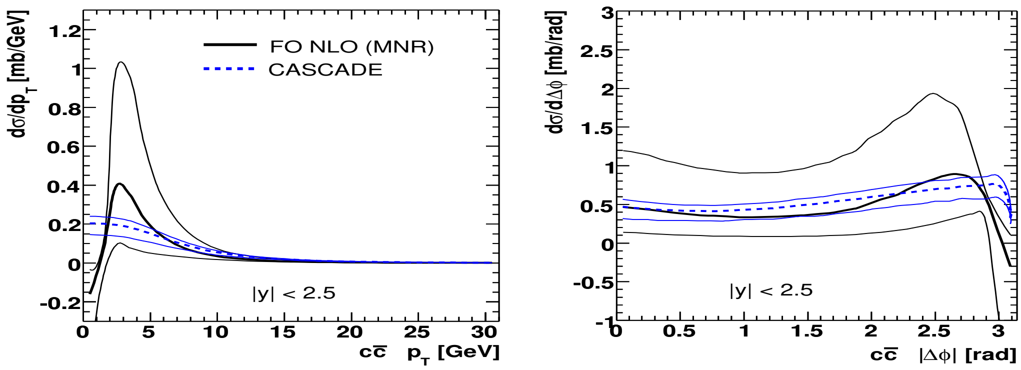


Fig. 9: Q-$\bar{Q}$ correlations for charm at LHC: p_T of the $c\bar{c}$ pair (left) and azimuthal angle $\Delta\phi$ between the c and the $\bar{c}$ (right). For both cross sections, the following kinematic cuts are applied: $|Y^c| < 2.5$, $|Y^{\bar{c}}| < 2.5$, $p_T^c > 3$ GeV, $p_T^{\bar{c}} > 6$ GeV.

and $p_T(Q\bar{Q}) = 0$, where the cancellation of soft and collinear divergencies occur, the fixed-order NLO calculation gives an unphysical negative cross section with next to a large positive peak. A larger binning would be needed to average this behavior and produce a more physical results. The CASCADE MC, has a more realistic behavior. Both calculations have a non-zero value at $\Delta\phi(Q\bar{Q}) = 0$ related to "gluon-splitting" events. A similar result was found for HERA as shown in Figure 11. This kind of distribution is expected to be well described by programs that merge NLO matrix elements to the parton-shower MC approach such as MC@NLO [25].

5 Conclusions

Heavy-flavour cross sections for HERA and LHC, obtained with fixed-order NLO programs, with matched massive/massless calculations and within the K_T-factorisation approach have been compared. Similar results are found for photoproduction at HERA and for the LHC. As expected the resummed calculations were found to be compatible with the fixed-order results but have smaller uncertainties at large p_T. Resummed calculations for charm in two different schemes (GM-VFNS and FONLL) are anyway somewhat incompatible both at HERA and LHC, suggesting that their uncertainty may be underestimated. The K_T-factorisation program CASCADE predicts larger cross sections than the other approaches at large p_T at LHC and for charm at HERA. The comparison for DIS was limited to FO-NLO and a MC program with leading order matrix elements. Large discrepancies, which deserve further investigations,

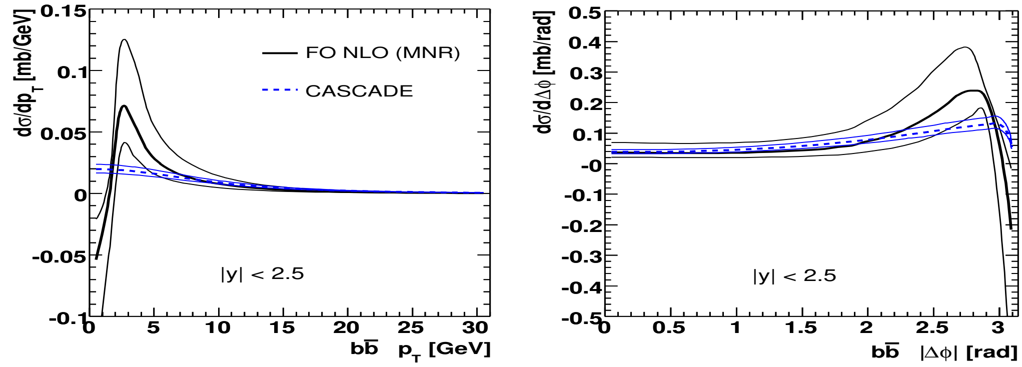


Fig. 10: Q-$\bar{Q}$ correlations for charm at LHC: p_T of the $b\bar{b}$ pair (left) and azimuthal angle $\Delta\phi$ between the b and the $\bar{b}$ (right). For both cross sections, the following kinematic cuts are applied: $|Y^b| < 2.5$, $|Y^{\bar{b}}| < 2.5$, $p_T^b > 3$ GeV, $p_T^{\bar{b}} > 6$ GeV.

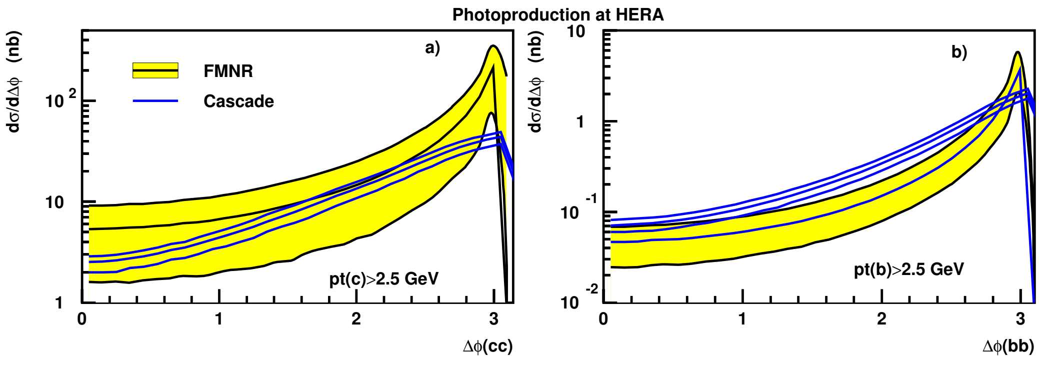


Fig. 11: Azimuthal Q-$\bar{Q}$ correlations in photoproduction at HERA for charm (a) and beauty (b). One of the two quarks was required to be in the "visible" region $p_T(Q) > 2.5$ GeV and $|Y(Q)| < 2$.

were found in this case. A comparison with experimental data would be needed for further understanding of the quality of the available calculations.

References

[1] Mangano, Michelangelo L. and Nason, Paolo and Ridolfi, Giovanni, Nucl. Phys. **B373**, 295 (1992).

[2] Frixione, Stefano and Mangano, Michelangelo L. and Nason, Paolo and Ridolfi, Giovanni, Nucl. Phys. **B412**, 225 (1994).

[3] Frixione, Stefano and Mangano, Michelangelo L. and Nason, Paolo and Ridolfi, Giovanni, Phys. Lett. **B348**, 633 (1995).

[4] Harris, B. W. and Smith, J., Phys. Rev. **D57**, 2806 (1998).

[5] Smith, J. and Harris, B. W., Nucl. Phys. Proc. Suppl. **51C**, 188 (1996).

[6] Cacciari, Matteo and Frixione, Stefano and Nason, Paolo, JHEP **03**, 006 (2001).

[7] Cacciari, Matteo and Greco, Mario and Nason, Paolo, JHEP **05**, 007 (1998).

[8] Kniehl, B. A. and Kramer, G. and Schienbein, I. and Spiesberger, H., Phys. Rev. **D71**, 014018 (2005).

[9] Kniehl, B. and Kramer, Gustav, Phys. Rev. **D71**, 094013 (2005).

[10] Kniehl, B. A. and Kramer, G. and Schienbein, I. and Spiesberger, H., Eur. Phys. J. **C41**, 199 (2005).

[11] Kramer, G. and Spiesberger, H., Eur. Phys. J. **C38**, 309 (2004).

[12] Jung, H., Comput. Phys. Commun. **143**, 100 (2002).

[13] Jung, Hannes, Comp. Phys. Commun. **86**, 147 (1995).

[14] Sjostrand, Torbjorn and Lonnblad, Leif and Mrenna, Stephen (2001).

[15] Stump, Daniel and others, JHEP **10**, 046 (2003).

[16] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C28**, 455 (2003).

[17] Alekhin, Sergey, Phys. Rev. **D68**, 014002 (2003).

[18] Aurenche, P. and Guillet, J. P. and Fontannaz, M., Z. Phys. **C64**, 621 (1994).

[19] Gluck, M. and Reya, E. and Vogt, A., Phys. Rev. **D46**, 1973 (1992).

[20] Peterson, C. and Schlatter, D. and Schmitt, I. and Zerwas, Peter M., Phys. Rev. **D27**, 105 (1983).

[21] Braaten, Eric and Cheung, King-man and Fleming, Sean and Yuan, Tzu Chiang, Phys. Rev. **D51**, 4819 (1995).

[22] Cacciari, Matteo and Nason, Paolo and Oleari, Carlo (2005).

[23] Kartvelishvili, V. G. and Likhoded, A. K. and Petrov, V. A., Phys. Lett. **B78**, 615 (1978).

[24] Carli, T. and Chiochia, V. and Klimek, K., JHEP **09**, 070 (2003).

[25] Frixione, Stefano and Nason, Paolo and Webber, Bryan R., JHEP **08**, 007 (2003).

4 Working group 4: Diffraction

List of participants in the working group:

L. Adamczyk, N. Amapane, V. Andreev, M. Arneodo, V. Avati, C. Avila, J. Bartels, R. Bellan, S. Bolognesi, M. Boonekamp, A. Bruni, A. Bunyatyan, G. Cerminara, B. Cox, R. Croft, A. De Roeck, M. Diehl, I. Efthymiopoulos, K. Eggert, F. Ferro, J. Forshaw, E. Gotsman, D. Goulianos, P. Grafström, M. Grothe, M. Groys, C. Hogg, G. Ingelman, X. Janssen, H. Jung, J. Kalliopuska, M. Kapishin, D. Kharzeev, V. Khoze, M. Klasen, H. Kowalski, G. Kramer, J. W. Lämsä, P. V. Landshoff, P. Laycock, E. Levin, A. Levy, L. Lönnblad, M. Lublinsky, L. Lytkin, D. Macina, T. Mäki, U. Maor, C. Mariotti, N. Marola, A. D. Martin, C. Marquet, V. Monaco, J. Monk, L. Motyka, S. Munier, M. Murray, E. Naftali, P. Newman, J. Nystrand, F. Oljemark, R. Orava, K. Österberg, M. Ottela, A. Panagiotou, R. Peschanski, A. Pilkington, K. Piotrzkowski, A. Proskuryakov, A. Prygarin, J. Raufeisen, M. Rijssenbeek, M. Ruspa, M. G. Ryskin, M. Ryynänen, A. Sabio-Vera, R. Sacchi, S. Schätzel, F.-P. Schilling, A. Sobol, M. Strikman, S. Tapprogge, M. Taševský, P. Van Mechelen, R. Venugopalan, K. Vervink, G. Watt, K. Wichmann, G. Wolf

Introduction to Diffraction

M. Arneodo [a], J. Bartels [b], A. Bruni [c], B. E. Cox [d], M. Diehl [e], J. Forshaw [d], M. Grothe [f], H. Kowalski [e], R. Orava [g], K. Piotrzkowski [h] and P. Van Mechelen [i]
[a]Università del Piemonte Orientale, Novara, and INFN-Torino, Italy
[b]II. Institut für Theoretische Physik, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany
[c]INFN Bologna, Via Irnerio 46, 40156 Bologna, Italy
[d]School of Physics and Astronomy, The University of Manchester, Manchester M139PL, United Kingdom
[e]Deutsches Elektronen-Synchroton DESY, 22603 Hamburg, Germany
[f]University of Torino and INFN-Torino, Italy; also at University of Wisconsin, Madison, WI, USA
[g]High Energy Physics Division, Department of Physical Sciences, University of Helsinki and Helsinki Institute of Physics, P. O. Box 64, 00014 University of Helsinki, Finland
[h]Institut de Physique Nucléaire, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
[i]Universiteit Antwerpen, Antwerpen, Belgium

Abstract
We summarize the main activities of the Working Group on Diffraction in this workshop, which cover a wide range of experimental, phenomenological and theoretical studies. Central themes are exclusive and inclusive diffraction at HERA and the LHC, multiple interactions and rapidity gap survival, and parton saturation.

1 Forward proton tagging at the LHC as a means to discover new physics

The use of forward proton tagging detectors at CMS and ATLAS as a means to search for and identify the nature of new physics at the LHC was one of the major topics of discussion at the workshop. The process of interest is the so-called “central exclusive” production process $pp \to p \oplus \phi \oplus p$, where $\oplus$ denotes the absence of hadronic activity (a “gap”) between the outgoing intact protons and the decay products of the central system ϕ. The final state therefore consists of *only* the decay products of the system ϕ, which can be seen in the central detectors, and the two outgoing protons, which must be detected at some point downstream of the interaction point where they emerge far enough from the LHC beams. To this end, the feasibility of installing proton tagging detectors at 420m from the interaction points of ATLAS and / or CMS, at a suitable time after the initial start-up of the LHC, is currently being assessed [1]. These would complement and increase the acceptance of the detectors already planned in the 220m / 240m region by CMS / TOTEM and ATLAS. The choice of the 420m region is set by the central system masses of interest; protons which lose approximately 60 GeV of their longitudinal momentum – the interesting range from the point of view of Higgs boson searches – emerge from the beam in this region.

The motivation for these studies stems from the unique properties of central exclusive production. Firstly, the mass of the central system ϕ can be measured to high accuracy by measuring the four-momenta of the outgoing protons alone, without reference to the central system (the so-called “missing mass method” [2]). The achievable mass resolution and the acceptance as a function of mass of the 420m detectors (in combination with the already planned 220m proton detectors) are discussed in detail in these proceedings [3, 4]. The resolution can be as good as 1 GeV for a Higgs boson of mass 140 GeV. As an example, in the case of a 140 GeV Standard Model Higgs decaying to two W bosons, and the subsequent leptonic decays of one or both of the Ws to leptons plus neutrinos, 6 events are expected with no modification of the level 1 trigger thresholds of ATLAS and CMS for 30 fb^{-1} of delivered luminosity. We discuss the trigger issues in more detail below. This number is expected to double if realistic changes are made to the leptonic trigger thresholds [5].

A second crucial advantage is that, to a good approximation, the central system ϕ is produced in the $J_z = 0$, C and P even state, and an absolute determination of the quantum numbers of any resonance is possible by measuring correlations between the outgoing proton momenta. Observation of any resonance production with associated proton tags, therefore, allows a determination of its quantum numbers, something that is difficult to do in any other process at the LHC. Such a determination could be made with only a few "gold-plated" events.

Thirdly, states which would otherwise be very difficult to detect in conventional channels can be detected in the central exclusive channel. Perhaps the best-studied example is the high $\tan\beta$ region of the MSSM, where over 100 signal events can be detected with backgrounds lower by an order of magnitude or more, within 30 fb^{-1} of delivered luminosity at the LHC [6]. There are extensions to the MSSM in which central exclusive production becomes in all likelihood the only method at the LHC of isolating the underlying physics. One example [7] is the case where there are non-vanishing CP phases in the gaugino masses and squark couplings. In such scenarios, the neutral Higgs bosons are naturally nearly degenerate for large values of $\tan\beta$ and charged Higgs masses around 150 GeV. In such scenarios, observing the mass spectrum using forward proton tagging may well be the only way to explore such a Higgs sector at the LHC. Explicit CP-violation in the Higgs sector can be observed as an asymmetry in the azimuthal distributions of the tagged protons [8].

From an experimental perspective, the key issue along with the mass resolution and acceptance is the level 1 trigger efficiency. The problem is that detectors at 420m from the interaction points of ATLAS or CMS are too far away to participate in a level 1 trigger decision without an increase in the trigger latency. This means that the central detectors, or forward detectors up to 220m, must be relied upon to keep candidate events until the signals from 420m can be used in higher level trigger decisions. A full description of the work done at the workshop is presented in [9, 10] in these proceedings. The most difficult case is that of a low-mass (120 GeV) Higgs boson decaying in the b-quark channel (a decay mode that will not be observed in any other measurement at the LHC). The relatively low transverse momenta of the b-jets necessitate L1 jet E_T thresholds as low as 40 GeV. Thresholds that low would result in a L1 trigger rate of more than 50 kHz, due to the QCD background, and thus would essentially saturate the available output bandwidth. The output rate of a 2-jet L1 trigger condition with thresholds of 40 GeV per jet can be kept at an acceptable level of order 1 kHz in the absence of pile-up (i.e. for a single proton-proton interaction per bunch crossing) by either using the TOTEM T1 and T2 detectors (or the ATLAS forward detectors) as vetoes – central exclusive events have no energy in these regions – or by requiring that a proton be seen in the TOTEM (or ATLAS) detectors at 220 m on one side of the interaction point. This gives a sufficient reduction of the QCD background event rate. At higher luminosities, up to $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, where pile-up is present, it is necessary to combine a 220m tag with additional conditions based on event topology and on H_T, the scalar sum of all L1 jet E_T values. These L1 trigger conditions result in signal efficiencies between 15% and 20%. A further 10% of the Higgs events can be retained by exploiting the muon-rich final state in the $H \to b\bar{b}$ mode, with no requirements on the forward detectors. Other interesting decay channels, such as WW and $\tau\tau$, should be possible at the highest luminosities ($1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$) since both ATLAS and CMS will trigger on such events routinely using only the central detectors.

As well as upgrading the proton tagging capabilities of ATLAS and CMS, there was also discussion of upgrading the very forward region of CMS to extend the pseudo-rapidity coverage up to $|\eta| \sim 11$. This would allow proton x values down to 10^{-8} to be probed, opening up an unexplored region of small-x parton dynamics [12].

In summary, central exclusive production provides an excellent means of measuring the masses of new particles with a precision at the 1 GeV level, irrespective of the decay mode of the particles. It also provides a clean way of unambiguously determining the quantum numbers of any resonances produced in the central exclusive process (including Standard Model and MSSM Higgs bosons) at the LHC. In certain regions of the MSSM, and indeed for any scenarios in which the new particles couple strongly to gluons, central exclusive production may be the discovery channel.[1] The challenge is to design and build proton tagging detectors with the capability to measure the momentum loss of the outgoing protons at the 1 GeV level.

2 Theory of diffractive Higgs production

It is a fact that the theoretical predictions for central exclusive production suffer from several sources of uncertainty. The theoretical framework is presented and critically assessed in the contribution by Forshaw [13]. The emphasis is on the calculations of the Durham group, which are performed within perturbative QCD. The use of perturbative QCD is shown to be justified, with around 90% of the contribution to the Standard Model Higgs production cross-section ($m_H = 120$ GeV) coming from the region where the gluon virtualities are all above 1 GeV.

One of the main sources of uncertainty in the perturbative calculation arises from a lack of knowledge of the proton's generalized, unintegrated gluon distribution function, and so far estimates are based upon theoretically motivated corrections to the more familiar gluon distribution function. It is hard to make an accurate assessment of the uncertainty arising from this source, but presently a factor of 2 uncertainty on the Higgs production cross section is probably not unrealistic. Measurements of exclusive diffraction at HERA can help constrain the generalized gluon distribution in kinematics similar to the one relevant for exclusive Higgs production at the LHC [14]. High-quality data are now available for $ep \to e\, J/\Psi\, p$. Exclusive production of Υ mesons and deeply virtual Compton scattering $ep \to e\gamma p$ involve smaller theoretical uncertainties, but are experimentally more demanding and will hopefully be explored in more detail with HERA II data.

Since the focus is on exclusive final states, such as $p \oplus H \oplus p$, it is necessary to sum the Sudakov logarithms which arise in perturbation theory. Importantly, one must go beyond summing the leading double logarithms and sum also the single logarithms. Without the single logs, one vastly underestimates the production rate. Unfortunately, perturbative emissions are not the only way to spoil the exclusive nature of the final state: extra particles can also be produced as a result of soft interactions between the colliding protons. To account for such soft interactions is clearly outside of the scope of perturbation theory and one is forced to resort to non-perturbative models. It is universally assumed that one can estimate the effect of forbidding additional particle production by simply multiplying the perturbative cross-section by an overall "gap survival" factor [15]. The two most sophisticated models of this factor are discussed in some detail and compared to each other in the contribution of Gotsman et al. [16]. It turns out that, although the approaches are different in many respects, they tend to predict very similar values for the gap survival factor. Nevertheless, both models are essentially multi-channel eikonal models and one would like to test them against data. Fortunately that is possible: data from HERA and the Tevatron already tend to support the theoretical models and future measurements at the LHC will allow one to further constrain them.

Uncertainties in the gluon densities and in our knowledge of gap survival can be reduced as we test our ideas against data, both at present colliders and at the LHC itself. Fortunately, these uncertainties essentially factorize (from the hard subprocess which produces the central system) into a universal "effective gluon luminosity" function. Thus one can hope to extract the important physics associated with the production of the central system by first measuring the luminosity function in a "standard candle" process. The ideal candidate is $pp \to p+\gamma\gamma+p$ [17] since the hard subprocess is well known ($gg \to \gamma\gamma$) and the effective gluon luminosity can be extracted over a wide kinematic range. In this way one might hope to extract the effective coupling of any centrally produced new physics to two gluons.

During the period of the workshop, Monte Carlo codes have been developed which simulate the theoretical predictions for both interesting signal processes and also the associated backgrounds. These codes are now routinely used, for example to help develop the case for the installation of low angle

[1] For a recent review of the physics case for FP420, see [11] and references therein

proton detectors at the LHC, and new processes are being added with time. A review and comparison of the various Monte Carlos is to be found in the contribution of Boonekamp et al. [18].

3 Diffractive structure functions and diffractive parton distributions

The cross section for the reaction $ep \to eXp$ can be expressed in terms of the diffractive structure functions F_2^D and F_L^D, in analogy to the way in which $d\sigma/dx\, dQ^2$ is related the structure functions F_2 and F_L for inclusive DIS, $ep \to eX$. The function F_2^D describes the proton structure in processes in which a fast proton is present in the final state; F_L^D corresponds to longitudinal polarization of the virtual photon. Since in diffractive events the proton typically loses a fraction of less than 0.02–0.03 of its initial momentum, the parton participating in a diffractive interaction has a fractional momentum which is also less than 0.02–0.03. Diffractive DIS thus probes the low-x structure of the proton, in a way complementary to that provided by non-diffractive DIS.

Diffractive structure functions, like the usual ones, can be expressed as the convolution of universal partonic cross sections and a specific type of parton distribution functions, the diffractive PDFs. This is the so called diffractive factorization theorem. Diffractive PDFs can be determined by means of QCD fits similar to those used for extracting the standard PDFs from the F_2 data.

Several measurements of F_2^D are available from the H1 and ZEUS collaborations. Three alternative approaches have been used to select diffractive events:

1. a fast proton is required in the final state; this can be done only be means of a proton spectrometer able to detect scattered protons which do not leave the beam pipe (see e.g. [19]);
2. a rapidity gap in the forward direction is required;
3. the different shape of the M_X distribution for diffractive and non-diffractive events is exploited.

Method 1 selects the reaction $ep \to eXp$ with a high degree of purity; the acceptance of proton spectrometers is however small, yielding comparatively small samples. Methods 2 and 3 select the reaction $ep \to eXY$, as opposed to $ep \to eXp$, with Y a proton or a low-mass system. Samples selected with these two methods may include some contamination from non-diffractive processes. Method 3 suppresses the contribution of subleading exchanges (i.e. Reggeon and pion exchanges, as opposed to Pomeron exchange), which is instead present in the samples selected with methods 1 and 2.

Results obtained with the three methods are presented and compared in these proceedings [21]. Methods 2 and 3 yield results for F_2^D which are higher than those obtained with the LPS by factors as large as 1.4, depending on the degree of forward coverage. This normalization difference is due to the proton-dissociative background (from $ep \to eXY$) and is relatively well understood. Having corrected for this effect, the results of the three methods exhibit, at present, a fair degree of agreement. Differences in the shapes of the Q^2, β and $x_{I\!P}$ dependences become however apparent especially when comparing the results obtained with method 3 and those obtained with methods 1 and 2. The origin of these differences is at present not clear. An urgent task for the HERA community is to understand these discrepancies and provide a consistent set of measurements of F_2^D.

Several NLO fits of the F_2^D data were discussed at the workshop [20, 21, 22]. The corresponding parameterizations are available in [23]. The diffractive PDFs are dominated by gluons, as expected given the low-x region probed, with the density of gluons larger than that of quarks by a factor 5–10. There are significant discrepancies between the results of the fits, reflecting, at least in part, the differences in the fitted data. In addition, Martin, Ryskin and Watt [22] argue that the leading-twist formula used in [20, 21] is inadequate in large parts of the measured kinematics, and use a modified expression which includes an estimate of power-suppressed effects. The discrepancies between the various diffractive PDFs, while not fully understood, are at the moment the best estimate of their uncertainties. Here as well, it is imperative that the HERA community provide a consistent set of diffractive PDFs. Not only are they important for our understanding of the proton structure, but also they are an essential input for any calculation of the cross sections for *inclusive* diffractive reactions at the LHC – which are interesting in themselves in addition to being a potentially dangerous background to the central *exclusive* production processes discussed in Sections 1 and 2.

No direct measurement exists of F_L^D. The dominant role played by gluons in the diffractive parton densities implies that the leading-twist F_L^D must also be relatively large. A measurement of F_L^D to even modest precision would provide an independent and theoretically very clean tool to verify our understanding of the underlying dynamics and to test the gluon density extracted indirectly in QCD fits from the scaling violations of F_2^D. This is discussed in [24].

4 Diffractive charm and dijet production at HERA

As mentioned in Section 2, the possibility to observe central exclusive processes depends critically on the survival probability of large rapidity gaps. This probability is not unity as a consequence of the rescattering between the spectator partons in the colliding hadrons: these interactions generate final-state particles which fill the would-be rapidity gap and slow down the outgoing proton/antiproton [16]. This is why the diffractive factorization theorem [25] is expected to fail for hadron-hadron scattering – and therefore also for resolved photoproduction, where the photon acts as a hadron.

In $p\bar{p}$ collisions at the Tevatron, breaking of diffractive factorization was indeed observed. The fraction of diffractive dijet events at CDF is a factor 3 to 10 smaller than that predicted on the basis of the diffractive parton densities obtained at HERA. Similar suppression factors were observed in all hard diffractive processes in proton-antiproton collisions.

In photoproduction processes, the situation is however far from clear at the moment. A recent ZEUS result [26] indicates that the cross section for diffractive photoproduction of D^* mesons, a process dominated by the *direct* photon component, is well described by NLO QCD predictions based on the diffractive PDFs. This lends support to the idea that in direct processes the photon is pointlike and that the diffractive factorization theorem holds in this case. Conversely, diffractive dijet data from H1 and ZEUS are better described by a global suppression of *both* the direct and resolved contribution. A discussion of how this might be understood is given in [27, 28], where a critical study of the factorization scheme and scale dependence of resolved and direct contributions is presented.

5 Multiple scattering at HERA and the LHC

A thorough analysis of the event structure at the LHC will have to take into account contributions from multiple parton interactions, i.e. from interactions involving more than one parton in each of the colliding protons. Such multiple interactions are expected to be particularly important in the region of small longitudinal momentum fractions and not too high momentum scales. At HERA there are several pieces of evidence that multiple interactions are present; the strongest one comes from the observation of diffractive final states in deep inelastic electron proton scattering. A useful tool for analyzing these multiple interactions are the so-called AGK cutting rules. During this workshop several groups have studied their application to HERA and to future LHC scattering processes.

In [29] the theoretical basis of the AGK rules in perturbative QCD has been outlined, and a few first applications to HERA and to LHC scattering processes have been addressed. Contribution [30] uses the AGK rules for deriving, from the measured *diffractive* structure function, absorptive corrections to the *inclusive* structure function F_2. An iterative scheme is then set up which leads to corrected parton densities: at low Q^2 and small x, they tend to be higher than those without absorptive corrections. In particular, they seem to weaken the trend of the gluon density becoming negative, which has been seen in the global parton analyses of both MRST2004 and CTEQ6. The study presented in [31] is based upon a specific saturation model that has been successfully applied both to the total $\gamma^* p$ cross section and to the diffractive process $\gamma^* p \to J/\Psi\, p$. An analysis of this model, based upon the AGK rules, leads to the conclusion that contributions of multiple interactions to F_2 are quite sizable, even for Q^2 as large as $40\ \mathrm{GeV}^2$.

6 Parton saturation: from HERA to the LHC

A key experimental finding of HERA is the strong rise of structure functions at small x, which implies a high density of small-x gluons in the proton. From theoretical considerations, it is clear that for sufficiently large parton densities, dynamics beyond what can be described by leading-twist factorization and linear DGLAP evolution must become important. If the associated momentum scale is high enough, the strong coupling is still small enough to serve as an expansion parameter, but at very high gluon densities the gluon potential can be so strong that the non-linear term $g_s f^{abc} A_\mu^b A_\nu^c$ in the gluon field strength is as large as the linear term $\partial_\mu A_\nu^a - \partial_\nu A_\mu^a$. High parton densities thus offer the possibility to study QCD in a strongly non-linear regime, and the effective theory of such a "color glass condensate" is reviewed in [32]. A possible link between the strong gluon fields in this description and QCD instantons is discussed in [33].

The theory and phenomenology of parton saturation are in rapid development, of which the workshop could only provide a snapshot. Data on both inclusive and diffractive deeply inclusive scattering, in particular their very similar energy dependence at given Q^2, suggest that saturation effects are relevant in HERA kinematics, see [34] and references therein. When saturation is important, the usual parton densities cease to be the key input quantities for describing physical processes. For many reactions a suitable quantity is instead the color-dipole cross section – a concept that has been successfully applied in HERA phenomenology. An important theoretical laboratory to study saturation effects is provided by the non-linear Balitsky-Kovchegov equation. In a contribution to the workshop, this equation has been applied to the color-dipole cross section for the proton [35]. To describe saturation in pp collisions in general requires non-perturbative functions that can be written as matrix elements of Wilson line operators; *one* of these functions is the color-dipole cross section just mentioned [32]. The formulation of suitable evolution equations for pp scattering is an active area of research [36].

7 Rapidity gaps in electroweak processes

Diffractive processes are characterized by rapidity gaps. Such gaps can however also originate from the exchange of a photon, a W or a Z boson (see for example [15]). Selecting events with large rapidity gaps filters out specific final states and, at the same time, leads to better-constrained event kinematics. The event rate is however lowered by the gap survival probability, as discussed in the previous sections.

The contribution by Amapane et al. [37] discusses the possibility to study the scattering of longitudinally polarized vector bosons (V_L) in pp collisions with the CMS detector at the LHC. $V_L V_L$ fusion may lead to Higgs production; should the Higgs boson not exist, the cross section for $V_L V_L$ scattering will deviate from the Standard Model prediction at high invariant masses of the $V_L V_L$ system. In all cases, $V_L V_L$ scattering should shed light on the mechanism behind the electro-weak symmetry breaking. Preliminary studies based on Pythia and a fast simulation of the CMS detector are encouraging. It will be interesting to investigate in more detail the potential of the rapidity gap signature for improved signal extraction and background control.

Large rapidity gaps at hadron colliders can also be due to photon exchange. In this case, a direct tagging of high-energy photon interactions can be achieved by using forward proton detectors [38]. Both photon-photon and photon-proton interactions at the LHC have been studied [39]. Some of these events can be used to scan the gap survival probability in impact parameter space, which would help to constrain models for gap survival. A reference point is given by single W boson photoproduction, which has been studied theoretically in this context [40] and is being investigated at HERA. Finally, diffractive photoproduction of Υ mesons, presently studied at HERA, can be accessed at the LHC in an extended range of small x. This will provide a very clean channel to study the generalized gluon distribution (see Section 2) and can be seen as a complement to measurements of the usual gluon distribution at very small x, for instance in forward jet production at the LHC.

Acknowledgments

We wish to thank all participants of the Working Group on Diffraction for their valuable contributions to this Workshop. This work has been supported by PPARC and the Royal Society in the UK, and the Italian Ministry for Education, University and Scientific Research under the program "Incentivazione alla mobilità di studiosi stranieri e italiani residenti all'estero".

References

[1] M. G. Albrow *et al.*, Proposal submitted to the LHCC: CERN-LHCC-2005-025, LHCC-I-015.

[2] M. G. Albrow and A. Rostovtsev, hep-ph/0009336.

[3] J. Kalliopuska *et al.*, these proceedings.

[4] V. Avati and K. Österberg, these proceedings.

[5] B. E. Cox *et al.*, hep-ph/0505240.

[6] A. B. Kaidalov, V. A. Khoze, A. D. Martin and M. G. Ryskin, Eur. Phys. J. C **33** (2004) 261 [hep-ph/0311023].

[7] J. R. Ellis, J. S. Lee and A. Pilaftsis, Phys. Rev. D **71** (2005) 075007 [hep-ph/0502251].

[8] V. A. Khoze, A. D. Martin and M. G. Ryskin, Eur. Phys. J. C **34** (2004) 327 [hep-ph/0401078].

[9] M. Arneodo *et al.*, these proceedings.

[10] G. Bruni, G. Iacobucci, L. Rinaldi and M. Ruspa, these proceedings.

[11] B. E. Cox, AIP Conf. Proc. **753** (2005) 103 [hep-ph/0409144].

[12] V. Andreev *et al.*, these proceedings.

[13] J. R. Forshaw, these proceedings.

[14] J. Collins *et al.*, these proceedings.

[15] J. D. Bjorken, Phys. Rev. D **47** (1993) 101.

[16] E. Gotsman, E. Levin, U. Maor, E. Naftali and A. Prygarin, these proceedings.

[17] V. A. Khoze, A. D. Martin, M. G. Ryskin and W. J. Stirling, Eur. Phys. J. C **38** (2005) 475 [hep-ph/0409037].

[18] M. Boonekamp *et al.*, these proceedings.

[19] R. Sacchi, these proceedings (web version only).

[20] M. Groys, A. Levy and A. Proskuryakov, these proceedings.

[21] P. Newman and F.-P. Schilling, these proceedings.

[22] A. Martin, M. Ryskin and G. Watt, these proceedings.

[23] F.-P. Schilling, these proceedings.

[24] P. Newman, these proceedings.

[25] J. C. Collins, Phys. Rev. D **57** (1998) 3051, Erratum-ibid. D **61** (2000) 019902 [hep-ph/9709499].

[26] ZEUS Collaboration, Paper 268 submitted to the XXII International Symposium on Lepton-Photon Interactions at High Energy, Uppsala, Sweden, June 2005, avalaible from http:www-zeus.desy.de ; I. Melzer-Pellman for the ZEUS Collaboration, talk at International Conference on the Structure and Interactions of the Photon (PHOTON 2005), Warsaw, Poland, August 2005.

[27] A. Bruni, M. Klasen, G. Kramer and S. Schätzel, these proceedings.

[28] M. Klasen and G. Kramer, hep-ph/0506121, to appear in J. Phys. G.

[29] J. Bartels, these proceedings.

[30] G. Watt, A. D. Martin and M. G. Ryskin, these proceedings.

[31] H. Kowalski, these proceedings.

[32] R. Venugopalan, these proceedings.

[33] F. Schrempp, these proceedings.

[34] M. Arneodo and M. Diehl, these proceedings.

[35] E. Gotsman, E. Levin, U. Maor and E. Naftali, these proceedings (web version only).

[36] M. Lublinsky, these proceedings (web version only).

[37] N. Amapane et al., these proceedings.

[38] K. Piotrzkowski, Phys. Rev. D **63** (2001) 071502 [hep-ex/0009065].

[39] K. Piotrzkowski, these proceedings (web version only).

[40] V. A. Khoze, A. D. Martin and M. G. Ryskin, Eur. Phys. J. C **24** (2002) 459 [hep-ph/0201301].

Diffraction for non-believers

Michele Arneodo[a] *and Markus Diehl*[b]
[a]Università del Piemonte Orientale, 28100 Novara, Italy
[b]Deutsches Elektronen-Synchroton DESY, 22603 Hamburg, Germany

Abstract
Diffractive reactions involving a hard scale can be understood in terms of quarks and gluons. These reactions have become a valuable tool for investigating the low-x structure of the proton and the behavior of QCD in the high-density regime, and they may provide a clean environment to study or even discover the Higgs boson at the LHC. In this paper we give a brief introduction to the description of diffraction in QCD. We focus on key features studied in ep collisions at HERA and outline challenges for understanding diffractive interactions at the LHC.

1 Introduction

In hadron-hadron scattering a substantial fraction of the total cross section is due to diffractive reactions. Figure 1 shows the different types of diffractive processes in the collision of two hadrons: in elastic scattering both projectiles emerge intact in the final state, whereas single or double diffractive dissociation corresponds to one or both of them being scattered into a low-mass state; the latter has the same quantum numbers as the initial hadron and may be a resonance or continuum state. In all cases, the energy of the outgoing hadrons a, b or the states X, Y is approximately equal to that of the incoming beam particles, to within a few per cent. The two (groups of) final-state particles are well separated in phase space and in particular have a large gap in rapidity between them.

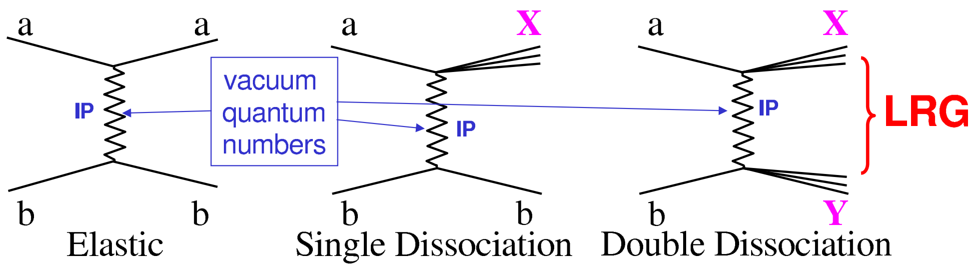


Fig. 1: Elastic scattering, single diffractive dissociation and double diffractive dissociation in the collision of two hadrons a and b. The two (groups of) final-state hadrons are separated by a large rapidity gap (LRG). The zigzag lines denote the exchange of a Pomeron ($I\!\!P$) in the t-channel. There are further graphs, not shown, with multiple Pomeron exchange.

Diffractive hadron-hadron scattering can be described within Regge theory (see e.g. [1]). In this framework, the exchange of particles in the t-channel is summed coherently to give the exchange of so-called "Regge trajectories". Diffraction is characterized by the exchange of a specific trajectory, the "Pomeron", which has the quantum numbers of the vacuum. Regge theory has spawned a successful phenomenology of soft hadron-hadron scattering at high energies. Developed in the 1960s, it predates the theory of the strong interactions, QCD, and is based on general concepts such as dispersion relations. Subsequently it was found that QCD perturbation theory in the high-energy limit can be organized following the general concepts of Regge theory; this framework is often referred to as BFKL after the authors of the seminal papers [2].

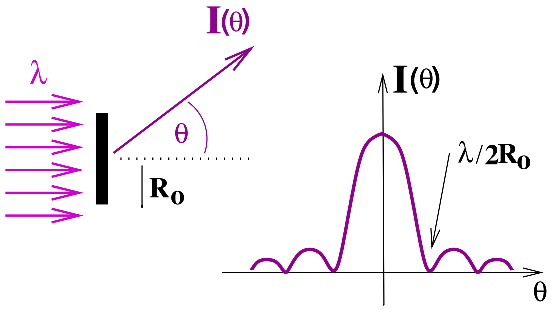


Fig. 2: Distribution of the intensity I in the diffraction of light of wavelength λ from a circular target of size R_0.

It is clear that a t-channel exchange leading to a large rapidity gap in the final state must carry zero net color: if color were exchanged, the color field would lead to the production of further particles filling any would-be rapidity gap. In QCD, Pomeron exchange is described by the exchange of two interacting gluons with the vacuum quantum numbers.

The effort to understand diffraction in QCD has received a great boost from studies of diffractive events in ep collisions at HERA (see e.g. [3] for further reading and references). The essential results of these studies are discussed in the present paper and can be summarized as follows:

- Many aspects of diffraction are well understood in QCD when a hard scale is present, which allows one to use perturbative techniques and thus to formulate the dynamics in terms of quarks and gluons. By studying what happens when the hard scale is reduced towards the non-perturbative region, it may also be possible to shed light on soft diffractive processes.
- Diffraction has become a tool to investigate low-momentum partons in the proton, notably through the study of diffractive parton densities in inclusive processes and of generalized parton distributions in exclusive ones. Diffractive parton densities can be interpreted as conditional probabilities to find a parton in the proton when the final state of the process contains a fast proton of given four-momentum. Generalized parton distributions, through their dependence on both longitudinal and transverse variables, provide a three-dimensional picture of the proton in high-energy reactions.
- A fascinating link has emerged between diffraction and the physics of heavy-ion collisions through the concept of saturation, which offers a new window on QCD dynamics in the regime of high parton densities.

Perhaps unexpectedly, the production of the Higgs boson in diffractive pp collisions is drawing more and more attention as a clean channel to study the properties of a light Higgs boson or even discover it. This is an example of a new theoretical challenge: to adapt and apply the techniques for the QCD description of diffraction in ep collisions to the more complex case of pp scattering at the LHC. A first glimpse of phenomena to be expected there is provided by the studies of hard diffraction in $p\bar{p}$ collisions at the Tevatron.

1.1 A digression on the nomenclature: why "diffraction" ?

Physics students first encounter the term "diffraction" in optics. Light of wavelength λ impinging on a black disk of radius R_0 produces on a distant screen a diffraction pattern, characterized by a large forward peak for scattering angle $\theta = 0$ (the "diffraction peak") and a series of symmetric minima and maxima, with the first minimum at $\theta_{\rm min} \simeq \pm\lambda/(2R_0)$ (Fig. 2). The intensity I as a function of the scattering angle θ is given by

$$\frac{I(\theta)}{I(\theta=0)} = \frac{[2J_1(x)]^2}{x^2} \simeq 1 - \frac{R_0^2}{4}(k\theta)^2, \tag{1}$$

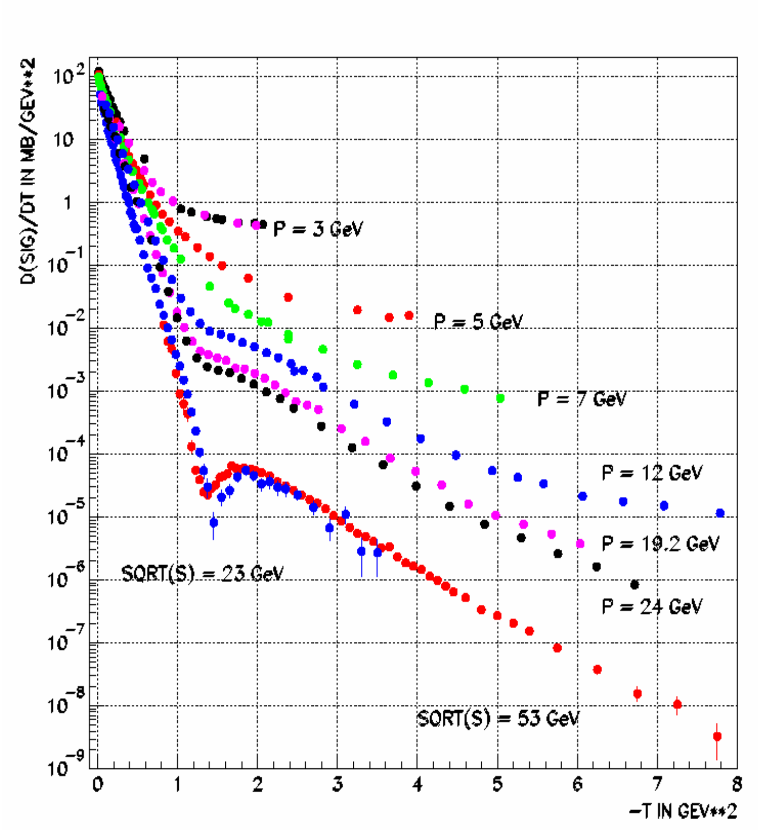


Fig. 3: Compilation of proton-proton elastic cross section data as a function of t. The symbol P indicates the momentum of the incoming proton in a fixed target experiment and $\sqrt{s}$ the center-of-mass energy in a pp collider setup.

where J_1 is the Bessel function of the first order and $x = kR_0 \sin\theta \simeq kR_0\,\theta$ with $k = 2\pi/\lambda$. The diffraction pattern is thus related to the size of the target and to the wavelength of the light beam.

As shown in Fig. 3, the differential cross section $d\sigma/dt$ for elastic proton-proton scattering, $pp \to pp$, bears a remarkable resemblance to the diffraction pattern just described (see e.g. [4]). At low values of $|t|$ one has

$$\frac{\frac{d\sigma}{dt}(t)}{\frac{d\sigma}{dt}(t=0)} \simeq e^{-b|t|} \simeq 1 - b\,(P\theta)^2, \tag{2}$$

where $|t| \simeq (P\theta)^2$ is the absolute value of the squared four-momentum transfer, P is the incident proton momentum and θ is the scattering angle. The t-slope b can be written as $b = R^2/4$, where once again R is related to the target size (or more precisely to the transverse distance between projectile and target). A dip followed by a secondary maximum has also been observed, with the value of $|t|$ at which the dip appears decreasing with increasing proton momentum. It is hence not surprising that the term diffraction is used for elastic pp scattering. Similar t distributions have been observed for the other diffractive reactions mentioned above, leading to the use of the term diffraction for all such processes.

1.2 Diffraction at HERA ?!

Significant progress in understanding diffraction has been made at the ep collider HERA, where 27.5 GeV electrons or positrons collide with 820 or 920 GeV protons. This may sound peculiar: diffraction is a typical hadronic process while ep scattering at HERA is an electro-weak reaction, where the electron

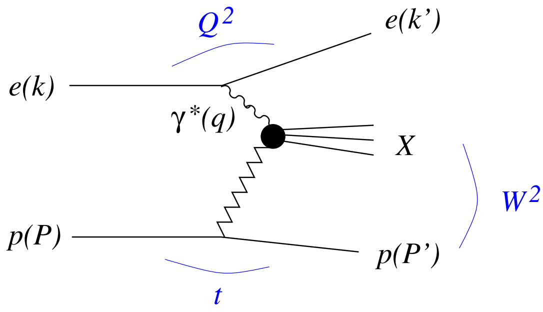


Fig. 4: Schematic diagram of inclusive diffractive DIS, $ep \to eXp$. Four-momenta are indicated in parentheses.

radiates a virtual photon (or a Z or W boson), which then interacts with the proton.[1] To understand this, it is useful to look at ep scattering in a frame where the virtual photon moves very fast (for instance in the proton rest frame, where the γ^* has a momentum of up to about 50 TeV at HERA). The virtual photon can fluctuate into a quark-antiquark pair. Because of its large Lorentz boost, this virtual pair has a lifetime much longer than a typical strong interaction time. In other words, the photon fluctuates into a pair long before the collision, and it is the pair that interacts with the proton. This pair is a small color dipole. Since the interaction between the pair and the proton is mediated by the strong interaction, diffractive events are possible.

An advantage of studying diffraction in ep collisions is that, for sufficiently large photon virtuality Q^2, the typical transverse dimensions of the dipole are small compared to the size of a hadron. Then the interaction between the quark and the antiquark, as well as the interaction of the pair with the proton, can be treated perturbatively. With decreasing Q^2 the color dipole becomes larger, and at very low Q^2 these interactions become so strong that a description in terms of quarks and gluons is no longer justified. We may then regard the photon as fluctuating into a vector meson – this is the basis of the well-known vector meson dominance model – and can therefore expect to see diffractive reactions very similar to those in hadron-hadron scattering.

A different physical picture is obtained in a frame where the incident proton is very fast. Here, the diffractive reaction can be seen as the deep inelastic scattering (DIS) of a virtual photon on the proton target, with a very fast proton in the final state. One can thus expect to probe partons in the proton in a very specific way. For suitable diffractive processes there are in fact different types of QCD factorization theorems, which bear out this expectation (see Sects. 2 and 3).

2 Inclusive diffractive scattering in *ep* collisions

Figure 4 shows a schematic diagram of inclusive diffractive DIS. The following features are important:

- The proton emerges from the interaction carrying a large fraction x_L of the incoming proton momentum. Diffractive events thus appear as a peak at $x_L \approx 1$, the diffractive peak, which at HERA approximately covers the region $0.98 < x_L < 1$ (see the left panel of Fig. 5). The right panel of Fig. 5 shows that large values of $|t|$ are exponentially suppressed, similarly to the case of elastic pp scattering we discussed in Sect. 1.1. These protons remain in the beam-pipe and can only be measured with detectors located inside the beam-pipe.
- The collision of the virtual photon with the proton produces a hadronic final state X with the photon quantum numbers and invariant mass M_X. A large gap in rapidity (or pseudorapidity) is present between X and the final-state proton. Figure 6 shows a typical diffractive event at HERA.

[1]For simplicity we will speak of a virtual photon in the following, keeping in mind that one can have a weak gauge boson instead.

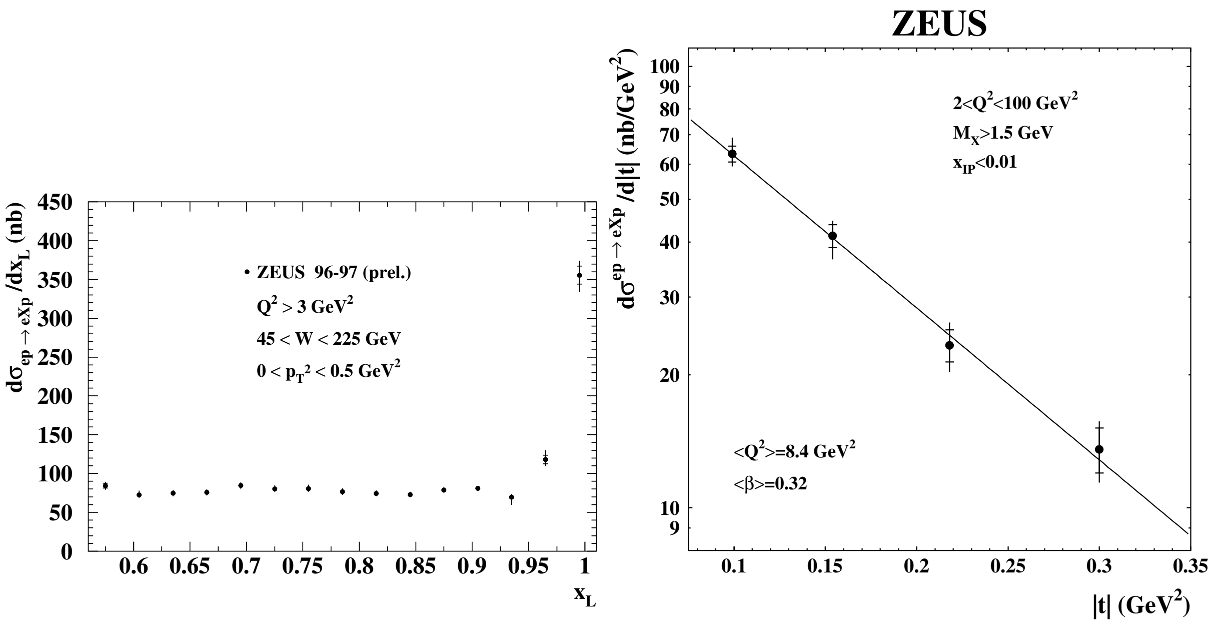


Fig. 5: Left: Differential cross section $d\sigma/dx_L$ for the process $ep \to eXp$ (from [5]). The diffractive peak at $x_L \approx 1$ is clearly visible. Right: Differential cross section $d\sigma/dt$ for the same process for $x_L > 0.99$ (from [6]). The average $|t|$ of this spectrum is $\langle |t| \rangle \approx 0.15$ GeV^2.

Diffractive ep scattering thus combines features of hard and soft scattering. The electron receives a large momentum transfer; in fact Q^2 can be in the hundreds of GeV^2. In contrast, the proton emerges with its momentum barely changed.

2.1 Diffractive structure functions

The kinematics of $\gamma^* p \to Xp$ can be described by the invariants $Q^2 = -q^2$ and $t = (P - P')^2$, and by the scaling variables $x_{I\!P}$ and β given by

$$x_{I\!P} = \frac{(P-P')\cdot q}{P \cdot q} = \frac{Q^2 + M_X^2 - t}{W^2 + Q^2 - M_p^2}, \qquad \beta = \frac{Q^2}{2(P-P')\cdot q} = \frac{Q^2}{Q^2 + M_X^2 - t}, \tag{3}$$

where $W^2 = (P+q)^2$ and the four-momenta are defined in Fig. 4. The variable $x_{I\!P}$ is the fractional momentum loss of the incident proton, related as $x_{I\!P} \simeq 1 - x_L$ to the variable x_L introduced above. The quantity β has the form of a Bjorken variable defined with respect to the momentum $P - P'$ lost by the initial proton instead of the initial proton momentum P. The usual Bjorken variable $x_B = Q^2/(2P\cdot q)$ is related to β and $x_{I\!P}$ as $\beta x_{I\!P} = x_B$.

The cross section for $ep \to eXp$ in the one-photon exchange approximation can be written in terms of diffractive structure functions $F_2^{D(4)}$ and $F_L^{D(4)}$ as

$$\frac{d\sigma^{ep\to eXp}}{d\beta\, dQ^2\, dx_{I\!P}\, dt} = \frac{4\pi\alpha_{\rm em}^2}{\beta Q^4}\left[\left(1 - y + \frac{y^2}{2}\right) F_2^{D(4)}(\beta, Q^2, x_{I\!P}, t) - \frac{y^2}{2} F_L^{D(4)}(\beta, Q^2, x_{I\!P}, t)\right], \tag{4}$$

in analogy with the way $d\sigma^{ep\to eX}/(dx_B\, dQ^2)$ is related to the structure functions F_2 and F_L for inclusive DIS, $ep \to eX$. Here $y = (P\cdot q)/(P\cdot k)$ is the fraction of energy lost by the incident lepton in the proton rest frame. The structure function $F_L^{D(4)}$ corresponds to longitudinal polarization of the virtual photon; its contribution to the cross section is small in a wide range of the experimentally accessible kinematic

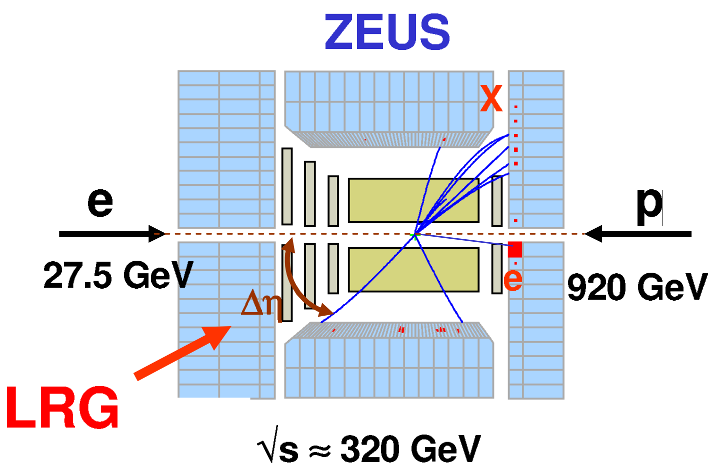


Fig. 6: A DIS event with a large rapidity gap (LRG) observed with the ZEUS detector at HERA. The scattered proton escapes into the beam-pipe. The symbol $\Delta\eta$ denotes the difference in pseudorapidity between the scattered proton and the most forward particle of the observed hadronic system X. Pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$ in terms of the polar angle θ measured with respect to the incoming proton direction, which is defined as "forward".

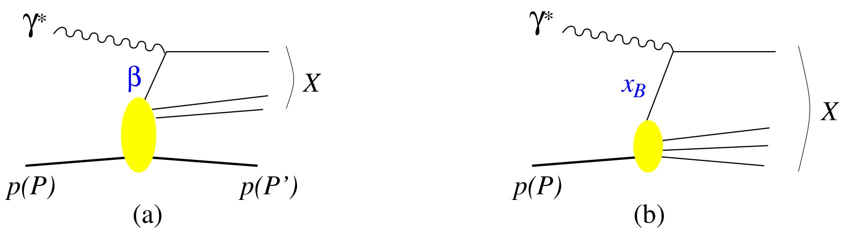


Fig. 7: Parton model diagrams for deep inelastic diffractive (a) and inclusive (b) scattering. The variable β is the momentum fraction of the struck quark with respect to $P - P'$, and x_B its momentum fraction with respect to P.

region (in particular at low y). The structure function $F_2^{D(3)}$ is obtained from $F_2^{D(4)}$ by integrating over t:

$$F_2^{D(3)}(\beta, Q^2, x_{I\!P}) = \int dt\, F_2^{D(4)}(\beta, Q^2, x_{I\!P}, t). \tag{5}$$

In a parton model picture, inclusive diffraction $\gamma^* p \to Xp$ proceeds by the virtual photon scattering on a quark, in analogy to inclusive scattering (see Fig. 7). In this picture, β is the momentum fraction of the struck quark with respect to the exchanged momentum $P - P'$ (indeed the allowed kinematical range of β is between 0 and 1). The diffractive structure function describes the proton structure in these specific processes with a fast proton in the final state. F_2^D may also be viewed as describing the structure of whatever is exchanged in the t-channel in diffraction, i.e. of the Pomeron (if multiple Pomeron exchange can be neglected). It is however important to bear in mind that the Pomeron in QCD cannot be interpreted as a particle on which the virtual photon scatters, as we will see in Sect. 2.5.

Figures 8 and 9 show recent H1 data [7] on $F_2^{D(3)}$ at fixed $x_{I\!P}$ as a function of β for different Q^2 bins, and as a function of Q^2 for different bins of β.[2] The data have two remarkable features:

- F_2^D is largely flat in the measured β range. Keeping in mind the analogy between β in diffractive DIS and x_B in inclusive DIS, this is very different from the behavior of the "usual" structure function F_2, which strongly decreases for $x_B \gtrsim 0.2$ (see Fig. 8).

[2]To be precise, the H1 data are for the so-called reduced diffractive cross section, which equals $F_2^{D(3)}$ if F_L^D can be neglected.

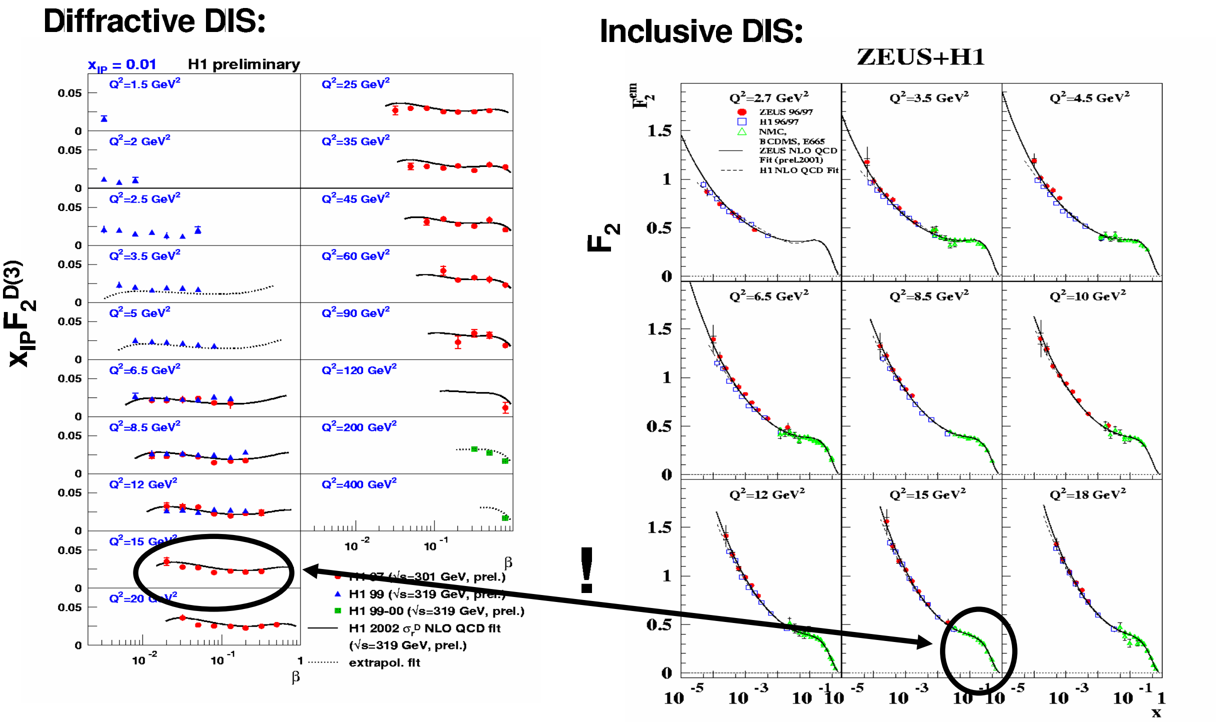


Fig. 8: Left: the diffractive structure function of the proton as a function of β (from [7]). Right: the structure function of the proton as a function of x_B (from [8]). The two highlighted bins show the different shapes of F_2^D and F_2 in corresponding ranges of β and x_B at equal Q^2.

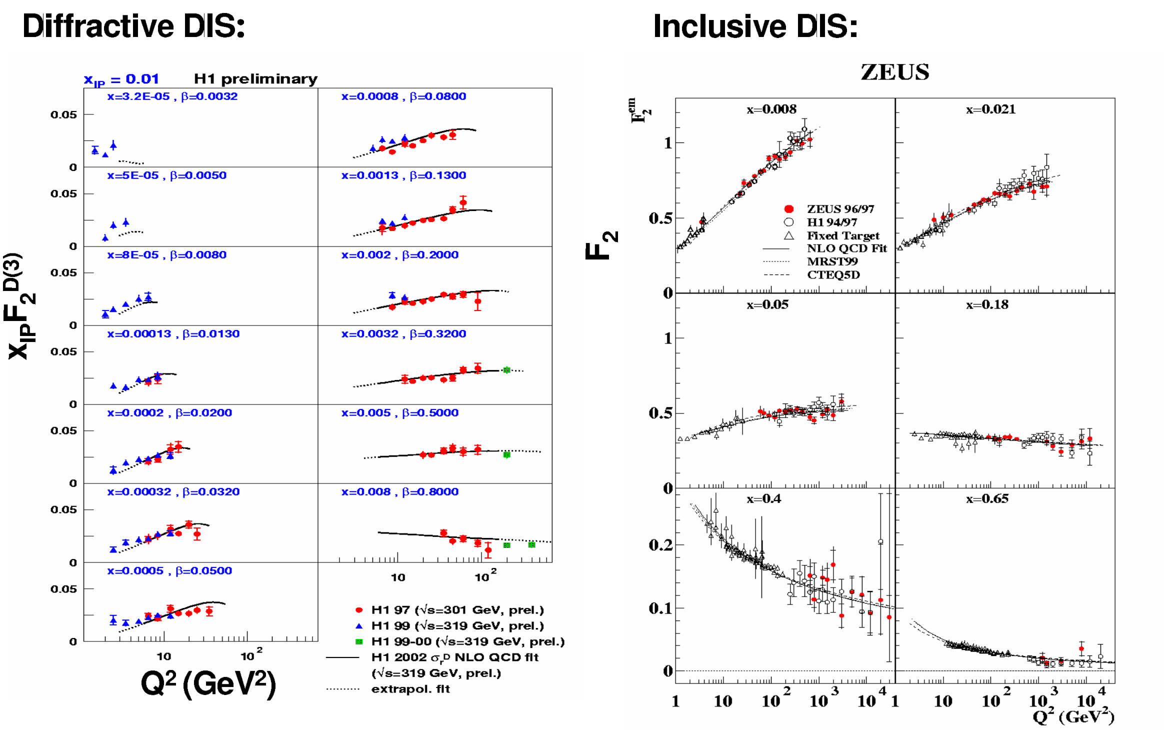


Fig. 9: Left: the diffractive structure function of the proton as a function of Q^2 (from [7]). Right: the structure function of the proton as a function of Q^2 (from [9]).

- The dependence on Q^2 is logarithmic, i.e. one observes approximate Bjorken scaling. This indicates the applicability of the parton model picture to inclusive $\gamma^* p$ diffraction. The structure function F_2^D increases with Q^2 for all β values except the highest. This is reminiscent of the scaling violations of F_2, except that F_2 rises with Q^2 only for $x_B \lesssim 0.2$ and that the scaling violations become negative at higher x_B (see Fig. 9). In the proton, negative scaling violations reflect the presence of the valence quarks radiating gluons, while positive scaling violations are due to the increase of the sea quark and gluon densities as the proton is probed with higher resolution. The F_2^D data thus suggest that the partons resolved in diffractive events are predominantly gluons. This is not too surprising if one bears in mind that these partons carry only a small part of the proton momentum: the struck quark in the diagram of Fig. 7a has a momentum fraction $\beta x_{I\!P} = x_B$ with respect to the incident proton, and $x_{I\!P} \lesssim 0.02 – 0.03$ in diffractive events.

2.2 Diffractive parton distributions

The conclusion just reached can be made quantitative by using the QCD factorization theorem for inclusive diffraction, $\gamma^* p \to Xp$, which formalizes the parton model picture we have already invoked in our discussion. According to this theorem, the diffractive structure function, in the limit of large Q^2 at fixed β, $x_{I\!P}$ and t, can be written as [10, 11, 12]

$$F_2^{D(4)}(\beta, Q^2, x_{I\!P}, t) = \sum_i \int_\beta^1 \frac{dz}{z}\, C_i\Big(\frac{\beta}{z}\Big)\, f_i^D(z, x_{I\!P}, t; Q^2), \tag{6}$$

where the sum is over partons of type i. The coefficient functions C_i describe the scattering of the virtual photon on the parton and are exactly the same as in inclusive DIS. In analogy to the usual parton distribution functions (PDFs), the diffractive PDFs $f_i^D(z, x_{I\!P}, t; Q^2)$ can be defined as operator matrix elements in a proton state, and their dependence on the scale Q^2 is given by the DGLAP evolution equations. In parton model language, they can be interpreted as conditional probabilities to find a parton i with fractional momentum $z x_{I\!P}$ in a proton, probed with resolution Q^2 in a process with a fast proton in the final state (whose momentum is specified by $x_{I\!P}$ and t).

During the workshop, several fits of the available F_2^D data were discussed which are based on the factorization formula (6) at next-to-leading order (NLO) in α_s [13, 14]. Figure 10 compares the diffractive PDFs from an earlier H1 fit [7] to those from the fit of the ZEUS data [15] by Schilling and Newman [13]. As expected the density of gluons is larger than that of quarks, by about a factor 5–10. Discrepancies between the two sets are evident, and it remains to be clarified to which extent they reflect differences in the fitted data. Martin, Ryskin and Watt [16] have argued that the leading-twist formula (6) is inadequate in large parts of the measured kinematics, and performed a fit to a modified expression which includes an estimate of power-suppressed effects. The discrepancies between the various diffractive PDFs, while not fully understood, may be taken as an estimate of the uncertainties on these functions at this point in time. A precise and consistent determination of the diffractive PDFs and their uncertainties is one of the main tasks the HERA community has to face in the near future. They are a crucial input for predicting cross sections of inclusive diffractive processes at the LHC.

2.3 Diffractive hard-scattering factorization

Like usual parton densities, diffractive PDFs are process-independent functions. They appear not only in inclusive diffraction but also in other processes where diffractive hard-scattering factorization holds. In analogy with Eq. (6), the cross section of such a process can be evaluated as the convolution of the relevant parton-level cross section with the diffractive PDFs. For instance, the cross section for charm production in diffractive DIS can be calculated at leading order in α_s from the $\gamma^* g \to c\bar{c}$ cross section and the diffractive gluon distribution. An analogous statement holds for jet production in diffractive DIS. Both processes have been analyzed at next-to-leading order in α_s.

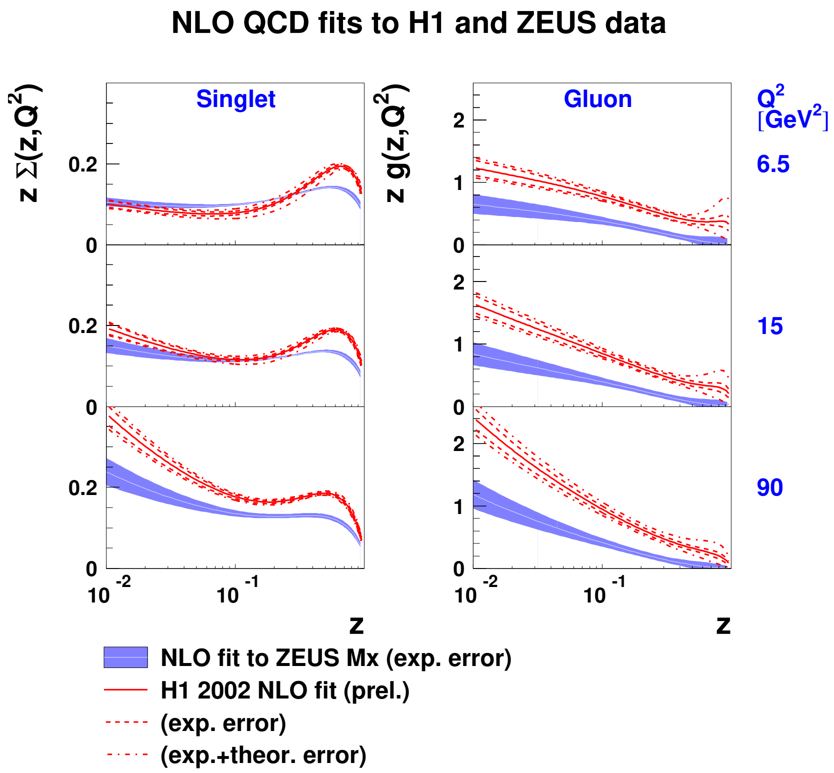
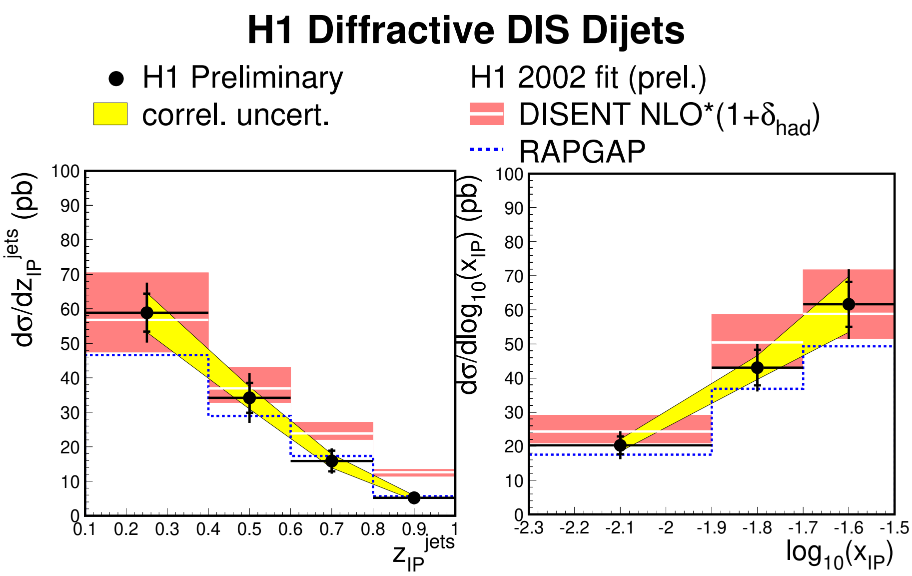


Fig. 10: Diffractive quark singlet and gluon distributions obtained from fits to H1 [7] and ZEUS [15] data (from [13]).



(a) (b)

Fig. 11: Cross section for dijet production in diffractive DIS, compared with the expectations based on the diffractive PDFs [7] (from [17]). The variable $z_{I\!P}^{\rm jets}$ estimates the fractional momentum of the parton entering the hard subprocess.

As an example, Fig. 11 shows a comparison between the measured cross sections for diffractive dijet production and the expectations based on diffractive PDFs extracted from a fit to F_2^D. These data lend support to the validity of hard-scattering factorization in diffractive $\gamma^* p$ interactions. For further discussion we refer the reader to [18].

2.4 Limits of diffractive hard-scattering factorization: hadron-hadron collisions

A natural question to ask is whether one can use the diffractive PDFs extracted at HERA to describe hard diffractive processes such as the production of jets, heavy quarks or weak gauge bosons in $p\bar{p}$ collisions at the Tevatron. Figure 12 shows results on diffractive dijet production from the CDF collaboration [19] compared to the expectations based on the diffractive PDFs [6, 7] from HERA. The discrepancy is spectacular: the fraction of diffractive dijet events at CDF is a factor 3 to 10 smaller than would be expected on the basis of the HERA data. The same type of discrepancy is consistently observed in all hard diffractive processes in $p\bar{p}$ events, see e.g. [20]. In general, while at HERA hard diffraction contributes a fraction of order 10% to the total cross section, it contributes only about 1% at the Tevatron.

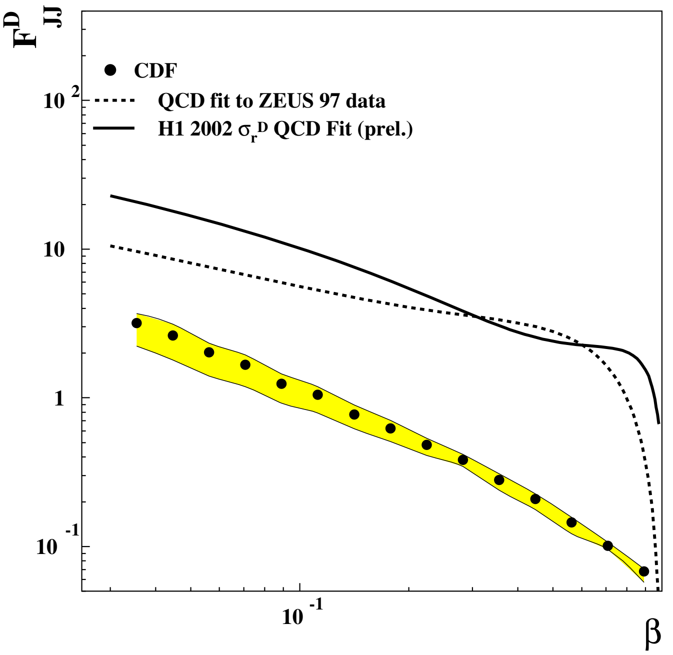


Fig. 12: CDF results for the cross section of diffractive dijet production with a leading antiproton in $p\bar{p}$ collisions (expressed in terms of a structure function F_{JJ}^D), compared with the predictions obtained from the diffractive PDFs [6] and [7] extracted at HERA (from [21]). See also the analogous plot in the original CDF publication [19].

In fact, diffractive hard-scattering factorization does not apply to hadron-hadron collisions [11, 12]. Attempts to establish corresponding factorization theorems fail because of interactions between spectator partons of the colliding hadrons. The contribution of these interactions to the cross section does not decrease with the hard scale. Since they are not associated with the hard-scattering subprocess (see Fig. 13), we no longer have factorization into a parton-level cross section and the parton densities of one of the colliding hadrons. These interactions are generally soft, and we have at present to rely on phenomenological models to quantify their effects [22].

The yield of diffractive events in hadron-hadron collisions is lowered precisely because of these soft interactions between spectator partons (often referred to as "reinteractions" or "multiple scatterings"). They can produce additional final-state particles which fill the would-be rapidity gap (hence the often-used term "rapidity gap survival"). When such additional particles are produced, a very fast proton

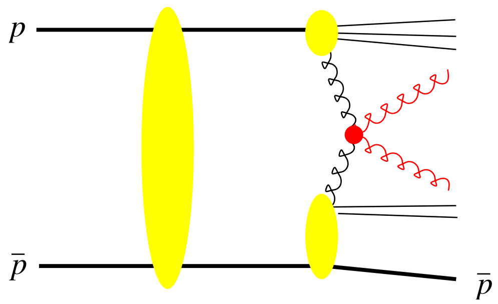


Fig. 13: Example graph for diffractive dijet production with a leading antiproton in a $p\bar{p}$ collision. The interaction indicated by the large vertical blob breaks hard diffractive factorization. It reduces the diffractive cross section, as explained in the text.

can no longer appear in the final state because of energy conservation. Diffractive factorization breaking is thus intimately related to multiple scattering in hadron-hadron collisions; understanding and describing this phenomenon is a challenge in the high-energy regime that will be reached at the LHC [23].

In pp or $p\bar{p}$ reactions, the collision partners are both composite systems of large transverse size, and it is not too surprising that multiple interactions between their constituents can be substantial. In contrast, the virtual photon in γ^*p collisions has small transverse size, which disfavors multiple interactions and enables diffractive factorization to hold. According to our discussion in Sect. 1.2, we may expect that for decreasing virtuality Q^2 the photon behaves more and more like a hadron, and diffractive factorization may again be broken. This aspect of diffractive processes in photoproduction at HERA was intensively discussed during the workshop, and findings are reported in [18].

2.5 Space-time structure: the Pomeron is not a particle

It is tempting to interpret diffractive γ^*p processes as the scattering of a virtual photon on a Pomeron which has been radiated off the initial proton. Diffractive DIS would then probe the distribution of partons in a “Pomeron target”. This is indeed the picture proposed by Ingelman and Schlein long ago [24].

This picture is however not supported by an analysis in QCD (see e.g. [25]). There, high-energy scattering is dominated by the exchange of two gluons, whose interaction is (in an appropriate gauge) described by ladder diagrams, as shown in Fig. 14. By analyzing these diagrams in time-ordered perturbation theory, one can obtain the dominant space-time ordering in the high-energy limit. The result depends on the reference frame, as illustrated in the figure. In the Breit frame, which is natural for a parton-model interpretation, the photon does *not* scatter off a parton in a pre-existing two-gluon system; in fact some of the interactions in the gluon ladder building up the Pomeron exchange take place long after the virtual photon has been absorbed. The picture in the Breit frame is however compatible with

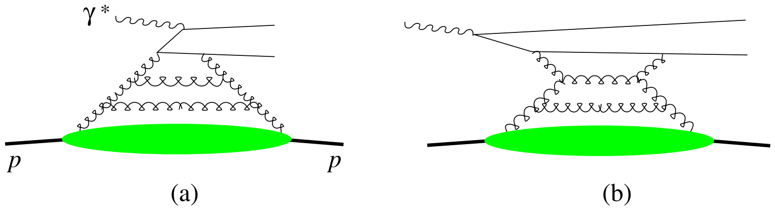


Fig. 14: Dominant time ordering for diffractive dissociation of a virtual photon in (a) the Breit frame, (b) the photon-proton center-of-mass, (c) the proton rest frame. The physical picture in (a) corresponds to the parton-model description of diffraction, and the one in (b) and (c) to the picture of the photon splitting into a quark-antiquark dipole which subsequently interacts with the proton.

the interpretation of diffractive parton densities given in Sect. 2.2, namely the probability to find a parton under the condition that subsequent interactions will produce a fast proton in the final state.

We note that the Ingelman-Schlein picture suggests that the diffractive structure function takes a factorized form $F_2^{D(4)} = f_{I\!P}(x_{I\!P}, t)\, F_2^{I\!P}(\beta, Q^2)$, where $f_{I\!P}$ is the "Pomeron flux" describing the emission of the Pomeron from the proton and its subsequent propagation, and where $F_2^{I\!P}$ is the "structure function of the Pomeron". Phenomenologically, such a factorizing ansatz works not too badly and is often used, but recent high-precision data have shown its breakdown at small $x_{I\!P}$ [15].

3 Exclusive diffractive processes

Let us now discuss diffractive processes where a real or virtual photon dissociates into a single particle. Since diffraction involves the exchange of vacuum quantum numbers, this particle can in particular be a vector meson (which has the same J^{PC} quantum numbers as the photon) – in this case the process is sometimes referred to as "elastic" vector meson production. Another important case is deeply virtual Compton scattering (DVCS), $\gamma^* p \to \gamma p$.[3] A striking feature of the data taken at HERA (Figs. 15 and 16) is that the energy dependence of these processes becomes steep in the presence of a hard scale, which can be either the photon virtuality Q^2 or the mass of the meson in the case of J/Ψ or Υ production. This is similar to the energy dependence of the $\gamma^* p$ total cross section (related by the optical theorem to forward Compton scattering, $\gamma^* p \to \gamma^* p$), which changes from flat to steep when going from real photons to Q^2 of a few GeV^2.

To understand this similarity, let us recall that in perturbative QCD diffraction proceeds by two-gluon exchange. The transition from a virtual photon to a real photon or to a quark-antiquark pair subsequently hadronizing into a meson is a short-distance process involving these gluons, provided that either Q^2 or the quark mass is large. In fact, in an approximation discussed below, the cross sections for DVCS and vector meson production are proportional to the square of the gluon distribution in the proton, evaluated at a scale of order $Q^2 + M_V^2$ and at a momentum fraction $x_{I\!P} = (Q^2 + M_V^2)/(W^2 + Q^2)$, where the vector meson mass M_V now takes the role of M_X in inclusive diffraction [28]. In analogy to the case of the total $\gamma^* p$ cross section, the energy dependence of the cross sections shown in Figs. 15 and 16 thus reflects the x and scale dependence of the gluon density in the proton, which grows with decreasing x with a slope becoming steeper as the scale increases.

There is however an important difference in how the gluon distribution enters the descriptions of inclusive DIS and of exclusive diffractive processes. The inclusive DIS cross section is related via the optical theorem to the imaginary part of the forward virtual Compton amplitude, so that the graphs in Fig. 17 represent the *cross section* of the inclusive process. Hence, the gluon distribution in Fig. 17a gives the *probability* to find *one* gluon in the proton (with any number of unobserved spectator partons going into the final state). In contrast, the corresponding graphs for DVCS and exclusive meson production in Fig. 18 represent the *amplitudes* of exclusive processes, which are proportional to the *probability amplitude* for first extracting a gluon from the initial proton and then returning it to form the proton in the final state. In the approximation discussed below, this probability amplitude is given by the gluon distribution. The cross sections of DVCS and exclusive meson production are then proportional to the *square* of the gluon distribution.

A detailed theoretical analysis of DVCS and exclusive meson production at large Q^2 shows that short-distance factorization holds, in analogy to the case of inclusive DIS. QCD factorization theorems [29] state that in the limit of large Q^2 (at fixed Bjorken variable x_B and fixed t) the Compton amplitude factorizes into a hard-scattering subprocess and a hadronic matrix element describing the emission and reabsorption of a parton by the proton target (see Fig. 18a). As shown in Fig. 18b, the analogous result for exclusive meson production involves in addition the quark-antiquark distribution amplitude of

[3] We do not discuss processes with diffractive dissociation of the proton in this paper, but wish to mention interesting studies of vector meson or real photon production at large $|t|$, where the proton predominantly dissociates, see e.g. [26].

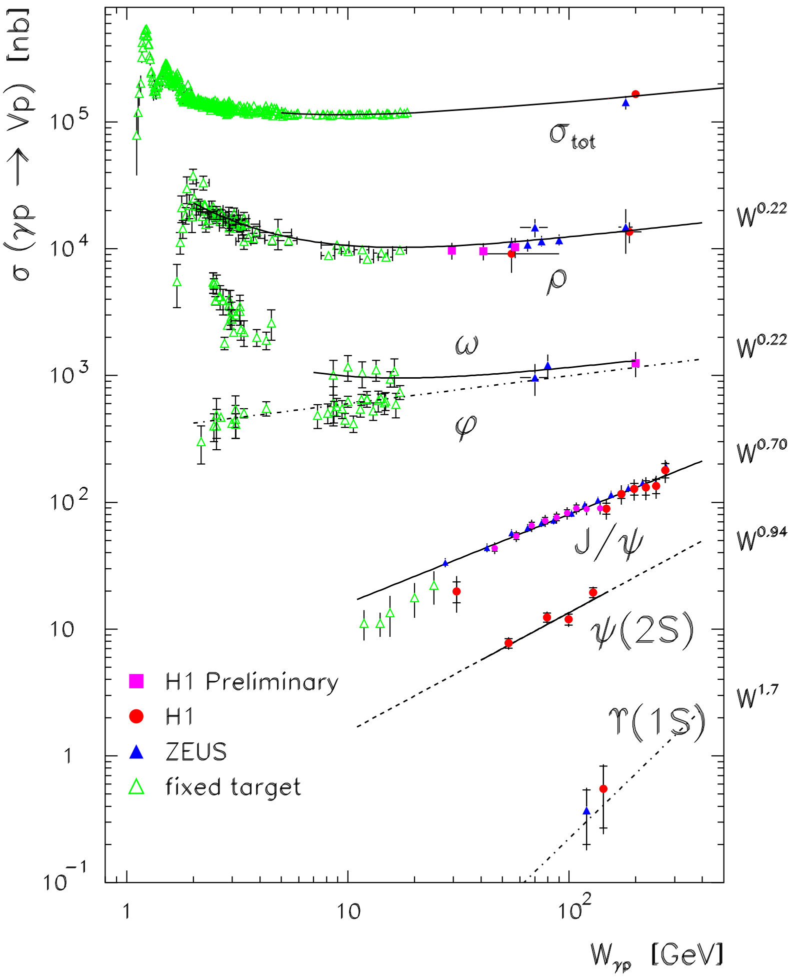


Fig. 15: Compilation of results on the cross section for vector meson photoproduction, $\gamma p \to Vp$, with $V = \rho, \omega, \phi, J/\Psi, \psi'$, Υ, as a function W. The total γp cross section $\sigma_{\rm tot}$ is also shown.

the meson (often termed the meson wave function) and thus a further piece of non-perturbative input.

The hadronic matrix elements appearing in the factorization formulae for exclusive processes would be the usual PDFs if the proton had the same momentum in the initial and final state. Since this is not the case, they are more general functions taking into account the momentum difference between the initial and final state proton (or, equivalently, between the emitted and reabsorbed parton). These "generalized parton distributions" (GPDs) depend on two independent longitudinal momentum fractions instead of a single one (compare Figs. 17a and 18a), on the transverse momentum transferred to the proton (whose square is $-t$ to a good approximation at high energy), and on the scale at which the partons are probed. The scale dependence of the GPDs is governed by a generalization of the DGLAP equations. The dependence on the difference of the longitudinal momenta (often called "skewness") contains information on correlations between parton momenta in the proton wave function. It can be neglected in the approximation of leading $\log x$ (then the GPDs at $t = 0$ reduce to the usual PDFs as anticipated above), but it is numerically important in typical HERA kinematics. The dependence on t allows for a very intuitive interpretation if a Fourier transformation is performed with respect to the transverse momentum transfer. We then obtain distributions depending on the impact parameter of the partons, which describe the two-dimensional distribution of the struck parton in the transverse plane, and on its longitudinal momentum fraction in the proton. The t dependence of exclusive diffractive processes thus provides unique information beyond the longitudinal momentum spectra encoded in the usual par-

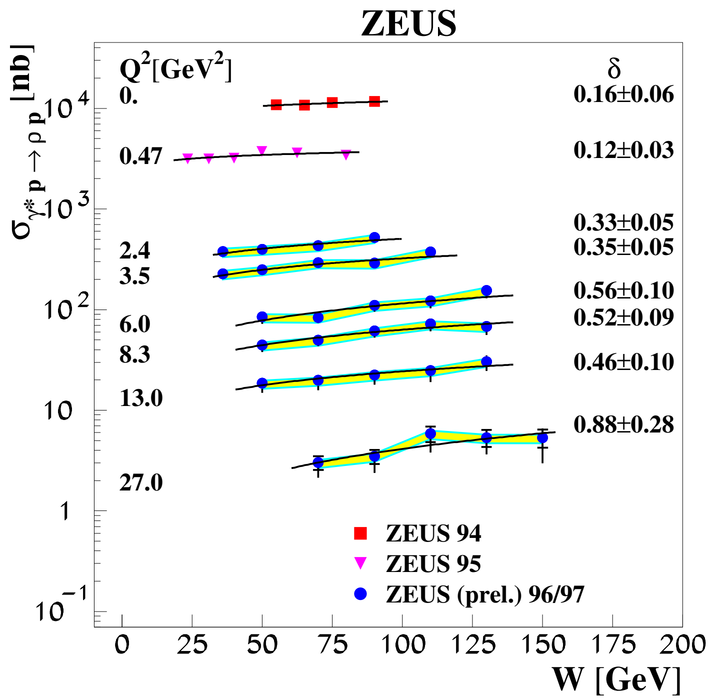


Fig. 16: Cross section for exclusive ρ production as a function of W (from [27]). The lines represent the result of fits to the data with the form $\sigma(\gamma^* p \to \rho p) \propto W^\delta$, yielding the exponents given in the figure.

ton densities. The study of the generalized parton distributions is a prime reason to measure DVCS and exclusive meson production in ep scattering. Detailed discussions and references can be found in the recent reviews [30, 31].

An observable illustrating the short-distance factorization in meson production at high Q^2 is the ratio of the ϕ and ρ production cross sections, shown in Fig. 19. At large Q^2 the process is described in terms of a light quark coupling to the photon and of the generalized gluon distribution. Using approximate flavor SU(3) symmetry between the ρ and ϕ wave functions, the only difference between the two channels is then due to different quark charge and isospin factors, which result in a cross section ratio of $2/9$.

3.1 High-energy factorization and the dipole picture

So far we have discussed the description of hard exclusive diffraction within short-distance, or collinear factorization. A different type of factorization is high-energy, or k_t factorization, which is based on the BFKL formalism. Here the usual or generalized gluon distribution appearing in the factorization formulae depends explicitly on the transverse momentum k_t of the emitted gluon. In collinear factorization, this k_t is integrated over in the parton distributions and set to zero when calculating the hard-scattering process (the partons are thus approximated as "collinear" with their parent hadron). Likewise, the meson wave functions appearing in k_t factorization explicitly depend on the relative transverse momentum between the quark and antiquark in the meson, whereas this is integrated over in the quark-antiquark distribution amplitudes (cf. Sect. 3) of the collinear factorization formalism. Only gluon distributions appear in k_t factorization, whereas collinear factorization formulae involve both quark and gluon distributions (see e.g. Sects. 8.1 and 8.2 in [30] for a discussion of this difference). We note that other factorization schemes have been developed, which combine features of the collinear and k_t factorization formalisms.

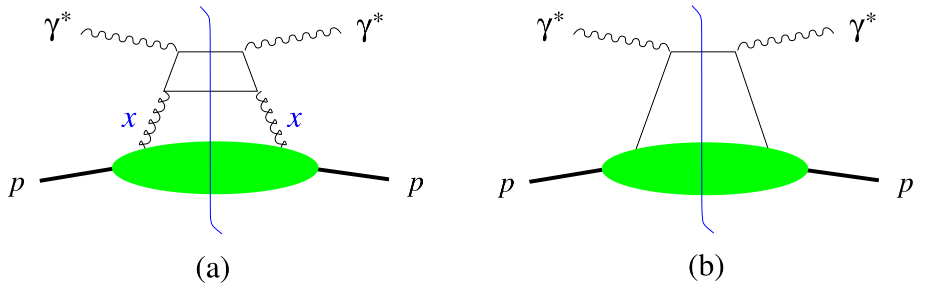


Fig. 17: Factorization of forward Compton scattering, which is related to the total inclusive structure function via the optical theorem, $\mathrm{Im}\,\mathcal{A}(\gamma^* p \to \gamma^* p) = \frac{1}{2}\sum_X |\mathcal{A}(\gamma^* p \to X)|^2 \propto \sigma(\gamma^* p \to X)$. The final state of the inclusive process is obtained by cutting the diagrams along the vertical line. The blobs represent the gluon or quark distribution in the proton. Graph (b) is absent in the k_t factorization formalism (see Sect. 3.1): its role is taken by graph (a) in the "aligned jet configuration", where the quark line joining the two photons carries almost the entire photon momentum.

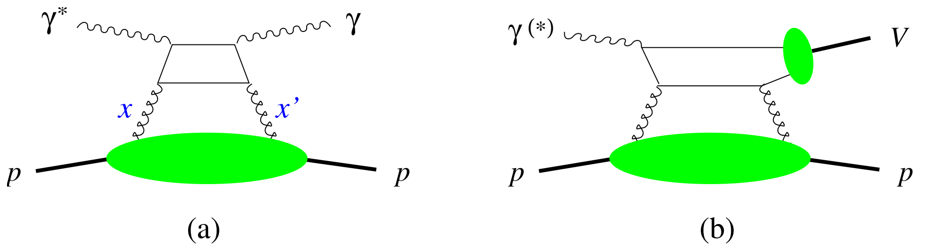


Fig. 18: (a) Factorization of deeply virtual Compton scattering, $\gamma^* p \to \gamma p$, which can be measured in the exclusive process $ep \to ep\gamma$. The blob represents the generalized gluon distribution, with x and x' denoting the momentum fractions of the gluons. (b) Factorization of exclusive meson production. The small blob represents the vector meson wave function. In the collinear factorization formalism, there are further graphs (not shown) involving quark instead of gluon exchange.

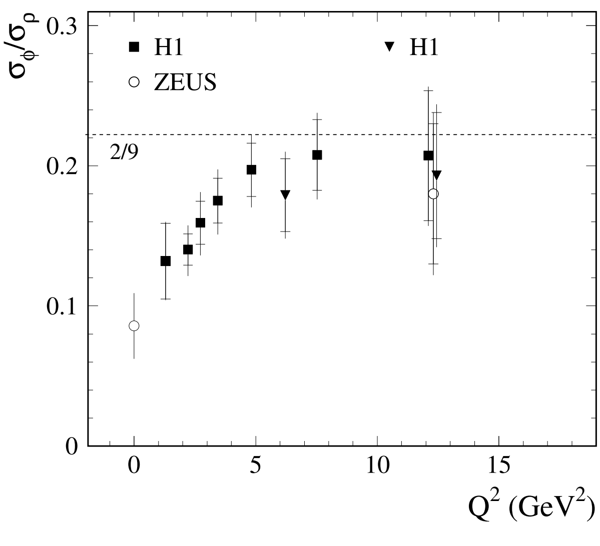


Fig. 19: The ratio of cross sections for $\gamma^* p \to \phi p$ and $\gamma^* p \to \rho p$ as a function of the photon virtuality (from [32]).

The two different types of factorization implement different ways of separating different parts of the dynamics in a scattering process. The building blocks in a short-distance factorization formula correspond to either small or large particle virtuality (or equivalently to small or large transverse momentum), whereas the separation criterion in high-energy factorization is the particle rapidity. Collinear and k_t factorization are based on taking different limits: in the former case the limit of large Q^2 at fixed x_B and in the latter case the limit of small x_B at fixed Q^2 (which must however be large enough to justify the use of QCD perturbation theory). In the common limit of large Q^2 and small x_B the two schemes give coinciding results. Instead of large Q^2 one can also take a large quark mass in the limits just discussed.

A far-reaching representation of high-energy dynamics can be obtained by casting the results of k_t factorization into a particular form. The different building blocks in the graphs for Compton scattering and meson production in Figs. 17a and 18 can be rearranged as shown in Fig. 20. The result admits a very intuitive interpretation in a reference frame where the photon carries large momentum (this may be the proton rest frame but also a frame where the proton moves fast, see Fig. 14): the initial photon splits into a quark-antiquark pair, which scatters on the proton and finally forms a photon or meson again. This is the picture we have already appealed to in Sect. 1.2.

In addition, one can perform a Fourier transformation and trade the relative transverse momentum between quark and antiquark for their transverse distance r, which is conserved in the scattering on the target. The quark-antiquark pair acts as a color dipole, and its scattering on the proton is described by a "dipole cross section" $\sigma_{q\bar{q}}$ depending on r and on $x_{I\!P}$ (or on x_B in the case of inclusive DIS). The wave functions of the photon and the meson depend on r after Fourier transformation, and at small r the photon wave function is perturbatively calculable. Typical values of r in a scattering process are determined by the inverse of the hard momentum scale, i.e. $r \sim (Q^2 + M_V^2)^{-1/2}$. An important result of high-energy factorization is the relation

$$\sigma_{q\bar{q}}(r, x) \propto r^2 x g(x) \tag{7}$$

at small r, where we have replaced the generalized gluon distribution by the usual one in the spirit of the leading $\log x$ approximation. A more precise version of the relation (7) involves the k_t dependent gluon distribution. The dipole cross section vanishes at $r = 0$ in accordance with the phenomenon of "color transparency": a hadron becomes more and more transparent for a color dipole of decreasing size.

The scope of the dipole picture is wider than we have presented so far. It is tempting to apply it outside the region where it can be derived in perturbation theory, by modeling the dipole cross section and the photon wave function at large distance r. This has been very been fruitful in phenomenology, as we will see in the next section.

The dipole picture is well suited to understand the t dependence of exclusive processes, parameterized as $d\sigma/dt \propto \exp(-b|t|)$ at small t. Figure 21 shows that b decreases with increasing scale $Q^2 + M_V^2$ and at high scales becomes independent of the produced meson. A Fourier transform from momentum to impact parameter space readily shows that b is related to the typical transverse distance between the

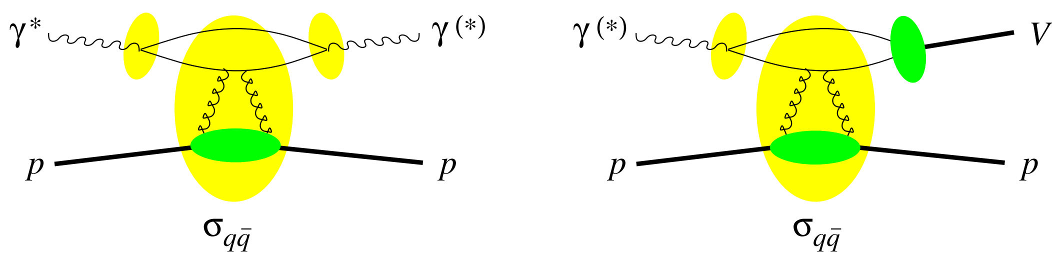


Fig. 20: The dipole representation of the amplitudes for Compton scattering (a) and for meson production (b), corresponding to the graphs in Figs. 17a and 18.

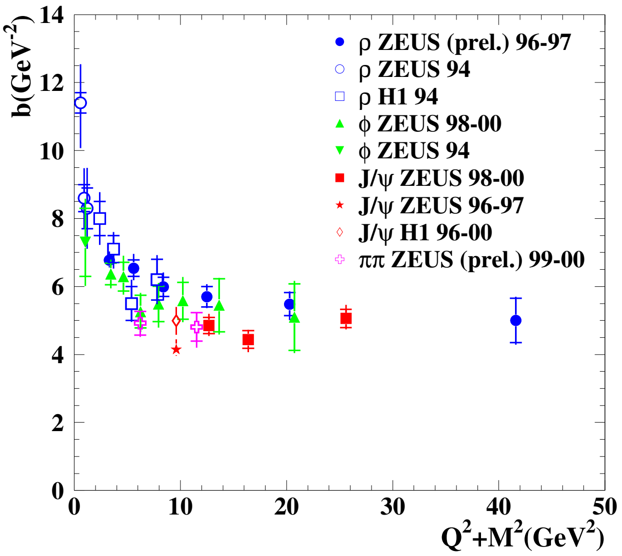


Fig. 21: The logarithmic slope of the t dependence at $t = 0$ for different meson production channels, as well as for non-resonant dipion production.

colliding objects, as anticipated by the analogy with optical diffraction in Sect. 1.1. At high scale, the $q\bar{q}$ dipole is almost pointlike, and the t dependence of the cross section is controlled by the t dependence of the generalized gluon distribution, or in physical terms, by the transverse extension of the proton. As the scale decreases, the dipole acquires a size of its own, and in the case of ρ or ϕ photoproduction, the values of b reflect the fact that the two colliding objects are of typical hadronic dimensions; similar values would be obtained in elastic meson-proton scattering.

3.2 Exclusive diffraction in hadron-hadron collisions

The concepts we have introduced to describe exclusive diffraction can be taken over to pp or $p\bar{p}$ scattering, although further complications appear in these processes. A most notable reaction is exclusive production of a Higgs boson, $pp \to pHp$, sketched in Fig. 22. The generalized gluon distribution is a central input in this description. The physics interest, theory description, and prospects to measure this process at the LHC have been discussed in detail at this workshop [33, 34]. A major challenge in the description of this process is to account for secondary interactions between spectator partons of the two projectiles, which can produce extra particles in the final state and hence destroy the rapidity gaps between the Higgs and final-state protons – the very same mechanism we discussed in Sect. 2.4.

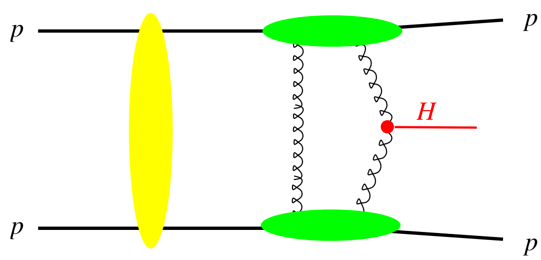


Fig. 22: Graph for the exclusive production of a Higgs boson in pp scattering. The horizontal blobs indicate generalized gluon distributions, and the vertical blob represents secondary interactions between the projectiles (cf. Fig. 13).

4 Parton saturation

We have seen that diffraction involves scattering on small-x gluons in the proton. Consider the density in the transverse plane of gluons with longitudinal momentum fraction x that are resolved in a process with hard scale Q^2. One can think of $1/Q$ as the "transverse size" of these gluons as seen by the probe. The number density of gluons at given x increases with increasing Q^2, as described by DGLAP evolution (see Fig. 23). According to the BFKL evolution equation it also increases at given Q^2 when x becomes smaller, so that the gluons become more and more densely packed. At some point, they will start to overlap and thus reinteract and screen each other. One then enters a regime where the density of partons saturates and where the linear DGLAP and BFKL evolution equations cease to be valid. If Q^2 is large enough to have a small coupling α_s, we have a theory of this non-linear regime called "color glass condensate", see e.g. [35]. To quantify the onset of non-linear effects, one introduces a saturation scale Q_s^2 depending on x, such that for $Q^2 < Q_s^2(x)$ these effects become important. For smaller values of x, the parton density in the target proton is higher, and saturation sets in at larger values of Q^2 as illustrated in Fig. 23.

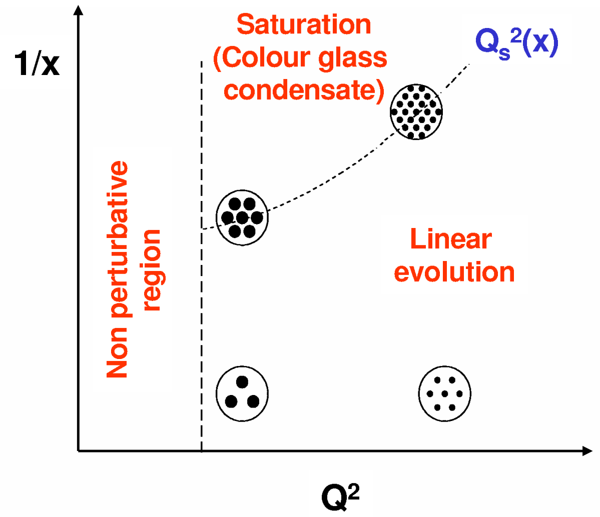


Fig. 23: Schematic view of the density of gluons in the transverse plane, as a function of the momentum fraction x and the resolution scale Q^2. Above the line given by $Q_s^2(x)$, saturation effects set in.

The dipole picture we introduced in Sect. 3.1 is well suited for the theoretical description of saturation effects. When such effects are important, the relation (7) between dipole cross section and gluon distribution ceases to be valid; in fact the gluon distribution itself is then no longer an adequate quantity to describe the dynamics of a scattering process. In a certain approximation, the evolution of the dipole cross section with x is described by the Balitsky-Kovchegov equation [36], which supplements the BFKL equation with a non-linear term taming the growth of the dipole cross section with decreasing x.

Essential features of the saturation phenomenon are captured in a phenomenological model for the dipole cross section, originally proposed by Golec-Biernat and Wüsthoff, see [37, 38]. Figure 24 shows $\sigma_{q\bar{q}}$ as a function of r at given x in this model. The dipole size r now plays the role of $1/Q$ in our discussion above. At small r the cross section rises following the relation $\sigma_{q\bar{q}}(r,x) \propto r^2 xg(x)$. At some value $R_s(x)$ of r, the dipole cross section is so large that this relation ceases to be valid, and $\sigma_{q\bar{q}}$ starts to deviate from the quadratic behavior in r. As r continues to increase, $\sigma_{q\bar{q}}$ eventually saturates at a value typical of a meson-proton cross section. In terms of the saturation scale introduced above, $R_s(x) = 1/Q_s(x)$. For smaller values of x, the initial growth of $\sigma_{q\bar{q}}$ with r is stronger because the gluon distribution is larger. The target is thus more opaque and as a consequence saturation sets in at lower r.

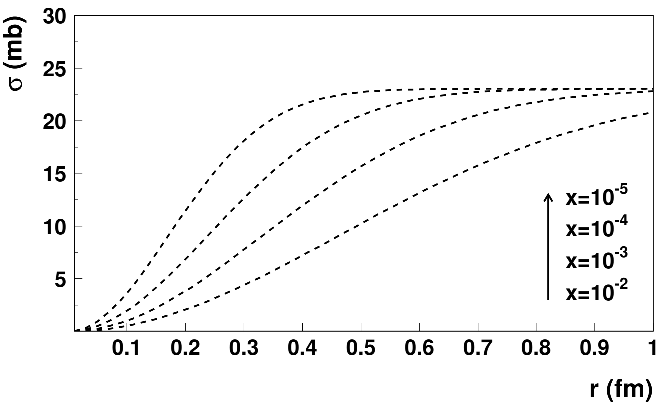


Fig. 24: The dipole cross section $\sigma_{q\bar{q}}$ in the Golec-Biernat Wüsthoff model as a function of dipole size r for different x (from [38]).

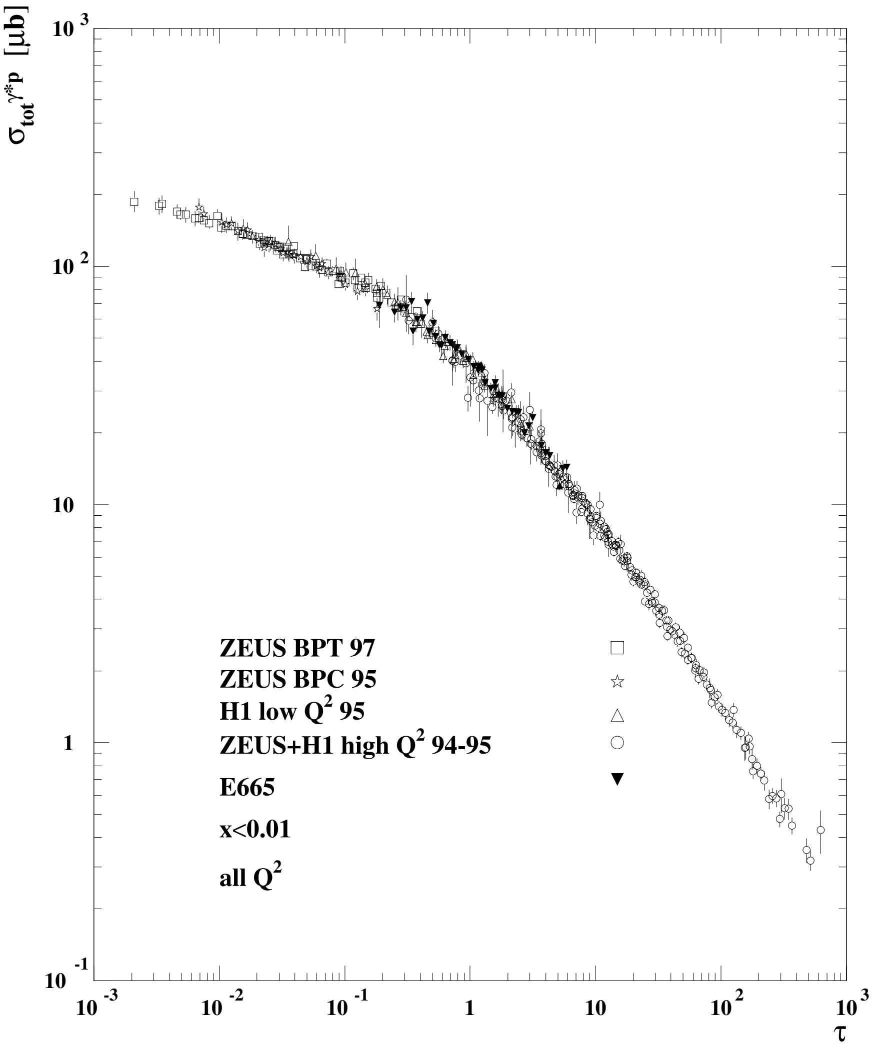


Fig. 25: Geometric scaling of the $\gamma^* p$ cross section in a single variable $\tau = Q^2/Q_s^2(x_B)$, as determined in [39]. The Q^2 of these data ranges from 0.045 to 450 GeV^2.

A striking feature found both in this phenomenological model [39] and in the solutions of the Balitsky-Kovchegov equation (see e.g. [40]) is that the total $\gamma^* p$ cross section only depends on Q^2 and x_B through a single variable $\tau = Q^2/Q_s^2(x_B)$. This property, referred to as geometric scaling, is well satisfied by the data at small x_B (see Fig. 25) and is an important piece of evidence that saturation effects are visible in these data. Phenomenological estimates find Q_s^2 of the order 1 GeV^2 for x_B around 10^{-3} to 10^{-4}.

The dipole formulation is suitable to describe not only exclusive processes and inclusive DIS, but also inclusive diffraction $\gamma^* p \to Xp$. For a diffractive final state $X = q\bar{q}$ at parton level, the theory description is very similar to the one for deeply virtual Compton scattering, with the wave function for the final state photon replaced by plane waves for the produced $q\bar{q}$ pair. The inclusion of the case $X = q\bar{q}g$ requires further approximations [37] but is phenomenologically indispensable for moderate to small β. Experimentally, one observes a very similar energy dependence of the inclusive diffractive and the total cross section in $\gamma^* p$ collisions at given Q^2 (see Fig. 26). The saturation mechanism implemented in the Golec-Biernat Wüsthoff model provides a simple explanation of this finding. To explain this aspect of the data is non-trivial. For instance, in the description based on collinear factorization, the energy dependence of the inclusive and diffractive cross sections is controlled by the x dependence of the ordinary and the diffractive parton densities. This x dependence is not predicted by the theory.

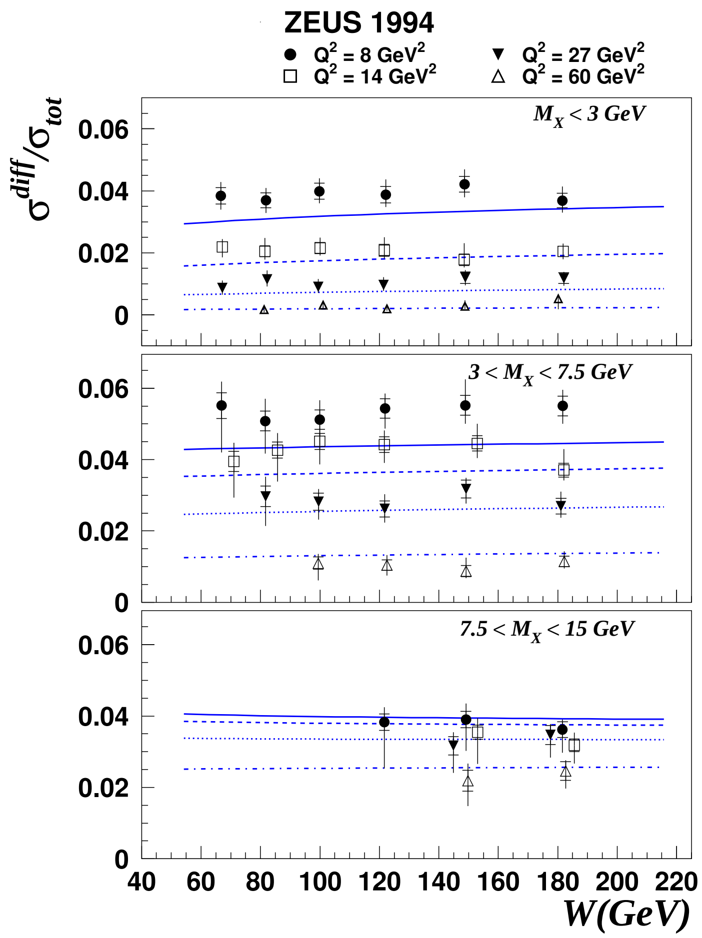


Fig. 26: Data on the ratio of diffractive and total $\gamma^* p$ cross sections compared with the result of the Golec-Biernat Wüsthoff model (from [37]).

The description of saturation effects in pp, pA and AA collisions requires the full theory of the color glass condensate, which contains concepts going beyond the dipole formulation discussed here and is e.g. presented in [35]. We remark however that estimates of the saturation scale $Q_s^2(x)$ from HERA data can be used to describe features of the recent data from RHIC [41].

5 A short summary

Many aspects of diffraction in ep collisions can be successfully described in QCD if a hard scale is present. A key to this success are factorization theorems, which render parts of the dynamics accessible to calculation in perturbation theory. The remaining non-perturbative quantities, namely diffractive PDFs and generalized parton distributions, can be extracted from measurements and contain specific information about small-x partons in the proton that can only be obtained in diffractive processes. To describe hard diffractive hadron-hadron collisions is more challenging since factorization is broken by rescattering between spectator partons. These rescattering effects are of interest in their own right because of their intimate relation with multiple scattering effects, which at LHC energies are expected to be crucial for understanding the structure of events in hard collisions. A combination of data on inclusive and diffractive ep scattering hints at the onset of parton saturation at HERA, and the phenomenology developed there is a helpful step towards understanding high-density effects in hadron-hadron collisions.

Acknowledgments

It is a pleasure to thank our co-convenors and all participants for the fruitful atmosphere in the working group on diffraction, and A. De Roeck and H. Jung for their efforts in organizing this workshop. We are indebted to A. Proskuryakov for Figs. 5 and 12, to P. Fleischmann for Fig. 15, and to A. Bonato, K. Borras, A. Bruni, J. Forshaw, M. Grothe, H. Jung, L. Motyka, M. Ruspa and M. Wing for valuable comments on the manuscript.

References

[1] P. D. B. Collins, *An Introduction to Regge Theory and High-Energy Physics*, Cambridge University Press, Cambridge, 1977.

[2] E. A. Kuraev, L. N. Lipatov and V. S. Fadin, Sov. Phys. JETP **44** (1976) 443; ibid. **45** (1977) 199; I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28** (1978) 822.

[3] J. R. Forshaw and D. A. Ross, *Quantum chromodynamics and the pomeron*, Cambridge Lect. Notes Phys. **9** (1997) 1;
A. Hebecker, Phys. Rept. **331** (2000) 1 [hep-ph/9905226];
M. Wüsthoff and A. D. Martin, J. Phys. G **25** (1999) R309 [hep-ph/9909362];
V. Barone and E. Predazzi, *High-energy particle diffraction*, Springer, 2002;
I. P. Ivanov, N. N. Nikolaev and A. A. Savin, hep-ph/0501034.

[4] K. Goulianos, Phys. Rept. **101** (1983) 169.

[5] G. Bruni *et al.*, these proceedings.

[6] S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **38** (2004) 43 [hep-ex/0408009].

[7] H1 Collaboration, Paper 89 submitted to the International Europhysics Conference on High Energy Physics (EPS 2003), July 2003, Aachen, Germany, available from http://www-h1.desy.de

[8] ZEUS Collaboration, Paper 1048 submitted to the 30th International Conference on High Energy Physics (ICHEP 2000), July–August 2000, Osaka, Japan, available from http://www-zeus.desy.de

[9] S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **21** (2001) 443 [hep-ex/0105090].

[10] L. Trentadue and G. Veneziano, Phys. Lett. B **323** (1994) 201.

[11] A. Berera and D. E. Soper, Phys. Rev. D **53** (1996) 6162 [hep-ph/9509239].

[12] J. C. Collins, Phys. Rev. D **57** (1998) 3051, Erratum ibid. D **61** (2000) 019902 [hep-ph/9709499].

[13] P. Newman and F.-P. Schilling, these proceedings.

[14] M. Groys, A. Levy and A. Proskuryakov, these proceedings.

[15] S. Chekanov *et al.* [ZEUS Collaboration], Nucl. Phys. B **713** (2005) 3 [hep-ex/0501060].

[16] A. D. Martin, M. G. Ryskin and G. Watt, Eur. Phys. J. C **37** (2004) 285 [hep-ph/0406224]; A. D. Martin, M. G. Ryskin and G. Watt, these proceedings.

[17] H1 Collaboration, Paper 6-0177 submitted to the 32nd International Conference on High Energy Physics (ICHEP 2004), August 2004, Beijing, China, available from http://www-h1.desy.de

[18] A. Bruni *et al.*, these proceedings [hep-ph/0509202].

[19] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **84** (2000) 5043.

[20] L. Alvero *et al.*, Phys. Rev. D **59** (1999) 074022 [hep-ph/9805268].

[21] A. Proskuryakov, Y. Yamazaki and M. Arneodo, presented at the HERA-LHC Workshop, 11–13 October 2004, CERN, http://agenda.cern.ch/fullAgenda.php?ida=a044163

[22] E. Gotsman *et al.*, these proceedings.

[23] J. Bartels, M. Salvadore and G. P. Vacca, Eur. Phys. J. C **42** (2005) 53 [hep-ph/0503049];
J. Bartels, these proceedings,
C. Buttar *et al.*, these proceedings.

[24] G. Ingelman and P. E. Schlein, Phys. Lett. B **152** (1985) 256.

[25] J. Bartels and M. G. Ryskin, Z. Phys. C **76** (1997) 241 [hep-ph/9612226].

[26] R. Enberg *et al.*, hep-ph/0407328;
S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **26** (2003) 389 [hep-ex/0205081];
A. Aktas *et al.* [H1 Collaboration], Phys. Lett. B **568** (2003) 205 [hep-ex/0306013].

[27] ZEUS Collaboration, Paper 594 submitted to the International Europhysics Conference on High Energy Physics (EPS 2001), July 2001, Budapest, Hungary, available from http://www-zeus.desy.de

[28] M. G. Ryskin, Z. Phys. C **57** (1993) 89.

[29] J. C. Collins, L. Frankfurt and M. Strikman, Phys. Rev. D **56** (1997) 2982 [hep-ph/9611433];
J. C. Collins and A. Freund, Phys. Rev. D **59** (1999) 074009 [hep-ph/9801262].

[30] M. Diehl, Phys. Rept. **388**, 41 (2003) [hep-ph/0307382].

[31] A. V. Belitsky and A. V. Radyushkin, hep-ph/0504030.

[32] C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **483** (2000) 360 [hep-ex/0005010].

[33] Contributions to these proceedings by J. Kalliopuska *et al.*; V. Avati and K. Österberg; M. Arneodo *et al.*; V. Andreev *et al.*

[34] J. Forshaw, these proceedings [hep-ph/0508274];
M. Boonekamp *et al.*, these proceedings.

[35] H. Weigert, hep-ph/0501087;
R. Venugopalan, these proceedings.

[36] I. Balitsky, Nucl. Phys. B **463** (1996) 99 [hep-ph/9509348];
Y. V. Kovchegov, Phys. Rev. D **61** (2000) 074018 [hep-ph/9905214].

[37] K. Golec-Biernat and M. Wüsthoff, Phys. Rev. D **60** (1999) 114023 [hep-ph/9903358].

[38] K. Golec-Biernat, Acta Phys. Polon. B **33** (2002) 2771 [hep-ph/0207188].

[39] A. M. Stasto, K. Golec-Biernat and J. Kwiecinski, Phys. Rev. Lett. **86** (2001) 596 [hep-ph/0007192].

[40] K. Golec-Biernat, L. Motyka and A. M. Stasto, Phys. Rev. D **65** (2002) 074037 [hep-ph/0110325];
M. Lublinsky, Eur. Phys. J. C **21** (2001) 513 [hep-ph/0106112];
S. Munier and R. Peschanski, Phys. Rev. Lett. **91** (2003) 232001 [hep-ph/0309177].

[41] D. Kharzeev and E. Levin, Phys. Lett. B **523** (2001) 79 [nucl-th/0108006];
A. Krasnitz, Y. Nara and R. Venugopalan, Nucl. Phys. A **727** (2003) 427 [hep-ph/0305112].

Diffractive Higgs Production: Experiment

Editors: B. E. Cox [a] and M. Grothe [b]
[a] School of Physics and Astronomy, The University of Manchester, Manchester M139PL, UK
[b] University of Torino and INFN-Torino, Italy; also at University of Wisconsin, Madison, WI, USA

TOTEM forward measurements: exclusive central diffraction

J. Kalliopuska, J.W. Lämsä, T. Mäki, N. Marola, R. Orava, K. Österberg, M. Ottela and S. Tapprogge

Abstract

In this contribution, we present a first systematic study of the precision of the momentum measurement of protons produced in the central exclusive diffractive processes, $pp \rightarrow p + X + p$, as well as the accuracy of the reconstructed mass for particle state X based on these proton measurements. The scattered protons are traced along the LHC beam line using the nominal LHC optics, accounting for uncertainties related to beam transport and proton detection.

To search for and precisely measure new particle states X with masses below 200 GeV, additional leading proton detectors are required at about 420 m from the interaction point in addition to the already approved detectors. Using these additional detectors, a mass resolution of the order of 1 GeV can be achieved for masses beyond $\sim$120 GeV.

TOTEM forward measurements: leading proton acceptance

V. Avati and K. Österberg

Abstract

We report about the acceptance of forward leading protons in Roman Pot stations placed along the LHC beam line. The TOTEM stations plus additional detectors at 420 m from the interaction point have been considered using the low–β^* optics V6.5 for LHC physics runs.

Diffractive Higgs: CMS/TOTEM Level-1 Trigger Studies

M. Arneodo, V. Avati, R. Croft, F. Ferro, M. Grothe, C. Hogg, F. Oljemark, K. Österberg and M. Ruspa

Abstract

Retaining events containing a Higgs Boson with mass around 120 GeV poses a special challenge to triggering at the LHC due to the relatively low transverse momenta of the decay products. We discuss the potential of including into the CMS trigger the TOTEM forward detectors and possible additional detectors at a distance of 420 m from the CMS interaction point. We find that the output rate of a 2-jet Level-1 trigger condition with thresholds sufficiently low for the decay products of a 120 GeV Higgs Boson can be limited to $\mathcal{O}(1)$ kHz for luminosities of up to $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ by including the TOTEM forward detectors in the Level-1 trigger.

Proposal to upgrade the very forward region at CMS

V. Andreev, A.Bunyatyan, H. Jung, M. Kapishin and L. Lytkin

Abstract

The possibilities of extending the acceptance of LHC experiments beyond 7 units of pseudorapidity are investigated. With additional detectors it would be possible to measure the particles with energies above 2 TeV in the pseudorapidity range between 7 and 11.

TOTEM forward measurements: exclusive central diffraction

J. Kalliopuska, J.W. Lämsä, T. Mäki, N. Marola, R. Orava, K. Österberg,† M. Ottela and S. Tapprogge‡*
High Energy Physics Division, Department of Physical Sciences, University of Helsinki and
Helsinki Institute of Physics, P.O. Box 64, FIN-00014 University of Helsinki, Finland

Abstract

In this contribution, we present a first systematic study of the precision of the momentum measurement of protons produced in the central exclusive diffractive processes, $p\,p \;\rightarrow\; p + X + p$, as well as the accuracy of the reconstructed mass for particle state X based on these proton measurements. The scattered protons are traced along the LHC beam line using the nominal LHC optics, accounting for uncertainties related to beam transport and proton detection.

To search for and precisely measure new particle states X with masses below 200 GeV, additional leading proton detectors are required at about 420 m from the interaction point in addition to the already approved detectors. Using these additional detectors, a mass resolution of the order of 1 GeV can be achieved for masses beyond $\sim$120 GeV.

1 Introduction

It has been recently suggested that the Higgs boson mass could be measured to an accuracy of $\mathcal{O}$(1 GeV) in the central exclusive diffractive process (CED) [1, 2]:

$$p\,p \rightarrow p\,+\,H\,+\,p \tag{1}$$

In contrast to this, the direct measurement of the Higgs boson mass, based on the two final state b-jets in $H \rightarrow b\bar{b}$, is estimated to yield a precision of $\mathcal{O}$(10 GeV). The precise reconstruction of the centrally produced system X, i.e. the Higgs mass in Eq. 1, is based on the four-momenta of the incoming ($p_{1,2}$) and scattered ($p'_{1,2}$) protons and since the two scattered protons are expected to have small transverse momenta, the following approximation for the mass of the centrally produced system can be made:

$$M^2 = (p_1 + p_2 - p'_1 - p'_2)^2 \approx \xi_1 \xi_2 s\,, \tag{2}$$

where $\xi_{1,2} = 1 - |\vec{p}'_{1,2}|/|\vec{p}_{1,2}|$ denote the momentum loss fractions of the two scattered protons.

The acceptance for forward leading protons for nominal LHC runs ($\beta^* \sim$ 0.5 m) is described in detail elsewhere (see [3]). This contribution focuses on the CED process and the precision with which the proton momenta and the mass of the centrally produced system can be reconstructed.

2 Leading proton uncertainties and transport

The study is done in multiple steps, which include the event generation (ExHuME [4] or PHOJET [5]), simulation of the interaction point (IP) region, tracking of the protons through the LHC beam line, a detector simulation and a proton momentum reconstruction algorithm using the detector information [6]. The following beam related uncertainties are inputs to the study[1]:

- pp interaction region width: $\sigma_{x,y}$ = 16 μm, σ_z = 5 cm,

* also Physics Department, Iowa State University, Ames, USA

† corresponding author: kenneth.osterberg@helsinki.fi

‡ currently at Institute of Physics, Johannes-Gutenberg Universität Mainz, Germany

[1] The reference system (x, y, z) used in the study corresponds to the reference orbit in the accelerator; the z-axis is tangent to the orbit and positive in the beam direction; the x-axis (horizontal) is negative toward the center of the ring.

- beam angular divergence: $\Theta_{x,y}$ = 30 μrad
- beam energy spread: $1.1 \cdot 10^{-4}$.

Concerning the detector response, only the horizontal plane is considered with the following inputs:

- The detector is assumed to be fully efficient at a distance $10\sigma_x(z)$ + 0.5 mm from the beam center[2], where $\sigma_x(z)$ is the horizontal beam width at distance z. The second term takes into account the distance from the bottom of the vacuum window to the edge of the fully sensitive detector area.
- For the protons within the fully sensitive detector area, a position reconstruction uncertainty is introduced by smearing the hit coordinates according to a Gaussian distribution with a σ of 10 μm.
- The uncertainty due to the beam position knowledge at each detector location is accounted for by smearing the hit coordinates by a correlated Gaussian distribution with a σ of 5 μm.

The transverse displacement $(x(z), y(z))$ of a scattered proton at a distance z from the IP is determined by tracing the proton along the LHC beam line using the MAD program [7]. The optics layout version 6.2 for nominal LHC runs (β^* = 0.5 m) with a 150 μrad horizontal crossing angle is used [8]. Although the study was carried out for CMS/TOTEM (IP5), the results should be equally valid for ATLAS (IP1).

3 Proton momentum reconstruction

The x-coordinate of the proton observed at any given location along the beam line, depends on three initial parameters of the scattered proton: its fractional momentum loss, ξ, its initial horizontal scattering angle, Θ^*_x, and its horizontal position of origin, x^*, at the IP. Consequently, more than one x-measurement of a particular proton is needed to constrain its parameters. In the procedure chosen, two x-measurements from a detector doublet are used to determine ξ and Θ^*_x, neglecting the x^* dependence. The effect of the x^* on the reconstructed proton momentum will be treated as an independent source of uncertainty.

To obtain a large acceptance in ξ, the following two detector locations, each consisting of a doublet of proton detectors, are chosen based on the LHC optics layout:

- 215 and 225 meters from IP5 ("215 m location"), and
- 420 and 430 meters from IP5 ("420 m location").

The 215 m location corresponds to a TOTEM approved Roman Pot location [9], while the 420 m location in the cryogenic section of the accelerator will require special design and further investigation.

Each detector doublet yields two observables, which are related to the horizontal offset and angle with respect to the beam axis. The ξ dependence of these observables has been derived by fitting a functional form to the simulated average values of ξ, as a function of the values of the two observables [6].

4 Acceptance and resolution on ξ and mass

The ξ and t acceptance of protons moving in the clockwise and counter-clockwise directions are slightly different due to differences in the optical functions, for details see [3]. As a summary: a proton from the CED process is seen when its ξ is between 0.025 (0.002) and 0.20 (0.015) for the 215 (420) location.

The relative resolution on ξ, $\Delta\xi/\xi = (\xi - \xi_{rec})/\xi$ as a function of ξ for protons produced in the CED process and seen in either the 215 m or the 420 m location is shown in Fig. 1 for protons circulating in the LHC both in the clockwise and counter-clockwise direction. Included are the separate effects from the uncertainty of the transverse IP position, the resolution of the proton detector, the beam energy uncertainty, the beam angular divergence at the IP, and the beam position resolution at the proton detector.

At both detector locations, major contributors to the over-all ξ resolution are the uncertainty of the transverse IP position and the resolution of the proton detector. In addition to these two uncertainties, the beam energy uncertainty contributes significantly to the resolution at the 420 m location.

[2]The LHC collimators extend to $6\sigma_x(z)$. The closest safe position can be assumed to lay anywhere between 10 and 15.

The acceptance as function of the mass of the centrally produced system is shown in Fig. 2a. Each leading proton is required to be within the acceptance of either the 215 or 420 m locations. Independently shown is the case (sub-set of above) where both protons are within the acceptance of the 420 m locations. In the mass range shown, there is no acceptance for detecting both protons at the 215 m location. The ξ_1-ξ_2 combinations result from the gluon density function in the proton and the mass of the centrally produced system (see Eq. 2). The ExHuME generator favours a harder gluon distribution than that of PHOJET. Thus, the Higgses are produced more centrally. This yields a higher acceptance for ExHuME.

The resolution effects of the two scattered protons are, in general, uncorrelated from each other. The only correlation comes from the production point, whose transverse component is determined by the rms spread of the beam at the IP and by an independent measurement using the Higgs decay products [10]. It can be determined to 10 μm or better, and therefore for the mass resolution of the centrally produced system, a 10 μm uncertainty on the transverse IP position is used. For the mass resolution, all other uncertainties are assumed to be uncorrelated between the two protons.

The mass resolutions for events with protons within the acceptance of the 420 m location on both sides, and for events with one proton within the acceptance of the 215 m location on one side and the other proton within the acceptance of the 420 m location on the other side (labelled "asym." in the figure) are shown as a function of the mass of the centrally produced system in Fig. 2b. The values quoted in the figure are based on Gaussian fits to the reconstructed mass distributions. The two-proton acceptance requirement imposes a restriction on the allowed ξ_1-ξ_2 combinations; as a result the mass resolutions obtained with ExHuME and PHOJET are very similar.

5 Conclusions

The first comprehensive study of the CED process at the LHC is reported. The study is based on detailed simulations along the LHC beam line of the diffractively scattered protons, accounting for the known sources of uncertainties related to beam transport and proton detection. The feasibility of measuring such events during nominal LHC runs for masses of the central system, X, below $\sim$200 GeV is addressed.

On the basis of this study, it is concluded that with an additional pair of leading proton detectors at $\pm$420 m from the interaction point, a Higgs boson with a mass of 120–180 GeV could be measured with a mass resolution of the order of 1 GeV. Such additional proton detectors would also enable large statistics of pure gluon jets to be collected, thereby turning the LHC into a gluon factory.

References

[1] A. De Roeck, V.A. Khoze, A.D. Martin, R. Orava and M.G. Ryskin, Eur. Phys. J. **C35**, 391 (2002).

[2] J. Kalliopuska, T. Mäki, N. Marola, R. Orava, K. Österberg, M. Ottela and S. Tapprogge, **HIP-2003-11/EXP** (2003).

[3] V. Avati and K. Österberg, "TOTEM forward measurements: leading proton acceptance", these proceedings.

[4] J. Monk and A. Pilkington, **hep-ph/0502077** (2005).

[5] R. Engel, Phys. Rev. **D51**, 3220 (1995).

[6] T. Mäki, Master thesis: "Exclusive production of Higgs boson at LHC collider: Higgs mass measurement via leading proton detection", Helsinki University of Technology (2003).

[7] The MAD program, Methodical Accelerator Design, **http://www.cern.ch/mad**.

[8] Lattice & Optics, **http://lhc-new-homepage.web.cern.ch/lhc-new-homepage/**.

[9] The TOTEM collaboration, TOTEM Technical Design Report, **CERN-LHCC-2004-002** (2004).

[10] P. Vanlaer, Hadron Collider Physics Conference, Les Diablerets, Switzerland (2005).

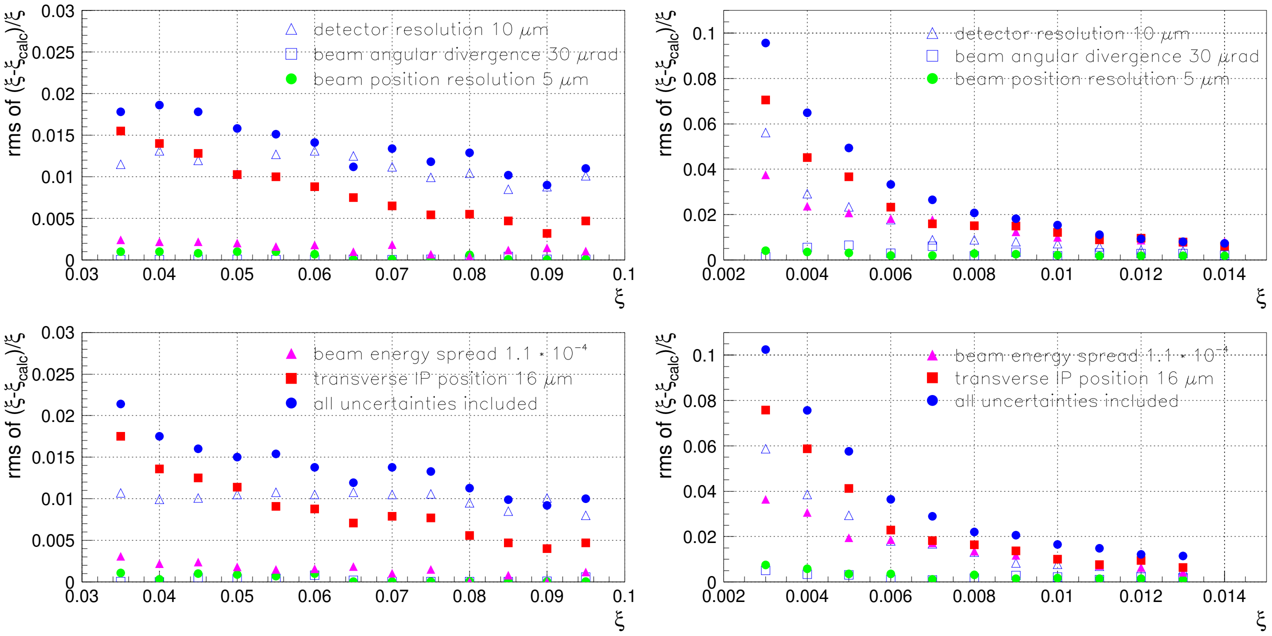


Fig. 1: Summary of all effects studied contributing to the over-all ξ resolution for the 215 m (left) and 420 m location (right). The upper and lower plots are for protons circulating clockwise and counter-clockwise along the LHC beam line, respectively. The t values of the protons used for each ξ bin is similar to the t distribution originating from central exclusive diffraction.

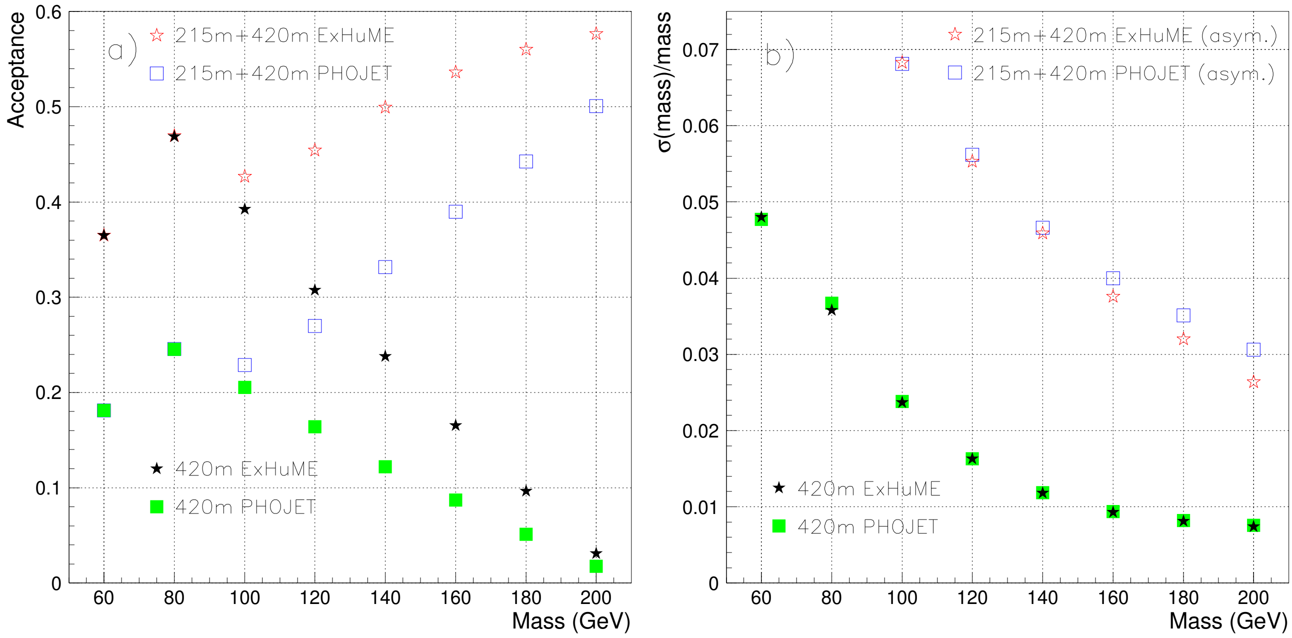


Fig. 2: a) Mass acceptance for events with protons within the acceptance of the 420 m location on each side of the interaction point ("420m"); and for events with protons within the combined acceptance of the 215 m and 420 m location on each side of the interaction point ("215m+420m"). b) Mass resolution for events with protons within the acceptance of the 420 m location on each side of the interaction point ("420m"); and for events with one proton within the acceptance of the 215 m location on one side of the interaction point and the other proton within the acceptance of the 420 m location on the other side ("215m+420m (asym.)"). ExHuME or PHOJET denotes the generator used for producing the central exclusive diffractive events.

TOTEM forward measurements: leading proton acceptance

V. Avati[a] and K. Österberg[b]
[a] CERN, Geneva, Switzerland and Case Western Reserve University, Cleveland, OH, USA
[b] Department of Physical Sciences, University of Helsinki and Helsinki Institute of Physics, Finland

Abstract
We report about the acceptance of forward leading protons in Roman Pot stations placed along the LHC beam line. The TOTEM stations plus additional detectors at 420 m from the interaction point have been considered using the low–β^* optics V6.5 for LHC physics runs.

1 Introduction

The TOTEM very forward detectors consist of telescopes of "Roman Pots" (RP) placed symmetrically on both sides of the interaction region IP5. The RP stations will be placed at 147 m and 220 m from IP5: each station is composed of two units separated by 2.5–4 m and each unit is equipped with two vertical and one horizontal silicon detector package. For more details on the TOTEM RPs, please refer to [1]. The possibility to add a detector in the cryogenic sections of the LHC is under investigation, therefore, we have included one more RP station at 420 m in these acceptance studies. This work is an update, due to the release of a new LHC optics, of previous studies done by the TOTEM Collaboration [2].

1.1 Low β^* optics acceptance study

The transverse displacement $(x(s), y(s))$[1] of a scattered leading proton (with momentum loss $\xi = \Delta P/P < 0$) at distance s from the interaction point (IP) is determined by tracking the proton through the accelerator lattice using the MAD-X program [3].

The new optics version 6.5 for the standard LHC physics runs is used. Notable changes (at IP5) from the previous versions are :

- β^* = 0.55 m (previously 0.5 m)
- Beam offset in the horizontal plane = 0.5 mm (previously zero)
- Horizontal crossing angle = 142 μrad (previously 150 μrad)

The protons at the IP are generated with flat distributions in the azimuthal angle ϕ, in Log($-\xi$) and in Log($-t$) in the kinematically allowed region of the ξ–t plane, i.e. for physical values of the scattering angle of the proton. The Mandelstam variable t is defined as $t = (p_{\text{orig}} - p_{\text{scatt}})^2$, where $p_{\text{orig(scatt)}}$ is the four-momentum of the incoming (scattered) proton. The scattering angle of the proton is physical when $t \geq t_0(\xi)$, where $t_0(\xi)$ is given by

$$t_0(\xi) = 2\left(P_{\text{orig}}^2 + m_p^2\right)\left[\sqrt{1 + \left(P_{\text{orig}}^2\left[\xi^2 + 2\xi\right]\right) / \left(P_{\text{orig}}^2 + m_p^2\right)} - 1\right] - 2\xi P_{\text{orig}}^2 \,. \qquad (1)$$

In Eq. 1, P_{orig} is the momentum of the incoming proton and m_p is the proton mass.

The transverse vertex position and the scattering angle at the IP are smeared assuming Gaussian distributions with widths given by the transverse beam size (16 μm) and the beam divergence (30 μrad).

To determine the acceptance of a RP station, the minimum distance of a detector to the beam and constraints imposed by the beam pipe or beam screen size are considered.

[1] The reference system (x,y,s) defines the reference orbit in the accelerator; the s-axis is tangent to the orbit and positive in the beam direction; the two other axes are perpendicular to the reference orbit. The x-axis (horizontal, bending plane) is negative toward the center of the ring.

The minimum distance of detector approach to the beam is proportional to the beam size:

$$x(y)_{min} = 10\sigma_{x(y)}{}^{beam} + c\,, \tag{2}$$

where c is a constant that takes into account the distance from the edge of the sensitive detector area to the bottom of the RP window ($\sim$ 0.5 mm). For the nominal transverse beam emittance ϵ = 3.75 $\mu\mathrm{m}\cdot\mathrm{rad}$ typical values of the horizontal detector distance are $\sim$1 mm (at 220 m) and $\sim$4 mm (at 420 m). In the results shown later, the detector shape has not been included. The beam pipe apertures can be found in the LHC–LAYOUT Database [4].

1.2 Results

Figs. 1–3 show the acceptance in Log($-\xi$), Log($-t$) for the RP stations at 220 and 420 m for the clockwise ("beam1") and counter–clockwise ("beam2") circulating beam.

One should note that these results refer to non-physical distributions in the variables ξ and t in order to have good statistics in each interval and describe all possible processes. To use these results in a general simulation program, the ϕ dependence has to be taken into account, since it is not negligible in many kinematical configurations. More detailed analysis such as detector alignment samples, collimator effects, etc. can be found in [5].

These results have been included in FAMOS (FAst MOntecarlo Simulation of the CMS detector) by M. Tasevsky ("Diffractive Higgs production", these proceedings) and they have been used in the CMS/TOTEM studies on triggering a diffractively produced light Higgs boson with the CMS Level-1 trigger ("Diffractive Higgs: CMS/TOTEM Level-1 Trigger Studies", these proceedings).

References

[1] The TOTEM collaboration, TOTEM Technical Design Report, **CERN-LHCC-2004-002 TOTEM-TDR-001** (2004).

[2] The TOTEM collaboration, Status Report to the LHCC **(http://www.cern.ch/totem)** (2002); J. Kalliopuska, J.W. Lämsä, T. Mäki, N. Marola, R. Orava, K. Österberg, M. Ottela and S. Tapprogge, "TOTEM forward measurements: exclusive central diffraction", these proceedings and J. Kalliopuska, T. Mäki, N. Marola, R. Orava, K. Österberg, M. Ottela and S. Tapprogge, **HIP-2003-11/EXP** (2003).

[3] The MAD-X Program, Methodical Accelerator Design, **http://www.cern.ch/mad**.

[4] LHC Layout, **http://lhclayout.web.cern.ch/lhclayout/**.

[5] V. Avati and K. Österberg, "Acceptance calculations methods for low–β^* optics", **TOTEM Internal Note 05-2 (http://www.cern.ch/totem)** (2005).

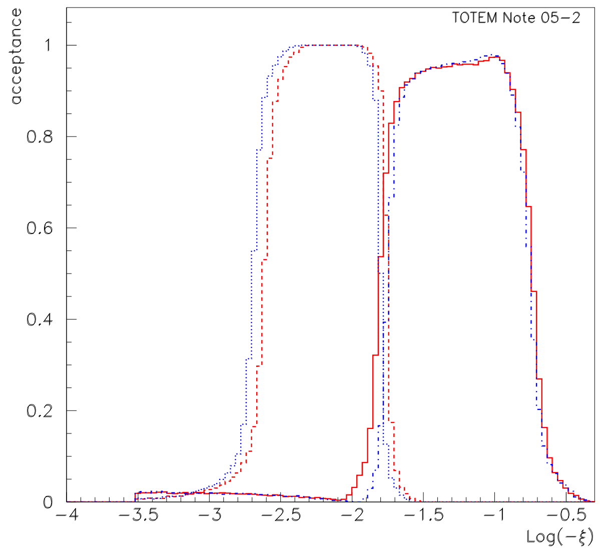


Fig. 1: Log($-\xi$) acceptance for: beam1 station at 220 m (solid-red) and 420 m (dashed-red) and beam2 station at 220 m (dashed-dotted-blue) and 420 m (dotted-blue).

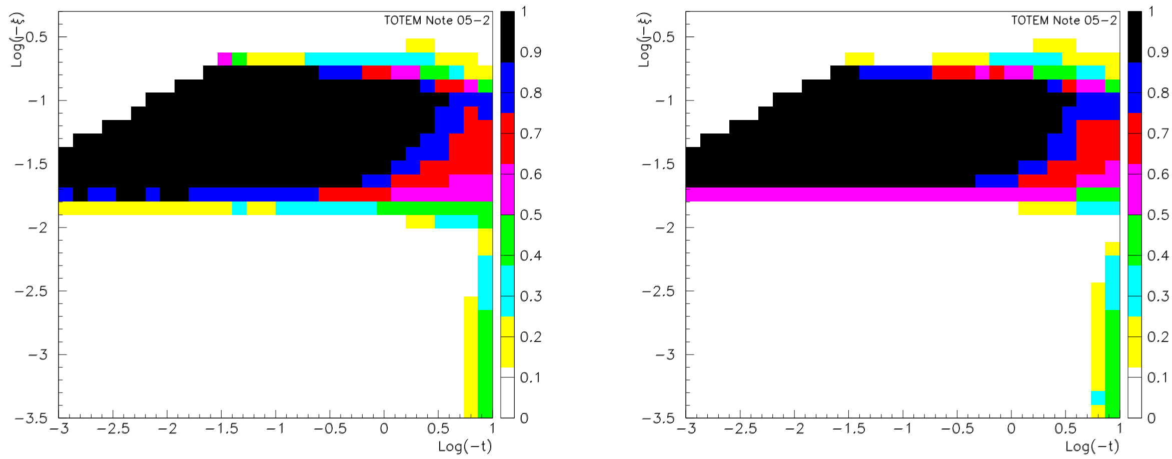


Fig. 2: Log($-\xi$) vs Log($-t$) acceptance for beam1 (left) and beam2 (right) stations at 220 m.

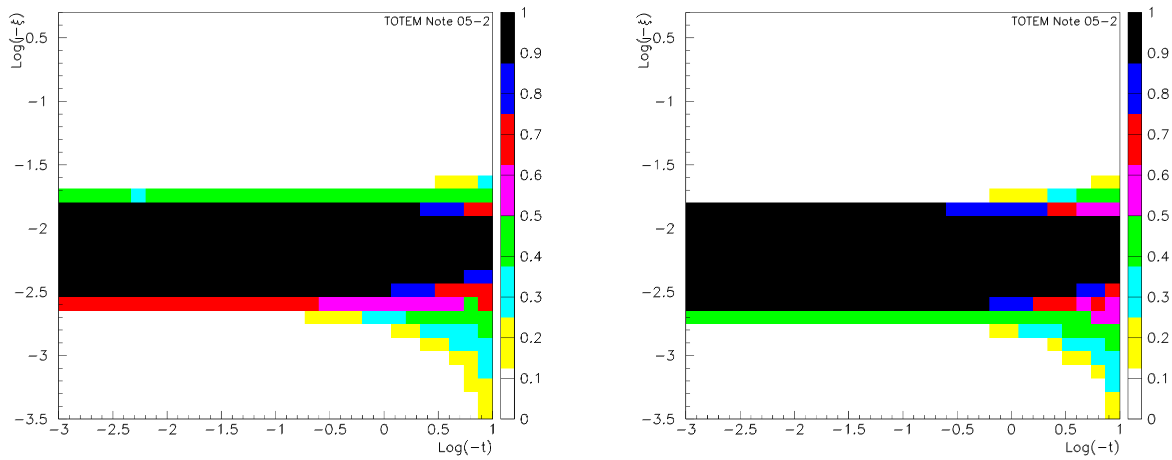


Fig. 3: Log($-\xi$) vs Log($-t$) acceptance for beam1 (left) and beam2 (right) stations at 420 m.

Diffractive Higgs: CMS/TOTEM Level-1 Trigger Studies

M. Arneodo[a], *V. Avati*[b], *R. Croft*[c], *F. Ferro*[d], *M. Grothe*[e*†], *C. Hogg*[f], *F. Oljemark*[g], *K. Österberg*[g], *M. Ruspa*[a]
[a]Università del Piemonte Orientale, Novara, and INFN-Torino, Italy
[b] European Organisation for Nuclear Research CERN, Geneva, Switzerland, and Case Western Reserve University, Cleveland, OH, USA
[c] University of Bristol, Bristol, UK
[d] University of Genova and INFN-Genova, Italy
[e] University of Torino and INFN-Torino, Italy; also at University of Wisconsin, Madison, WI, USA
[f] University of Wisconsin, Madison, WI, USA
[g] Department of Physical Sciences, University of Helsinki, and Helsinki Institute of Physics, Finland

Abstract
Retaining events containing a Higgs Boson with mass around 120 GeV poses a special challenge to triggering at the LHC due to the relatively low transverse momenta of the decay products. We discuss the potential of including into the CMS trigger the TOTEM forward detectors and possible additional detectors at a distance of 420 m from the CMS interaction point. We find that the output rate of a 2-jet Level-1 trigger condition with thresholds sufficiently low for the decay products of a 120 GeV Higgs Boson can be limited to $\mathcal{O}(1)$ kHz for luminosities of up to $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ by including the TOTEM forward detectors in the Level-1 trigger.

1 Introduction

A Higgs Boson with mass close to the current exclusion limit poses a special challenge to triggering at the LHC. The dominant decay of a Standard Model Higgs Boson of mass $\sim$120 GeV is into two b-quarks and generates 2 jets with at most 60 GeV transverse momentum, p_T, each. The so far considered Level-1 (L1) trigger tables of CMS [1] are optimized for events with high p_T; the necessity of keeping the overall L1 rate at acceptable levels requires thresholds in two-jet events above p_T =100 GeV per jet. Conversely, triggering is not a problem should the mass of the Higgs Boson be sufficiently high so that its final states are rich in high p_T leptons, as is the case for $H \rightarrow WW^{\star}$.

In order to retain a potential Higgs signal with mass close to the current exclusion limit, information beyond that from the central CMS detector needs to be included in the L1 trigger. A proton that scatters diffractively at the CMS interaction point (IP) may be detected by Roman Pot (RP) detectors further downstream. Roman Pot detectors up to 220 m downstream of CMS will be part of the TOTEM experiment [2]. Information from TOTEM will be available to the CMS L1 trigger. Furthermore, detectors at up to 420 m distance from the IP are currently discussed as part of the FP420 project [3]. Including information from them into the CMS L1 trigger is however not possible without an increase in the trigger latency.

This article discusses the effect of including the TOTEM forward detectors and/or those planned at 420 m distance on rate and selection efficiency of the CMS L1 trigger. All results reported in the following are preliminary; further studies are still on-going at the time of writing.

* Work supported by the Italian Ministry for Education, University and Scientific Research under the program "Incentivazione alla mobilità di studiosi stranieri e italiani residenti all'estero".

† Corresponding author: Monika.Grothe@cern.ch

2 Experimental apparatus

The CMS trigger system is designed to reduce the input rate of 10^9 interactions per second at the nominal LHC luminosity of $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ to an output rate of not more than 100 Hz. This reduction of 10^7 is achieved in two steps, by the CMS L1 trigger (output rate 100 kHz) and the CMS Higher-Level Trigger (HLT). The L1 trigger carries out its data selection algorithms with the help of three principal components: the Calorimeter Trigger, the Muon Trigger and the Global Trigger. The decision of the Calorimeter Trigger is based on the transverse energy, E_T, information of the CMS calorimeters (pseudorapidity coverage $|\eta| < 5$). A L1 jet consists of 3×3 regions, each with 4×4 trigger towers, where the E_T in the central region is above the E_T in any of the outer regions. A typical L1 jet has dimensions $\Delta\eta \times \Delta\phi = 1 \times 1$, where ϕ is the azimuthal angle. The E_T reconstructed by the L1 trigger for a given jet corresponds on the average only to 60% of its true E_T. All studies in this article use calibrated jet E_T values, obtained from the reconstructed value by means of an η and E_T dependent correction.

The TOTEM experiment [2] will have two identical arms, one at each side of the CMS IP. Each arm will comprise two forward tracker telescopes, T1 (Cathode Strip Chambers) and T2 (Gas Electron Multipliers), as well as Silicon detectors housed in RP stations along the LHC beam-line. The TOTEM detectors will provide input data to the Global Trigger of the CMS L1 trigger. Track finding in T1 and T2 (combined coverage $3.2 < |\eta| < 6.6$) for triggering purposes is optimized with respect to differentiating between beam-beam events that point back to the IP and beam-gas and beam-halo events that do not. The TOTEM RP stations will be placed at a distance of ± 147 m and ± 220 m from the CMS IP. Each station will consist of two units, 2.5 m and 4 m apart, each with one horizontally and two vertically movable pots equipped with Silicon strip detectors. The possibility of implementing a cut on ξ in the L1 trigger is currently under investigation.

The fractional momentum loss, ξ, of diffractively scattered protons peaks at $\xi = 0$ ("diffractive peak"). The combination of CMS and TOTEM will permit to measure protons that have undergone a fractional momentum loss $0.2 > \xi > 0.02$. Detectors at a distance of 420 m, in the cryogenic region of the LHC ring, are currently being considered by the FP420 project [3]. They would provide a coverage of $0.02 > \xi > 0.002$, complementary to that of the TOTEM detectors, but cannot be included in the L1 trigger without an increase in the L1 latency of 3.2 μs. A special, long latency running mode might be feasible at lower luminosities. This option is currently under investigation. Using detectors at 420 m in the L1 trigger is included as an option in the studies discussed in this article.

The studies discussed in the following assume that the RP detectors are 100% efficient in detecting all particles that emerge at a distance of at least $10\sigma_{beam} + 0.5$ mm from the beam axis. Their acceptance was calculated by way of a simulation program that tracks particles through the accelerator lattice [4]. This has been done for the nominal LHC optics, the so-called low-β^* optics, version V6.5. Further details can be found in [5]. All Monte Carlo samples used in the following assume LHC bunches with 25 ns spacing.

3 Level-1 trigger rates and signal efficiencies

We consider here perhaps the most challenging case, that of a low-mass (120 GeV) Standard Model Higgs Boson, decaying into two b-jets. There, the jets have transverse energies of at most 60 GeV. In order to retain as large a signal fraction as possible, as low an E_T threshold as possible is desirable. In practice, the threshold value cannot be chosen much lower than 40 GeV per jet. The L1 trigger applies cuts on the calibrated E_T value of the jet. Thus, a threshold of 40 GeV corresponds to 20 to 25 GeV in reconstructed E_T, i.e. to values where noise effects start becoming sizable.

In the trigger tables forseen for the first LHC running period, a L1 2-jet rate of $\mathcal{O}(1)$ kHz is planned. For luminosities of $10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and above, the rate from standard QCD processes for events with at least 2 central jets ($|\eta| < 2.5$) with $E_T > 40$ GeV is above this. Thus additional conditions need to be employed in the L1 trigger to reduce the rate from QCD processes. The efficiency of several conditions was investigated and, in the following, the corresponding rate reduction factors are always quoted with respect to the rate of QCD events that contain at least 2 central jets with $E_T > 40$ GeV per jet. These conditions are:

1) Condition based on additional central detector quantities available to the Calorimeter Trigger.
2) Condition based on T1 and T2 as vetoes.
3) Condition based on the RP detectors at ± 220 m and ± 420 m distance from the CMS IP.
4) Condition based on the Muon Trigger.

The QCD background events were generated with the Pythia Monte Carlo generator.

At higher luminosities more than one interaction takes place per bunch crossing; the central exclusive production of a Higgs boson is overlaid with additional, typically soft events, the so-called pile-up. In order to assess the effect when the signal is overlaid with pile-up, a sample of 500,000 pile-up events was generated with Pythia. This sample includes inelastic as well as elastic and diffractive events. Pythia underestimates the number of final state protons in this sample. The correction to the Pythia leading proton spectrum described in [6] was used to obtain the results discussed in the following.

The effect from beam-halo and beam-gas events on the L1 rate is not yet included in the studies discussed here. Preliminary estimates suggest that the size of their contribution is such that the conclusions of this article are not invalidated.

Table 1 summarizes the situation for luminosities between $10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Given a target rate for events with 2 central L1 jets of $\mathcal{O}(1)$ kHz, a total rate reduction between a factor 20 at $1 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and 200 at $1 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is necessary.

Luminosity $[\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	# Pile-up events per bunch crossing	L1 2-jet rate [kHz] for $E_T > 40$GeV per jet	Total reduction needed	Reduction when requiring track in RP detectors: at 220 m	at 220 m, $\xi < 0.1$	at 420 m	at 220 m & 420 m (asymmetric)	at 220 m & 420 m (asymmetric), $\xi < 0.1$
1×10^{32}	0	2.6	2	370				
1×10^{33}	3.5	26	20	7	15	27	160	380
2×10^{33}	7	52	40	4	10	14	80	190
5×10^{33}	17.5	130	100	3	5	6	32	75
1×10^{34}	35	260	200	2	3	4	17	39

Table 1: Reduction of the rate from standard QCD processes for events with at least 2 central L1 jets with $E_T > 40$ GeV, achievable with requirements on the tracks seen in the RP detectors. Additional rate reductions can be achieved with the H_T condition and with a topological condition (see text). Each of them yields, for all luminosities listed, an additional reduction by about a factor 2.

3.1 Condition based on central CMS detector quantities

In addition to the E_T values of individual L1 jets, the CMS Calorimeter Trigger has at its disposal the scalar sum, H_T, of the E_T values of all jets. Requiring that essentially all the E_T be concentrated in the two central L1 jets with highest E_T, i.e. $[E_T^1 + E_T^2]/H_T > 0.9$ (H_T condition), corresponds to imposing a rapidity gap of at least 2.5 units with respect to the beam direction. This condition reduces the rate of QCD events by approximately a factor 2, independent of the presence of pile-up and with only a small effect on the signal efficiency.

3.2 Condition based on TOTEM detectors T1 and T2

Using T1 and T2 as vetoes in events with 2 central L1 jets imposes the presence of a rapidity gap of at least 4 units. This condition suppresses QCD background events by several orders of magnitude. At luminosities low enough so that not more than one interaction takes place per bunch crossing, the signal efficiency is very high ($> 90\%$). In the presence of pile-up, the signal efficiency falls rapidly. The non-diffractive component in pile-up events tends quickly to fill in the rapidity gap in the Higgs production process. Only about 20 (5) % of signal events survive in the presence of 1 (2) pile-up event(s).

3.3 Condition based on Roman Pot detectors

Demanding that a proton be seen in the RP detectors at 220 m results in excellent suppression of QCD background events in the absence of pile-up. This is demonstrated in Figure 1 for a luminosity of $10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. There, the rate of QCD background events with at least 2 central L1 jets with E_T above a threshold is shown as function of the threshold value. The two histograms reflect the rate without and with the requirement that a proton be seen in the RP detectors at 220 m. The rate of QCD background events containing at least 2 central L1 jets with $E_T > 40$ GeV each is reduced by a factor ~ 370. At $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, where on the average 7 pile-up events overlay the signal event, the diffractive component in the pile-up causes the reduction to decrease to a factor ~ 4, and at $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, to a factor ~ 2, as can be seen from table 1.

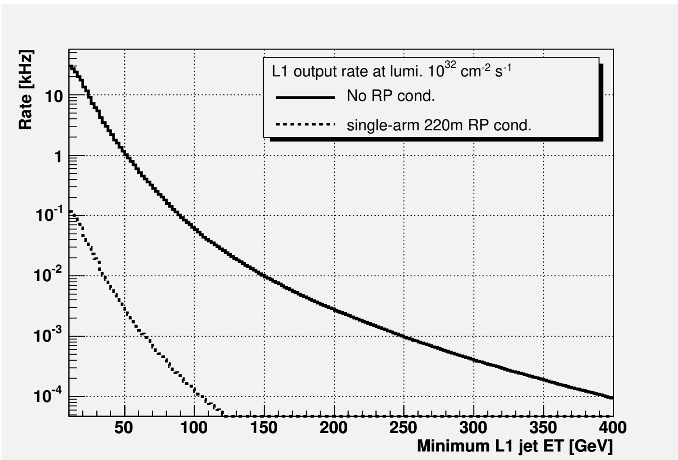


Fig. 1: L1 rate for the QCD background at a luminosity of $10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ as function of the L1 threshold value when requiring at least 2 central L1 jets with E_T above threshold.

Table 1 summarizes the reduction factors achieved with different conditions for tracks in the RP detectors: a track in the RP detectors at 220 m distance on one side of the IP (single-arm 220 m), without and with a cut on ξ, a track in the RP detectors at 420 m distance on one side of the IP (single-arm 420 m), a track in the RP detectors at 220 m and 420 m distance (asymmetric). Because the detectors at 220 m and 420 m have complementary coverage in ξ, the last condition in effect selects events with two tracks of very different ξ value, in which one track is seen at 220 m distance on one side of the IP and a second track is seen on the other side at 420 m. If not by the L1 trigger, these asymmetric events can be selected by the HLT and are thus of highest interest. The effect on the acceptance of the RP detectors of a collimator located in front of the LHC magnet Q5, which will be operative at higher luminosities, has not been taken into account in table 1.

A further reduction of the QCD rate could be achieved with the help of a topological condition. The 2-jet system has to balance the total momentum component of the two protons along the beam axis. In signal events with asymmetric ξ values, the proton seen on one side in the RP detectors at 220 m distance is the one with the larger ξ and thus has lost more of its initial momentum component along the beam axis. Hence the jets tend to be located in the same η-hemisphere as the RP detectors that detect this proton. A trigger condition requiring that $[\eta^{jet1} + \eta^{jet2}] \times \mathrm{sign}(\eta^{220m\ RP}) > 0$ would reduce the QCD background by a factor 2, independent of pile-up, and with no loss in signal efficiency.

A reduction of the QCD rate to levels compatible with the trigger bandwidth requirements by including RP detectors at a distance of 220 m from the CMS IP thus appears feasible for luminosities up to $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, as long as a ξ cut can be administered in the L1 trigger such that the accepted events can be restricted to the diffractive peak region around $\xi = 0$. Higher luminosities would necessitate inclusion of the RP detectors at 420 m distance in the L1 trigger.

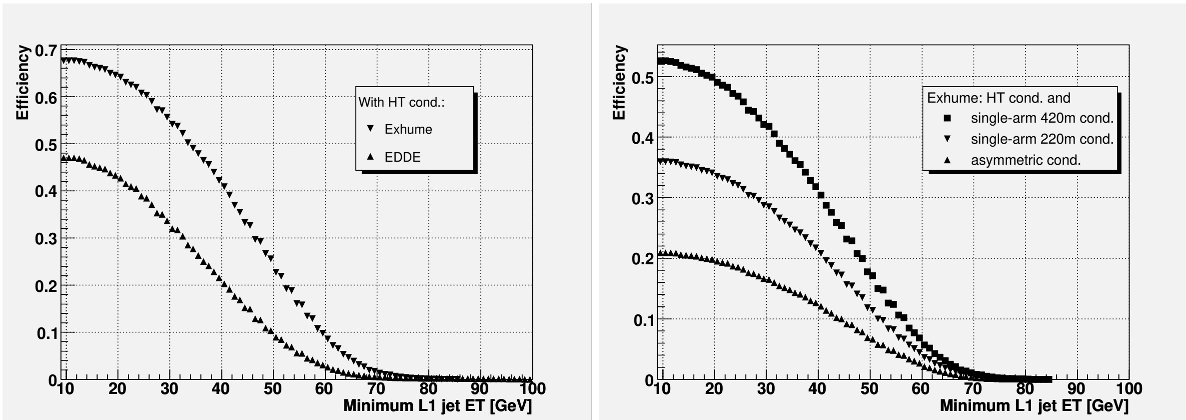


Fig. 2: L1 selection efficiency as function of the E_T threshold value when requiring at least 2 central L1 jets with E_T above threshold. All plots are for the non-pile-up case and the H_T condition (see text) has been applied. Left: Comparison between the EDDE and Exhume Monte Carlo generators, without applying any additional RP conditions. Right: Comparison of the effect of different RP conditions on the efficiency in the Exhume Monte Carlo sample.

In order to study the effect of the L1 trigger selection on the Higgs signal, signal samples of 20,000 events with central exclusive production of a Higgs Boson were generated with the Monte-Carlo programs EDDE [7] (version 1.1) and Exhume [8] (version 0.9). Figure 2 shows the L1 selection efficiency as a function of the E_T threshold values when requiring at least 2 central L1 jets with E_T above threshold. The histograms show the case when no pile-up is present. The presence of pile-up has only a small effect on the efficiency curves. The plot on the left-hand side compares the efficiency curves obtained for EDDE and Exhume. For a threshold of 40 GeV per jet, Exhume yields an efficiency of about 40%. As a consequence of its less central jet η distribution (see [9]), the efficiency for EDDE is about 20% lower than the one of Exhume. The plot on the right-hand side overlays the efficiency curves obtained with Exhume when including three different RP detector conditions in the L1 2-jet trigger: single-arm 220 m, single-arm 420 m and the asymmetric 220 & 420 m condition. At an E_T threshold of 40 GeV per jet, the single-arm 220 m (420 m) condition results in an efficiency of the order 20% (30%), the asymmetric condition in one of 15%. This also means that even without the possibility of including the RP detectors at 420 m distance from the CMS IP in the L1 trigger, 15% of the signal events can be triggered with the single-arm 220 m condition, but will have a track also in the 420 m detectors which can be used in the HLT.

3.4 Condition based on the Muon Trigger

An alternative trigger strategy may be to exploit the relatively muon-rich final state from B-decays. We estimate that up to 10% of the signal events could be retained using this technique. Further investigations are underway at the time of writing.

4 Conclusions

Retaining a Higgs Boson with mass around 120 GeV poses a special challenge to triggering at the LHC. The relatively low transverse momenta of its decay products necessitate L1 jet E_T thresholds as low as 40 GeV. Thresholds that low would result in a L1 trigger rate of more than 50 kHz, essentially saturating the available output bandwidth.

The results we presented in this article are preliminary and should be taken as a snapshot of our present understanding. They can be summarized as follows: The output rate of a 2-jet L1 trigger condition with thresholds of 40 GeV per jet can be kept at an acceptable $\mathcal{O}(1)$ kHz by including the TOTEM forward detectors in the CMS L1 trigger. In the absence of pile-up, either using the TOTEM T1 and T2 detectors as vetoes or requiring that a proton be seen in the TOTEM RP detectors at 220 m on one side of the CMS IP (single-sided 220 m condition) results in a sufficient reduction of the QCD event rate that dominates the L1 trigger output rate. At higher luminosities, up to $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, where pile-up is present, it is necessary to combine the single-sided 220 m condition with conditions based on event topology and on H_T, the scalar sum of all L1 jet E_T values. Going to even higher luminosities, up to $1 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, would necessitate additional L1 trigger conditions, such as inclusion of RP detectors at 420 m distance from the CMS IP, which, however, would require an increase in the L1 trigger latency. These L1 trigger conditions result in signal efficiencies between 15% and 20%.

We expect no trigger problems for final states rich in high p_T leptons, such as the WW decay modes of the Standard Model Higgs Boson.

Acknowledgement

We are grateful to Sridhara Dasu for sharing with us his expertise on the CMS L1 trigger simulation and for providing considerable pratical help with it. We would like to thank him and Dan Bradley for their invaluable help with producing the Monte Carlo event samples used in this article.

References

[1] “CMS: The TRIDAS project – Technical Design Report Vol.1: The trigger systems”, CMS collab., CERN-LHCC-2000-038.
“CMS: The TRIDAS project – Technical Design Report Vol.2: Data Acquisition and High-Level Trigger”, CMS collab., CERN-LHCC-2002-026.

[2] “TOTEM Technical Design Report”, TOTEM collab., CERN-LHCC-2004-002 and Addendum to the TOTEM-TDR, TOTEM collab., CERN-LHCC-2004-020.

[3] “FP420: A proposal to investigate the feasibility of installing proton tagging detectors in the 420 m region of the LHC”, M.G.Albrow et al., CERN-LHCC-2005-025, LHCC-I-015.
available from: glodwick.hep.man.ac.uk/brian/fp420loi.doc

[4] The MAD-X Program, Methodical Accelerator Design, available from http://www.cern.ch/mad.

[5] “TOTEM forward measurements: leading proton acceptance”, V. Avati, K. Österberg, these proc.
“TOTEM forward measurements: exclusive central diffraction”, J. Kalliopuska et al., these proc.

[6] “Leading proton production in ep and pp experiments: how well do high-energy physics Monte Carlos reproduce the data ?”, G.Bruni et al., these proc.

[7] “EDDE Monte Carlo event generator”, R.A. Ryutin, hep-ph/0409180.

[8] “Exhume: A Monte Carlo event generator for exclusive diffraction”, J. Monk, A. Pilkington, hep-ph/0502077.

[9] "Monte-Carlo generators for central exclusive diffraction", M. Boonekamp et al., these proc.

Proposal to upgrade the very forward region at CMS

V. Andreev[1], *A.Bunyatyan*[2,3], *H. Jung*[4], *M. Kapishin*[5], *L. Lytkin*[3,5]
[1] Lebedev Physics Institute, Moscow, [2] Yerevan Physics Institute,
[3] MPI-K Heidelberg, [4] DESY Hamburg, [5] JINR Dubna

Abstract
The possibilities of extending the acceptance of LHC experiments beyond 7 units of pseudorapidity are investigated. With additional detectors it would be possible to measure the particles with energies above 2 TeV in the pseudorapidity range between 7 and 11.

1 Introduction

At the LHC experiments, CMS and ATLAS, the acceptance for forward energy measurements is limited to about 5 units of pseudorapidity. The acceptance of CMS detector will be extended by proposed CASTOR calorimeter, which will cover the angular range $5.4 < \eta < 6.7$. Already with this device small-x parton dynamics can be studied down to very small x-values of $10^{-6} - 10^{-7}$ with Drell-Yan, prompt photon and jet events at small invariant masses of the order of $M \sim 10$ GeV.

In the present work we investigate the technical possibilities of extending the angular acceptance for forward energy measurement beyond 7 units of pseudorapidity. Extending the acceptance down to $\eta \sim 11$, x-values down to 10^{-8} can be reached, which is a completely unexplored region of phase space. In this region, effects coming from new parton dynamics are expected to show up, as well as effects coming from very high density gluonic systems, where saturation and recombination effects will occur. In this region of phase space, a breakdown of the usual factorization formalism is expected, and multiple interactions will be dominant (see Fig 1) [1]. The full angular coverage from the central to the most forward region allows a systematic study of the transition from single particle exchange processes to complex systems and a systematic understanding of non-linear and collective phenomena.

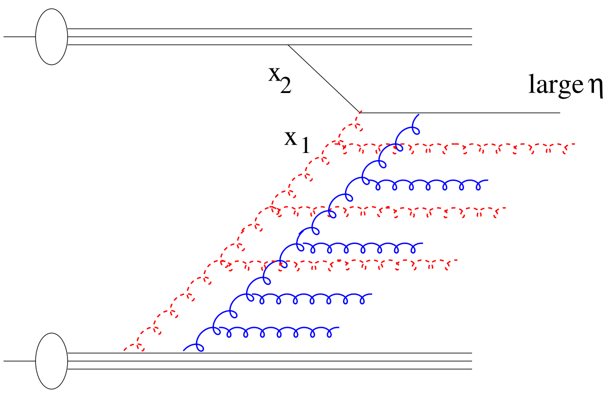


Fig. 1: Schematic picture of multiple interactions at small x

The interest in the very forward region of phase space is not only motivated by the fundamental understanding of QCD in a new phase of matter, but is also important for the further understanding of high energetic cosmic rays [2].

2 Tracking and energy measurement in the very forward region

For this study the geometry of the beam-line around the CMS detector up to 150m from interaction point has been implemented in the GEANT-3 [3] simulation program. The PYTHIA Monte-Carlo genera-

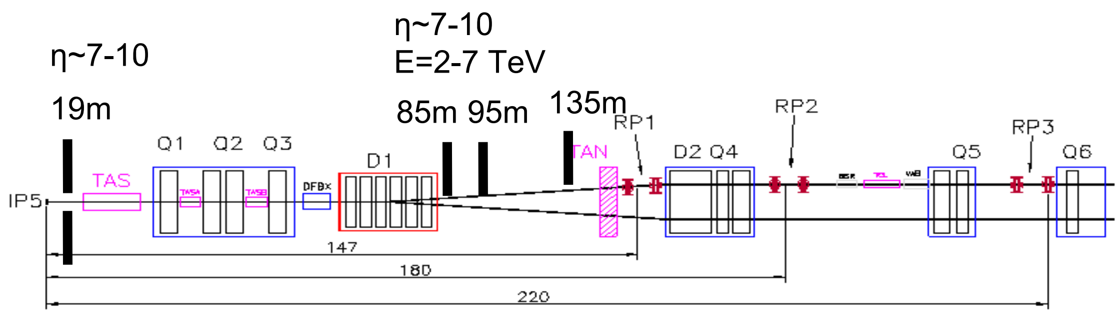


Fig. 2: Schematic view of forward beamline at CMS detector up to 220m. The positions of proposed tracking devices at 19m, 85m and 95m and calorimeter at 135m are indicated.

tor program [4] was used to generate charged particles produced in the interaction point, which were subsequently fed into the beam-line simulation.

The main restriction for additional installations is the very limited space available between magnetic elements. Up to about 80m there is no space for a calorimeter, and there one can only consider the installation of tracking devices, such as Roman Pots or micro-stations [5]. On the other hand, to be able to measure the particle momenta, the tracking devices should be placed after bending magnets.

Therefore the idea is to have tracking devices before and after dipoles, to be able to measure both the integrated particle flow and the particle momenta. The free space after 135 m can be used for a calorimeter.

Taking into account the limited space available for new detectors, the background conditions and magnetic field, the following strategy is proposed (Fig.2):

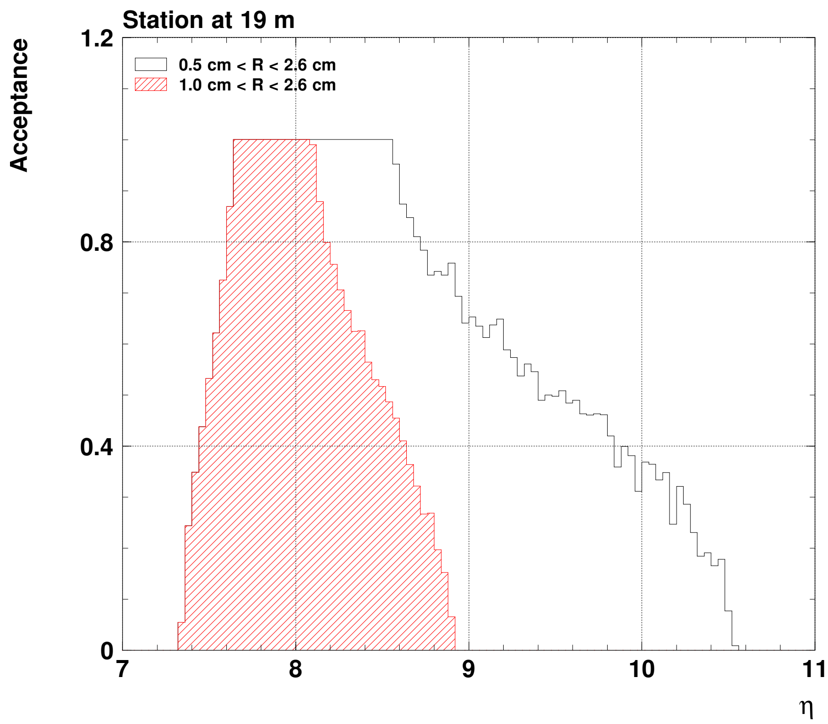


Fig. 3: Acceptance of micro-station detector at 19m as a function of pseudorapidity.

- the 25 cm space in front of TAS absorber at 19 m from interaction point can be used to install two

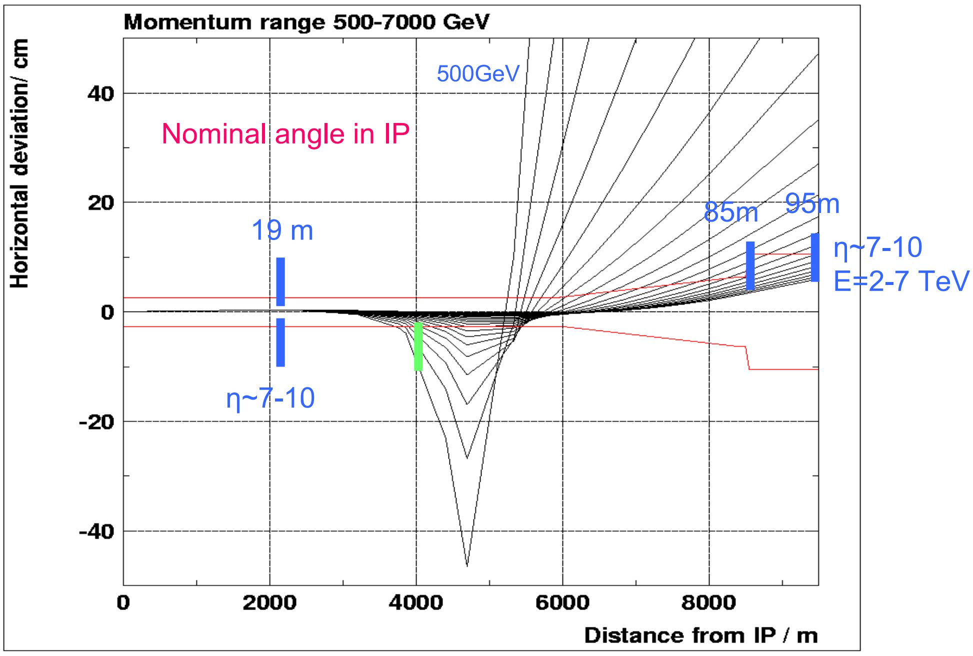


Fig. 4: The trajectory of particles in momentum range 500-7000 GeV, scattered from interaction point at 0 deg. The positions of proposed tracking devices at 19m, 85m and 95m are indicated.

micro-stations with two half-ring radiation hard silicon or diamond detectors approaching the beam horizontally up to 5-10mm. At this position the particles can not be separated by their momenta, thus the micro-stations will measure the charged particle flow integrated up to energies of ~ 7 TeV in the pseudorapidity range between 7.3 and 9 or 10.5, depending on how close the counter can go to the beam (see Fig. 3). The detector will also provide accurate position measurement which is necessary for linking with roman-pots/micro-stations installed further down the beam line. Combining the position and time-of-flight measurement of these micro-stations with the event vertex measured in central detector will allow to suppress beam-wall background and pile-up events;

- a combination of two horizontal roman-pots/micro-stations can be installed behind the dipole magnets D1 at 85m and 95m. The detectors are the half-rings and approach the beam horizontally from one side up to 10mm. These detectors will cover the pseudorapidity range above 8 units (see Fig. 5). The particles with energies below 2 TeV will escape the detector acceptance, as shown in Fig. 4 and 5.
- a hadronic calorimeter at 135m (in front of TAN iron absorber) with a minimal distance to the beam of 10cm (radius of beam-pipe) will measure the energy in the range 2-5.5 TeV and pseudorapidity between 7 and 11 (see Fig.6). This can be a sandwich type calorimeter with radiation hard sensitive layers, with a transverse size up to 1×1 m^2 and depth about 7-9 hadronic interaction length. Optionally one can consider to instrument the TAN absorber with sensitive layers.
  The calorimeter covers basically the same kinematic range as the micro-stations at 85 and 95m, but it is needed for energy measurement. In addition it can be used for redundancy, background subtraction and cross calibration.

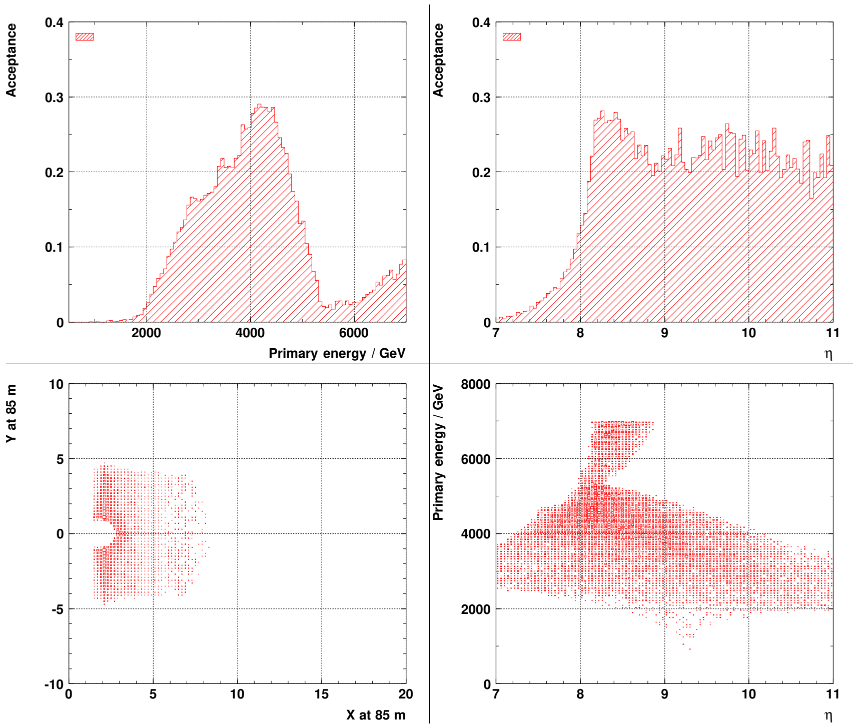


Fig. 5: Acceptance of roman pots/micro-station detector at 85 m and 95m as a function of energy and pseudorapidity.

The acceptances of the proposed detectors as function of energy and pseudorapidity are summarized in table 1. The result shows that with the proposed installations it will be possible to measure the energy flow in the energy range between 2 and 7 GeV and the pseudorapidity range between 7 and 11.

To be able to measure the particles with momenta below 2 TeV one would need to install detectors in the cold area between the quadrupoles at 40-50m. This will require essential modifications of cryogenic lines, and can be considered for a future upgrade program.

	0.5–7 TeV	2–5.5 TeV
two roman-pots/micro-stations at 85 and 95m		
$\eta = 7 - 10$	11%	21%
$\eta = 7 - 8$	10%	10–20%
$\eta = 8 - 9$	15-25%	30–55%
$\eta = 9 - 11$	20-25%	55–60%
Calorimeter at 135m		
$\eta = 7 - 8$	15%	25%
$\eta = 8 - 9$	20-25%	35–55%
$\eta = 9 - 11$	25-40%	45–60%

Table 1: Acceptance as a function of E_p and η

3 Conclusions

We have studied the possibilities of extending the angular acceptance for forward energy measurement at LHC. With additional roman-pots/micro-stations and a calorimeter it will be possible to measure the forward energy in the rapidity range between 7 and 11 units. Such installation will be a valuable addition to the LHC physics program.

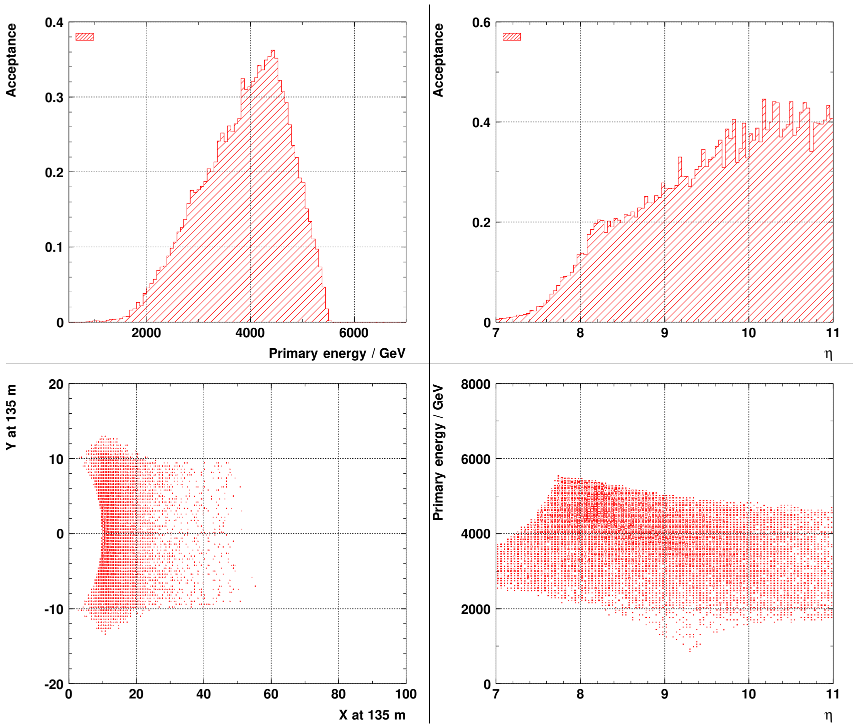


Fig. 6: Acceptance of calorimeter at 135 m as a function of energy and pseudorapidity.

References

[1] J. Bartels, *AGK cutting rules for pp collisions*. Talk at HERA-LHC workshop, March 2005.

[2] J. Ellis, *LHC - physics*. Plenary talk at HERA-LHC workshop, March 2005.

[3] Brun, R. and Bruyant, F. and Maire, M. and McPherson, A. C. and Zanarini, P., *GEANT3*. CERN-DD/EE/84-1.

[4] Sjostrand, Torbjorn and others, Comput. Phys. Commun. **135**, 238 (2001). Hep-ph/0010017.

[5] Nomokonov, V. P., *The microstation concept for forward physics*. Prepared for 1st Workshop on Forward Physics and Luminosity Determination at LHC, Helsinki, Finland, 31 Oct - 4 Nov 2000.

Diffractive Higgs production: theory

Jeff Forshaw
Particle Physics Group, School of Physics & Astronomy,
University of Manchester, Manchester, M13 9PL. United Kingdom.

Abstract
We review the calculation for Higgs production via the exclusive reaction $pp \to p + H + p$. In the first part we review in some detail the calculation of the Durham group and emphasise the main areas of uncertainty. Afterwards, we comment upon other calculations.

1 Introduction

Our aim is to compute the cross-section for the process $pp \to p + H + p$. We shall only be interested in the kinematic situation where all three final state particles are very far apart in rapidity with the Higgs boson the most central. In this "diffractive" situation the scattering protons lose only a very small fraction of their energy, but nevertheless enough to produce the Higgs boson. Consequently, we are in the limit where the incoming protons have energy E much greater than the Higgs mass m_H and so we will always neglect terms suppressed by powers of m_H/E. In the diffractive limit cross-sections do not fall as the beam energy increases as a result of gluonic (spin-1) exchanges in the t-channel.

Given the possibility of instrumenting the LHC to detect protons scattered through tiny angles with a high resolution [1, 2, 3, 4], diffractive production of any central system X via $pp \to p + X + p$ is immediately of interest if the production rate is large enough. Even if X is as routine as a pair of high p_T jets we can learn a great deal about QCD in a new regime [2, 3, 5, 6]. But no doubt the greatest interest arises if X contains "new physics" [7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]. The possibility arises to measure the new physics in a way that is not possible using the LHC general purpose detectors alone. For example, its invariant mass may be measured most accurately, and the spin and CP properties of the system may be explored in a manner more akin to methods hitherto thought possible only at a future linear collider. Our focus here is on the production of a Standard Model Higgs boson [7, 8, 13, 18, 19]. Since the production of the central system X effectively factorizes, our calculation will be seen to be of more general utility.

Most of the time will be spent presenting what we shall call the "Durham Model" of central exclusive production [7, 8]. It is based in perturbative QCD and is ultimately to be justified a posteriori by checking that there is not a large contribution arising from physics below 1 GeV. A little time will also be spent explaining the non-perturbative model presented by the Saclay group [13] and inspired by the original paper of Bialas and Landshoff [20]. Even less time will be devoted to other approaches which can be viewed, more-or-less, as hybrids of the other two [18, 19].

Apart from the exclusive process we study here, there is also the possibility to produce the new physics in conjunction with other centrally produced particles, e.g. $pp \to p + H + X + p$. This more inclusive channel typically has a much higher rate but does not benefit from the various advantages of exclusive production. Nevertheless, it must be taken into account in any serious phenomenological investigation into the physics potential of central exclusive production [21, 22]

2 The Durham Model

The calculation starts from the easier to compute parton level process $qq \to q + H + q$ shown in Figure 1. The Higgs is produced via a top quark loop and a minimum of two gluons need to be exchanged in order that no colour be transferred between the incoming and outgoing quarks. Quark exchange in the

t-channel leads to contributions which are suppressed by an inverse power of the beam energy and so the diagram in Figure 1 is the lowest order one. Our strategy will be to compute only the imaginary part of the amplitude and we shall make use of the Cutkosky rules to do that – the relevant cut is indicated by the vertical dotted line in Figure 1. There is of course a second relevant diagram corresponding to the Higgs being emitted from the left-hand gluon. We shall assume that the real part of the amplitude is negligible, as it will be in the limit of asymptotically high centre-of-mass energy when the quarks are scattered through small angles and the Higgs is produced centrally.

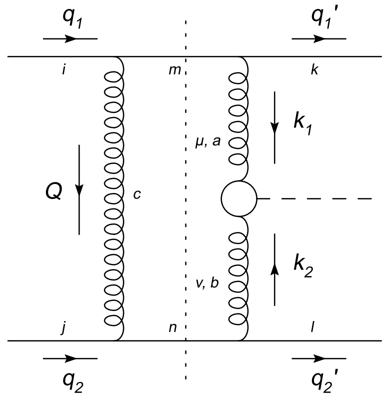


Fig. 1: The relevant Feynman graph for $qq \to q + H + q$.

The calculation can be further simplified by making use of the eikonal approximation for those vertices which couple the gluons to the external quarks. The gluons are very soft and so, modulo corrections which are suppressed by the inverse of the beam energy, we can approximate the qqg vertices by $2g\tau^a_{ij} q_{1,2}\delta_{\lambda,\lambda'}$, where τ^a is a Gell-Mann matrix, g is the QCD coupling and the Kronecker delta tells us that the quark does not change its helicity. The calculation of the amplitude is now pretty straightforward:

$$\begin{aligned}\mathrm{Im}A^{ik}_{jl} &= \frac{1}{2} \times 2 \int d(PS)_2\ \delta((q_1 - Q)^2)\delta((q_2 + Q)^2) \\ &\quad \frac{2gq_1^\alpha\ 2gq_{2\alpha}}{Q^2}\ \frac{2gq_1^\mu}{k_1^2}\ \frac{2gq_2^\nu}{k_2^2}\ V^{ab}_{\mu\nu}\ \tau^c_{im}\tau^c_{jn}\tau^a_{mk}\tau^b_{nl}\ . \qquad (1)\end{aligned}$$

The factor of $1/2$ is from the cutting rules and the factor of 2 takes into account that there are two diagrams. The phase-space factor is

$$d(PS)_2 = \frac{s}{2} \int \frac{d^2\mathbf{Q_T}}{(2\pi)^2}\ d\alpha d\beta \qquad (2)$$

where we have introduced the Sudakov variables via $Q = \alpha q_1 + \beta q_2 + Q_T$. The delta functions fix the cut quark lines to be on-shell, which means that $\alpha \approx -\beta \approx \mathbf{Q_T}^2/s \ll 1$ and $Q^2 \approx Q_T^2 \equiv -\mathbf{Q_T}^2$. As always, we are neglecting terms which are energy suppressed such as the product $\alpha\beta$. For the Higgs production vertex we take the Standard Model result:

$$V^{ab}_{\mu\nu} = \delta^{ab}\left(g_{\mu\nu} - \frac{k_{2\mu}k_{1\nu}}{k_1 \cdot k_2}\right) V \qquad (3)$$

where $V = m_H^2\alpha_s/(4\pi v)F(m_H^2/m_t^2)$ and $F \approx 2/3$ provided the Higgs is not too heavy. The Durham group also include a NLO K-factor correction to this vertex. After averaging over colours we have

$$\tau^c_{im}\tau^c_{jn}\tau^a_{mk}\tau^b_{nl} \to \frac{\delta^{ab}}{4N_c^2}.$$

We can compute the contraction $q_1^\mu V_{\mu\nu}^{ab} q_2^\nu$ either directly or by utilising gauge invariance which requires that $k_1^\mu V_{\mu\nu}^{ab} = k_2^\nu V_{\mu\nu}^{ab} = 0$. Writing[1] $k_i = x_i q_i + k_{iT}$ yields

$$q_1^\mu V_{\mu\nu}^{ab} q_2^\nu \approx \frac{k_{1T}^\mu}{x_1} \frac{k_{2T}^\nu}{x_2} V_{\mu\nu}^{ab} \approx \frac{s}{m_H^2} k_{1T}^\mu k_{2T}^\nu V_{\mu\nu}^{ab} \tag{4}$$

since $2k_1 \cdot k_2 \approx x_1 x_2 s \approx m_H^2$. Note that it is as if the gluons which fuse to produce the Higgs are transversely polarized, $\epsilon_i \sim k_{iT}$. Moreover, in the limiting case that the outgoing quarks carry no transverse momentum $Q_T = -k_{1T} = k_{2T}$ and so $\epsilon_1 = -\epsilon_2$. This is an important result; it clearly generalizes to the statement that the centrally produced system should have a vanishing z-component of angular momentum in the limit that the protons scatter through zero angle (i.e. $q_{iT}'^2 \ll Q_T^2$). Since we are experimentally interested in very small angle scattering this selection rule is effective. One immediate consequence is that the Higgs decay to b-quarks may now be viable. This is because, for massless quarks, the lowest order $q\bar{q}$ background vanishes identically (it does not vanish at NLO). The leading order $b\bar{b}$ background is therefore suppressed by a factor $\sim m_b^2/m_H^2$. Beyond leading order, one also needs to worry about the $b\bar{b}g$ final state.

Returning to the task in hand, we can write the colour averaged amplitude as

$$\frac{\mathrm{Im}A}{s} \approx \frac{N_c^2-1}{N_c^2} \times 4\alpha_s^2 \int \frac{d^2\mathbf{Q_T}}{\mathbf{Q_T}^2 \mathbf{k_{1T}}^2 \mathbf{k_{2T}}^2} \frac{-\mathbf{k_{1T}} \cdot \mathbf{k_{2T}}}{m_H^2} V. \tag{5}$$

Using $d^3\mathbf{q_1}' d^3\mathbf{q_2}' d^3\mathbf{q_H} \delta^{(4)}(q_1 + q_2 - q_1' - q_2' - q_H) = d^2\mathbf{q_{1T}}' d^2\mathbf{q_{2T}}' dy\ E_H$ (y is the rapidity of the Higgs) the cross-section is therefore

$$\frac{d\sigma}{d^2\mathbf{q_{1T}}' d^2\mathbf{q_{2T}}' dy} \approx \left(\frac{N_c^2-1}{N_c^2}\right)^2 \frac{\alpha_s^6}{(2\pi)^5} \frac{G_F}{\sqrt{2}} \left[\int \frac{d^2\mathbf{Q_T}}{2\pi} \frac{\mathbf{k_{1T}} \cdot \mathbf{k_{2T}}}{\mathbf{Q_T}^2 \mathbf{k_{1T}}^2 \mathbf{k_{2T}}^2} \frac{2}{3}\right]^2 \tag{6}$$

and for simplicity here we have taken the large top mass limit of V (i.e. $m_t \gg m_H$). We are mainly interested in the forward scattering limit whence

$$\frac{\mathbf{k_{1T}} \cdot \mathbf{k_{2T}}}{\mathbf{Q_T}^2 \mathbf{k_{1T}}^2 \mathbf{k_{2T}}^2} \approx -\frac{1}{\mathbf{Q_T}^4}.$$

As it stands, the integral over Q_T diverges. Let us not worry about that for now and instead turn our attention to how to convert this parton level cross-section into the hadron level cross-section we need.[2]

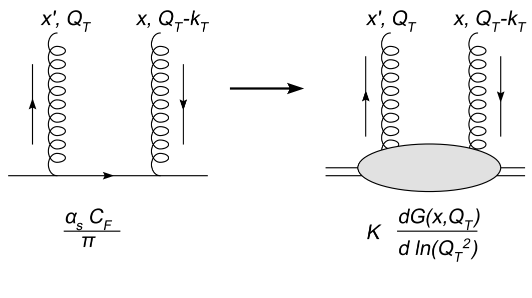


Fig. 2: The recipe for replacing the quark line (left) by a proton line (right).

What we really want is the hadronic matrix element which represents the coupling of two gluons into a proton, and this is really an off-diagonal parton distribution function [23]. At present we don't have

[1] We can do this because $x_i \sim m_H/\sqrt{s}$ whilst the other Sudakov components are $\sim Q_T^2/s$.

[2] We note that (6) was first derived by Bialas and Landshoff, except that they made a factor of 2 error in the Higgs width to gluons.

much knowledge of these distributions, however we do know the diagonal gluon distribution function. Figure 2 illustrates the Durham prescription for coupling the two gluons into a proton rather than a quark. The factor K would equal unity if $x' = x$ and $k_T = 0$ which is the diagonal limit. That we should, in the amplitude, replace a factor of $\alpha_s C_F/\pi$ by $\partial G(x, Q_T)/\partial \ln Q_T^2$ can be easily derived starting from the DGLAP equation for evolution off an initial quark distribution given by $q(x) = \delta(1-x)$. The Durham approach makes use of a result derived in [24] which states that in the case $x' \ll x$ and $k_T^2 \ll Q_T^2$ the off-diagonality can be approximated by a multiplicative factor, K. Assuming a Gaussian form factor suppression for the k_T-dependence they estimate that

$$K \approx e^{-bk_T^2/2} \frac{2^{2\lambda+3}}{\sqrt{\pi}} \frac{\Gamma(\lambda + 5/2)}{\Gamma(\lambda + 4)} \tag{7}$$

and this result is obtained assuming a simple power-law behaviour of the gluon density, i.e. $G(x, Q) \sim x^{-\lambda}$. For the production of a 120 GeV Higgs boson at the LHC, $K \sim 1.2 \times e^{-bk_T^2/2}$. In the cross-section, the off-diagonality therefore provides an enhancement of $(1.2)^4 \approx 2$. Clearly the current lack of knowledge of the off-diagonal gluon is one source of uncertainty in the calculation. We also do not really know what to take for the slope parameter b. It should perhaps have some dependence upon Q_T and for $Q_T \sim 1.5$ GeV, which it will turn out is typical for a 120 GeV scalar Higgs, one might anticipate the same k_T-dependence as for diffractive J/ψ production which is well measured, i.e. $b \approx 4$ GeV^{-2}.

Thus, after integrating over the transverse momenta of the scattered protons we have

$$\frac{d\sigma}{dy} \approx \frac{1}{256\pi b^2} \frac{\alpha_s G_F \sqrt{2}}{9} \left[\int \frac{d^2\mathbf{Q_T}}{\mathbf{Q_T}^4} f(x_1, Q_T) f(x_2, Q_T) \right]^2 \tag{8}$$

where $f(x, Q) \equiv \partial G(x, Q)/\partial \ln Q^2$ and we have neglected the exchanged transverse momentum in the integrand. Notice that in determining the total rate we have introduced uncertainty in the normalisation arising from our lack of knowledge of b. This uncertainty, as we shall soon see, is somewhat diminished as the result of a similar b-dependence in the gap survival factor.

Now it is time to worry about the fact that our integral diverges in the infra-red. Fortunately we have missed some crucial physics. The lowest order diagram is not enough, virtual graphs possess logarithms in the ratio Q_T/m_H which are very important as $Q_T \to 0$; these logarithms need to be summed to all orders. This is Sudakov physics: thinking in terms of real emissions we must be sure to forbid real emissions into the final state. Let's worry about real gluon emission off the two gluons which fuse to make the Higgs. The emission probability for a single gluon is (assuming for the moment a fixed coupling α_s)

$$\frac{C_A \alpha_s}{\pi} \int_{Q_T^2}^{m_H^2/4} \frac{dp_T^2}{p_T^2} \int_{p_T}^{m_H/2} \frac{dE}{E} \sim \frac{C_A \alpha_s}{4\pi} \ln^2 \left(\frac{m_H^2}{Q_T^2} \right).$$

The integration limits are kinematic except for the lower limit on the p_T integral. The fact that emissions below Q_T are forbidden arises because the gluon not involved in producing the Higgs completely screens the colour charge of the fusing gluons if the wavelength of the emitted radiation is long enough, i.e. if $p_T < Q_T$. Now we see how this helps us solve our infra-red problem: as $Q_T \to 0$ so the screening gluon fails to screen and real emission off the fusing gluons cannot be suppressed. To see this argument through to its conclusion we realise that multiple real emissions exponentiate and so we can write the non-emission probability as

$$e^{-S} = \exp \left(-\frac{C_A \alpha_s}{\pi} \int_{Q_T^2}^{m_H^2/4} \frac{dp_T^2}{p_T^2} \int_{p_T}^{m_H/2} \frac{dE}{E} \right). \tag{9}$$

As $Q_T \to 0$ the exponent diverges and the non-emission probability vanishes faster than any power of Q_T. In this way our integral over Q_T becomes

$$\int \frac{dQ_T^2}{Q_T^4} f(x_1, Q_T) f(x_2, Q_T)\, e^{-S} \tag{10}$$

which is finite.

There are two loose ends to sort out before moving on. Firstly, note that emission off the screening gluon is less important since there are no associated logarithms in m_H/Q_T. Secondly, (9) is correct only so far as the leading double logarithms. It is of considerable practical importance to correctly include also the single logarithms. To do this we must re-instate the running of α_s and allow for the possibility that quarks can be emitted. Including this physics means we ought to use

$$e^{-S} = \exp\left(-\int_{Q_T^2}^{m_H^2/4} \frac{dp_T^2}{p_T^2}\frac{\alpha_s(p_T^2)}{2\pi}\int_0^{1-\Delta} dz\,\left[zP_{gg}(z) + \sum_q P_{qg}(z)\right]\right) \tag{11}$$

where $\Delta = 2p_T/m_H$, and $P_{gg}(z)$ and $P_{qg}(z)$ are the leading order DGLAP splitting functions. To correctly sum all single logarithms requires some care in that what we want is the distribution of gluons in Q_T with no emission up to m_H, and this is in fact [25]

$$\tilde{f}(x, Q_T) = \frac{\partial}{\partial \ln Q_T^2}\left(e^{-S/2}\, G(x, Q_T)\right).$$

The integral over Q_T is therefore

$$\int \frac{dQ_T^2}{Q_T^4}\tilde{f}(x_1, Q_T)\tilde{f}(x_2, Q_T) \tag{12}$$

which reduces to (10) in the double logarithmic approximation where the differentiation of the Sudakov factor is subleading.

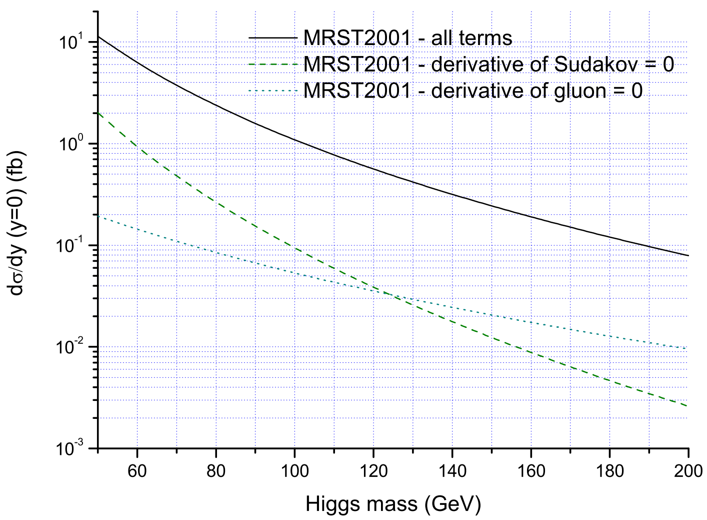


Fig. 3: The Higgs cross-section at zero rapidity, and the result obtained if one were to assume that $\partial G(x, Q)/\partial Q = 0$ or that $\partial S/\partial Q = 0$.

The numerical effect of correctly including the single logarithms is large. For production of a 120 GeV Higgs at the LHC, there is a factor ~ 30 enhancement compared to the double logarithmic approximation, with a large part of this coming from terms involving the derivative of the Sudakov. Figure 3 shows just how important it is to keep those single logarithmic terms coming from differentiation of the Sudakov factor. For the numerical results we used the MRST2001 leading order gluon [26], as included in LHAPDF [27]. Here and elsewhere (unless otherwise stated), we use a NLO QCD K-factor of 1.5 and the one-loop running coupling with $n_f = 4$ and $\Lambda_{\rm QCD} = 160$ MeV. As discussed in the next paragraph, we also formally need an infra-red cut-off Q_0 for the Q_T-integral; we take $Q_0 = 0.3$ GeV although as we shall see results are insensitive to Q_0 provided it is small enough. Finally, all our results include an overall multiplicative "gap survival factor" of 3% (gap survival is discussed shortly).

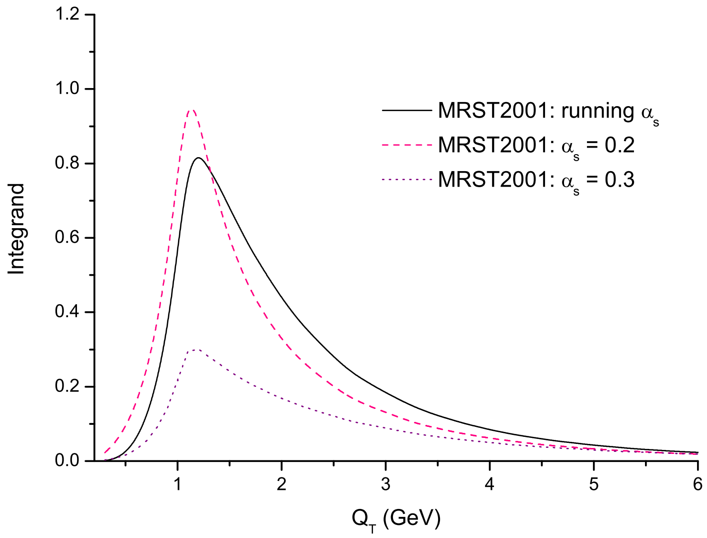


Fig. 4: The integrand of the Q_T integral for three different treatments of α_s and $m_H = 120$ GeV.

Formally there is the problem of the pole in the QCD coupling at $p_T = \Lambda_{\rm QCD}$. However, this problem can be side-stepped if the screening gluon has "done its job" sufficiently well and rendered an integrand which is peaked at $Q_T \gg \Lambda_{\rm QCD}$ since an infra-red cutoff on p_T can then safely be introduced. We must be careful to check whether or not this is the case in processes of interest. Indeed, a saddle point estimate of (10) reveals that

$$\exp(\langle \ln Q_T \rangle) \sim \frac{m_H}{2} \exp\left(-\frac{c}{\alpha_s}\right) \tag{13}$$

where c is a constant if the gluon density goes like a power of Q_T^2. Clearly there is a tension between the Higgs mass, which encourages a large value of the loop momentum, and the singular behaviour of the $1/Q_T^4$ factor which encourages a low value. Also, as α_s reduces so real emission is less likely and the Sudakov suppression is less effective in steering Q_T away from the infra-red. Putting in the numbers one estimates that $\exp(\langle \ln Q_T^2 \rangle) \approx 4$ GeV^2 for the production of a 120 GeV scalar at the LHC which is just about large enough to permit an analysis using perturbative QCD. Figure 4 provides the quantitative support for these statements in the case of a Higgs of mass 120 GeV. The integrand of the Q_T integral in equation (12) is shown for both running and fixed α_s. We see that the integrand peaks just above 1 GeV and that the Sudakov factor becomes increasingly effective in suppressing the cross-section as α_s increases. Although it isn't too easy to see on this plot, the peak does move to higher values of Q_T as α_s increases in accord with (13). This plot also illustrates quite nicely that the cross-section is pretty much insensitive to the infra-red cutoff for $Q_0 < 1$ GeV and this is made explicit in Figure 5.

Discussion of the infra-red sensitivity would not be complete without returning to the issue of the unintegrated gluon density. In all our calculations we model the off-diagonality as discussed below equation (7) and we shan't discuss this source of uncertainty any further here.[3] Figure 6 shows the gluon

[3] We actually assume a constant enhancement factor of 1.2 per gluon density.

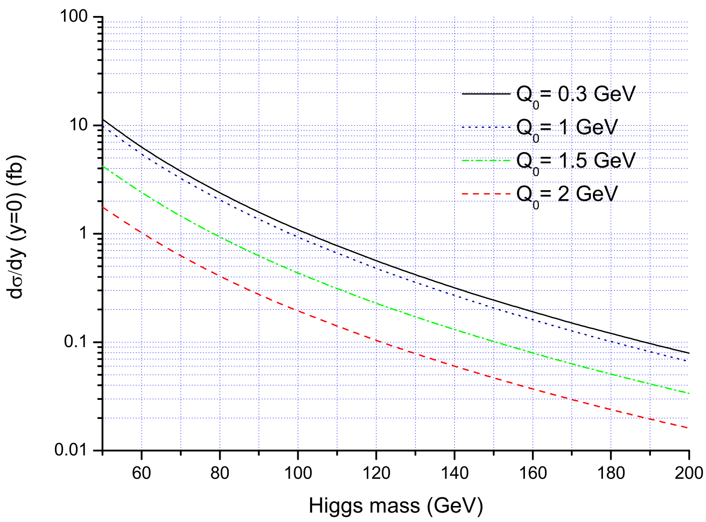


Fig. 5: The Higgs cross-section dependence upon the infra-red cutoff Q_0.

density $G(x, Q)$ as determined in four recent global fits (rather arbitrarily chosen to illustrate the typical variety) [26, 28, 29, 30]. Apart from the Fermi2002 fit, they are all leading order fits. Now, none of these parameterisations go down below $Q = 1$ GeV, so what is shown in the figure are the gluons extrapolated down to $Q = 0$. We have extrapolated down assuming that the gluon and its derivative are continuous at $Q = 1$ GeV and that $G(x, Q) \sim Q^2$ at $Q \to 0$.[4] The gluons plotted in Figure 6 are all determined at $x = 0.01$ which would be the value probed in the production of a 120 GeV Higgs at $y = 0$ at the LHC. The key point is to note that it is hard to think of any reasonable parameterisation of the gluon below 1 GeV which could give a substantial contribution to the cross-section. The Sudakov factor suppresses the low Q^2 region and also the size of the gluon and its derivative are crucial, and one cannot keep both of these large for $Q < 1$ GeV. Figure 7 shows the integrand of the Q_T integral for different fits to the gluon. In all cases the contribution below 1 GeV is small, although there are clearly important uncertainties in the cross-section. These uncertainties are better seen in Figure 8 which illustrates that one might anticipate a factor of a few uncertainty from this source.

We note that although a variety of parameterizations are presented in Figure 8 the way that the actual Q_T dependence of the integrand is obtained is the same in each case. In [31, 32] the uncertainties arising from the way the unintegrated parton densities are obtained from the integrated ones are examined. Here we have followed the prescription presented in [33] which amounts to performing one backward step in a DGLAP parton shower. However, it is known that such showers tend to underestimate the hardness of, for example, the W/Z $p_\perp$ spectra in hadron colliders unless a large intrinsic transverse momentum is added to the perturbative $k_\perp$ distribution of the colliding partons [34, 35]. In [32] it was shown that adding such an intrinsic transverse momentum would harden the Q_T distribution of the integrand in (12) for small Q_T which in turn lowers the cross-section for central exclusive Higgs production by a factor 2 (for a Gaussian intrinsic transverse momentum with $\langle k_\perp^2 \rangle = 2$ GeV^2). Investigations into how one could use unintegrated gluon densities obtained by CCFM [36] and LDC [37] evolution for central exclusive Higgs production have also been performed [32]. However, as discussed in more detail elsewhere in these proceedings [23], the available parameterizations, which are all fitted to HERA data only, are not constrained enough to allow for reliable predictions for Higgs production at the LHC.

[4] To be precise we extrapolate assuming $G(x, Q) \sim Q^{2+(\gamma-2)Q}$.

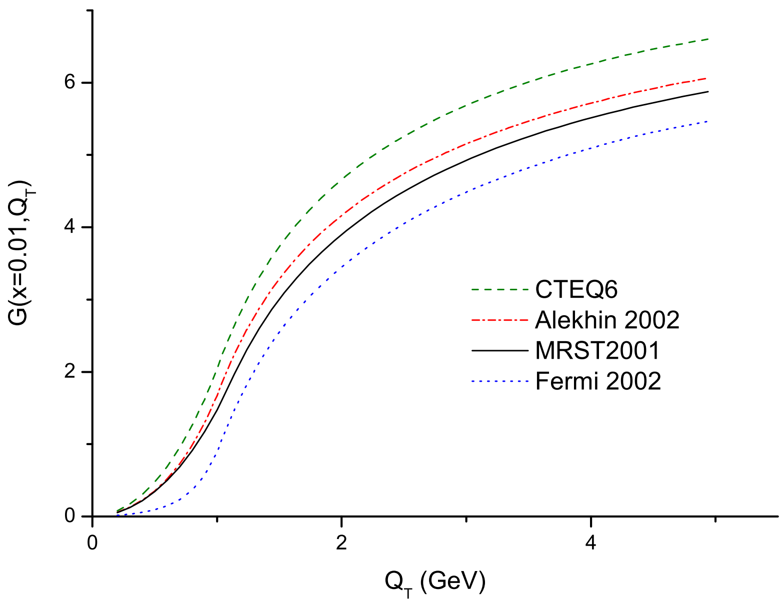


Fig. 6: The gluon density function in four different parameterisations.

This is perhaps a good place to mention pseudo-scalar production, as might occur in an extension to the Standard Model. The scalar product, $\mathbf{k_{1T}} \cdot \mathbf{k_{2T}}$, in (6) now becomes $(\mathbf{k_{1T}} \times \mathbf{k_{2T}}) \cdot \mathbf{n}$, where $\mathbf{n}$ is a unit vector along the beam axis. After performing the angular integral the only surviving terms are proportional to the vector product of the outgoing proton transverse momenta, i.e. $\mathbf{q_1}' \times \mathbf{q_2}'$. Notice that this term vanishes, in accord with the spin-0 selection rule, as $\mathbf{q_i}' \to 0$. Notice also that the integrand now goes like $\sim 1/Q_T^6$ (in contrast to the $1/Q_T^4$ in the scalar case). As a result c in (13) is larger (in fact it is linearly proportional to the power of Q_T) and the mean value of Q_T smaller. This typically means that pseudo-scalar production is not really accessible to a perturbative analysis.

The Sudakov factor has allowed us to ensure that the exclusive nature of the final state is not spoilt by perturbative emission off the hard process. What about non-perturbative particle production? The protons can in principle interact quite apart from the perturbative process discussed hitherto and this interaction could well lead to the production of additional particles. We need to account for the probability that such emission does not occur. Provided the hard process leading to the production of the Higgs occurs on a short enough timescale, we might suppose that the physics which generates extra particle production factorizes and that its effect can be accounted for via an overall factor multiplying the cross-section we have just calculated. This is the “gap survival factor”. Gap survival is discussed in detail elsewhere in these proceedings and so we’ll not dwell on it here [38].

The gap survival, S^2, is given by

$$d\sigma(p + H + p|\text{no soft emission}) = d\sigma(p + H + p) \times S^2$$

where $d\sigma(p+H+p)$ is the differential cross-section computed above. The task is to estimate S^2. Clearly this is not straightforward since we cannot utilize QCD perturbation theory. Let us at this stage remark that data on a variety of processes observed at HERA, the Tevatron and the LHC can help us improve our understanding of “gap survival”.

The model presented here provides a good starting point for understanding the more sophisticated treatments [39, 40, 41]. Dynamically, one expects that the likelihood of extra particle production will be greater if the incoming protons collide at small transverse separation compared to collisions at larger separations. The simplest model which is capable of capturing this feature is one which additionally

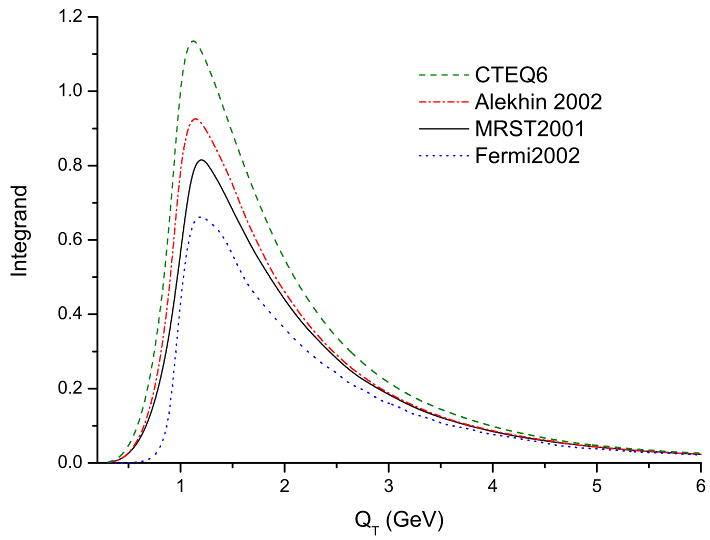


Fig. 7: The integrand of the Q_T integral for four recent global fits to the gluon.

assumes that there is a single soft particle production mechanism, let us call it a "re-scattering event", and that re-scattering events are independent of each other for a collision between two protons at transverse separation r. In such a model we can use Poisson statistics to model the distribution in the number of re-scattering events per proton-proton interaction:

$$P_n(r) = \frac{\chi(r)^n}{n!} \exp(-\chi(r)) \; . \tag{14}$$

This is the probability of having n re-scattering events where $\chi(r)$ is the mean number of such events for proton-proton collisions at transverse separation r. Clearly the important dynamics resides in $\chi(r)$; we expect it to fall monatonically as r increases and that it should be much smaller than unity for r much greater than the QCD radius of the proton. Let us for the moment assume we know $\chi(r)$, then we can determine S^2 via

$$S^2 = \frac{\int dr \;\; d\sigma(r) \;\; \exp(-\chi(r))}{\int dr \;\; d\sigma(r)} \tag{15}$$

where $d\sigma(r)$ is the cross-section for the hard process that produces the Higgs expressed in terms of the transverse separation of the protons. Everything except the r dependence of $d\sigma$ cancels when computing S^2 and so we need focus only on the dependence of the hard process on the transverse momenta of the scattered protons ($\mathbf{q_i}'$), these being Fourier conjugate to the transverse position of the protons, i.e.

$$\begin{aligned} d\sigma(r) \quad &\propto \quad [(\int d^2\mathbf{q_1}' \, e^{i\mathbf{q_1}'\cdot\mathbf{r}/2} \;\exp(-b\mathbf{q_1}'^2/2)) \times (\int d^2\mathbf{q_2}' \, e^{-i\mathbf{q_2}'\cdot\mathbf{r}/2} \;\exp(-b\mathbf{q_2}'^2/2))]^2 \\ &\propto \quad \exp\left(-\frac{r^2}{2b}\right) \; . \end{aligned} \tag{16}$$

Notice that since the b here is the same as that which enters into the denominator of the expression for the total rate there is the aforementioned reduced sensitivity to b since as b decreases so does S^2 (since the collisions are necessarily more central) and what matters is the ratio S^2/b^2.

It remains for us to determine the mean multiplicity $\chi(r)$. If there really is only one type of re-scattering event[5] independent of the hard scattering, then the inelastic scattering cross-section can be

[5]Clearly this is not actually the case, but such a "single channel eikonal" model has the benefit of being simple.

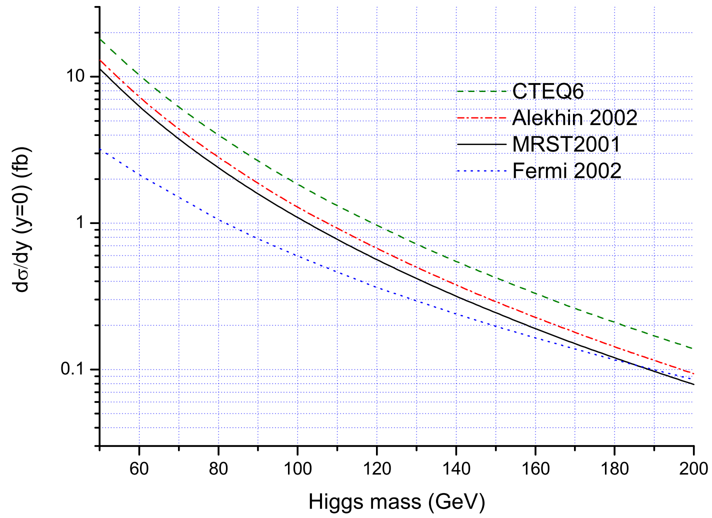


Fig. 8: The Higgs cross-section for four recent global fits to the gluon.

written

$$\sigma_{\text{inelastic}} = \int d^2\mathbf{r}(1 - \exp(-\chi(r))), \tag{17}$$

from which it follows that the elastic and total cross-sections are

$$\sigma_{\text{elastic}} = \int d^2\mathbf{r}(1 - \exp(-\chi(r)/2))^2, \tag{18}$$

$$\sigma_{\text{total}} = 2\int d^2\mathbf{r}(1 - \exp(-\chi(r)/2)). \tag{19}$$

There is an abundance of data which we can use to test this model and we can proceed to perform a parametric fit to $\chi(r)$. This is essentially what is done in the literature, sometimes going beyond a single-channel approach. Suffice to say that this simple approach works rather well. Moreover, it also underpins the models of the underlying event currently implemented in the PYTHIA [42] and HERWIG [43, 44] Monte Carlo event generators which have so far been quite successful in describing many of the features of the underlying event [45, 46, 47]. Typically, models of gap survival predict S^2 of a few percent at the LHC. Although data support the existing models of gap survival there is considerable room for improvement in testing them further and in so doing gaining greater control of what is perhaps the major theoretical uncertainty in the computation of exclusive Higgs production. In all our plots we took $S^2 = 3\%$ which is typical of the estimates in the literature for Higgs production ath the LHC.

3 Other Models

We'll focus in this section mainly on the model presented by what we shall call the Saclay group [13]. The model is a direct implementation of the original Bialas-Landshoff (BL) calculation [20] supplemented with a gap survival factor. It must be emphasised that BL did not claim to have computed for an exclusive process, indeed they were careful to state that "additional...interactions...will generate extra particles...Thus our calculation really is an inclusive one".

Equation (6) is the last equation that is common to both models. BL account for the coupling to the proton in a very simple manner: they multiply the quark level amplitude by a factor of 9 (which corresponds to assuming that there are three quarks in each proton that are able to scatter off each other).

Exactly like the Durham group they also include a form factor suppression factor $\exp(-bq_{iT}'^2)$ for each proton at the cross-section level with $b = 4$ GeV^{-2}. Since BL are not interested in suppressing radiation, they do have a problem with the infra-red since there is no Sudakov factor. They dealt with this by following the earlier efforts of Landshoff and Nachtmann (LN) in replacing the perturbative gluon propagators with non-perturbative ones [48, 49]:

$$\frac{g^2}{k^2} \to A \exp(-k^2/\mu^2).$$

Rather arbitrarily, $g^2 = 4\pi$ was assumed, except for the coupling of the gluons to the top quark loop, where $\alpha_s = 0.1$ was used.

Following LN, μ and A are determined by assuming that the $p\bar{p}$ elastic scattering cross-section at high energy can be approximated by the exchange of two of these non-perturbative gluons between the 3×3 constituent quarks: the imaginary part of this amplitude determines the total cross-section for which there are data which can be fitted to. In order to carry out this procedure successfully, one needs to recognize that a two-gluon exchange model is never going to yield the gentle rise with increasing centre-of-mass energy characteristic of the total cross-section. BL therefore also include an additional "reggeization" factor of $s^{\alpha(t)-1}$ in the elastic scattering amplitude where

$$\alpha(t) = 1 + \epsilon + \alpha' \, t$$

is the pomeron trajectory which ensures that a good fit to total cross-section data is possible for $\epsilon = 0.08$ and $\alpha' = 0.25$ GeV^{-2}. In this way the two-gluon system is modelling pomeron exchange. They found that $\mu \approx 1$ GeV and $A \approx 30$ GeV^{-2} gave a good fit to the data. Similarly, the amplitude for central Higgs production picks up two reggeization factors.

The inclusive production of a Higgs boson in association with two final state protons is clearly much more infra-red sensitive than the exclusive case where the Sudakov factor saves the day. Having said that, the Saclay model does not include the Sudakov suppression factor. Instead it relies upon the behaviour of the non-perturative gluon propagators to render the Q_T integral finite. As a result, the typical Q_T is much smaller than in the Durham case. Indeed it may be sufficiently small to make the approximation $Q_T^2 \gg q_{iT}'^2$ invalid which means that the spin-0 selection rule is no longer applicable.

Pulling everything together, the Saclay model of the cross-section for $pp \to p + H + p$ gives

$$\begin{aligned} \frac{d\sigma}{d^2\mathbf{q_{1T}}' d^2\mathbf{q_{2T}}' dy} \approx\; & S^2 \left(\frac{N_c^2 - 1}{N_c^2}\right)^2 \frac{\alpha_s^2}{(2\pi)^5} \left(\frac{g^2}{4\pi}\right)^4 \frac{G_F}{\sqrt{2}} e^{-bq_{1T}'^2} e^{-bq_{2T}'^2} \\ & x_1^{2-2\alpha(q_{1T}'^2)} x_2^{2-2\alpha(q_{2T}'^2)} \left[9 \int \frac{d^2\mathbf{Q_T}}{2\pi} \mathbf{Q_T}^2 \left(\frac{A}{g^2}\right)^3 \exp(-3\mathbf{Q_T}^2/\mu^2) \frac{2}{3} \right]^2 \end{aligned} \quad (20)$$

The reggeization factors depend upon the momentum fractions x_1 and x_2 which satisfy $x_1 x_2 s = m_H^2$ and $y = \frac{1}{2} \ln(x_1/x_2)$. The only difference[6] between this and the original BL result is the factor of S^2. Integrating over the final state transverse momenta and simplifying a little gives

$$\frac{d\sigma}{dy} \approx S^2 \frac{\pi}{b + 2\alpha' \ln(1/x_1)} \frac{\pi}{b + 2\alpha' \ln(1/x_2)} \left(\frac{N_c^2 - 1}{N_c^2}\right)^2 \frac{G_F}{\sqrt{2}} \frac{\alpha_s^2}{(2\pi)^5} \frac{1}{(4\pi)^4} \left(\frac{s}{m_H^2}\right)^{2\epsilon} \frac{1}{g^4} \left[\frac{A^3 \mu^4}{3}\right]^2 . \quad (21)$$

Figure 9 shows how the Saclay model typically predicts a rather larger cross-section with a weaker dependence upon m_H than the Durham model. The weaker dependence upon m_H arises because the Saclay model does not contain the Sudakov suppression, which is more pronounced at larger m_H, and also because of the choice $\epsilon = 0.08$. A larger value would induce a correspondingly more rapid fall. The

[6] Apart from the factor 2 error previously mentioned.

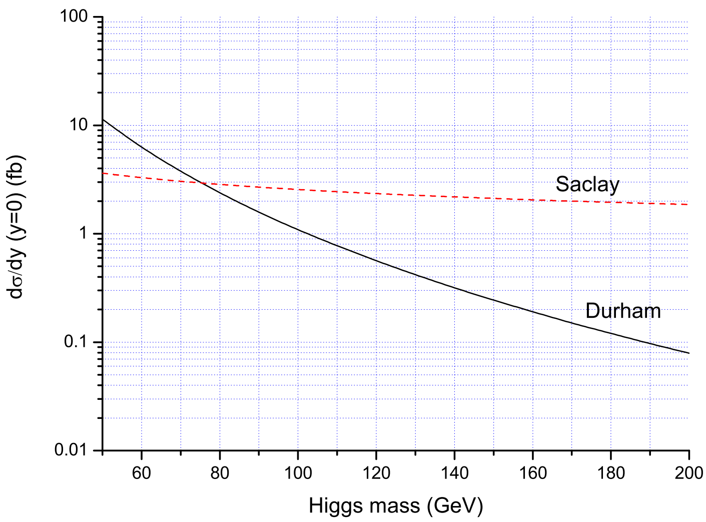


Fig. 9: Comparing dependence upon m_H of the Saclay and Durham predictions. $S^2 = 3\%$ in both cases.

Durham use of the gluon density function does indeed translate into an effective value of ϵ subtantially larger than 0.08. This effect is also to be seen in the dependence of the model predictions upon the centre-of-mass energy as shown in Figure 10. We have once again assumed a constant $S^2 = 3\%$ in this figure despite the fact that one does expect a dependence of the gap survival factor upon the energy.

Figure 11 compares the rapidity dependence of the Higgs production cross-section in the two models. The Saclay prediction is almost y-independent. Indeed the only y-dependence is a consequence of $\alpha' \neq 0$. In both models the calculations are really only meant to be used for centrally produced Higgs bosons, i.e. $|y|$ not too large since otherwise one ought to revisit the approximations implicit in taking the high-energy limit. Nevertheless, the Durham prediction does anticipate a fall as $|y|$ increases, and this is coming because one is probing larger values of x in the gluon density. In contrast, the Saclay prediction does not anticipate this fall and so a cutoff in rapidity needs to be introduced in quoting any cross-section integrated over rapidity. In Figure 11 a cut on $x_{1,2} < 0.1$ is made (which is equivalent to a cut on $|y| < 2.5$) for the Saclay model. After integrating over rapidity, the Durham model predicts a total cross-section of 2 fb for the production of a 120 GeV Higgs boson at the LHC whilst the Saclay model anticipates a cross-section a factor ~ 5 larger.

The essentially non-perturbative Saclay prediction clearly has some very substantial uncertainties associated with it. The choice of an exponentially falling gluon propagator means that there is no place for a perturbative component. However, as the Durham calculation shows, there does not seem to be any good reason for neglecting contributions from perturbatively large values of Q_T. It also seems entirely reasonable to object on the grounds that one should not neglect the Sudakov suppression factor and that including it would substantially reduce the cross-section.

In [18], the Sudakov factor of equation (11) is included, with the rest of the amplitude computed following Bialas-Landshoff. The perturbative Sudakov factor is also included in the approach of [19], albeit only at the level of the double logarithms. This latter approach uses perturbative gluons throughout the calculation but Regge factors are included to determine the coupling of the gluons into the protons, i.e. rather than the unintegrated partons of the Durham model. In both cases the perturbative Sudakov factor, not suprisingly, is important.

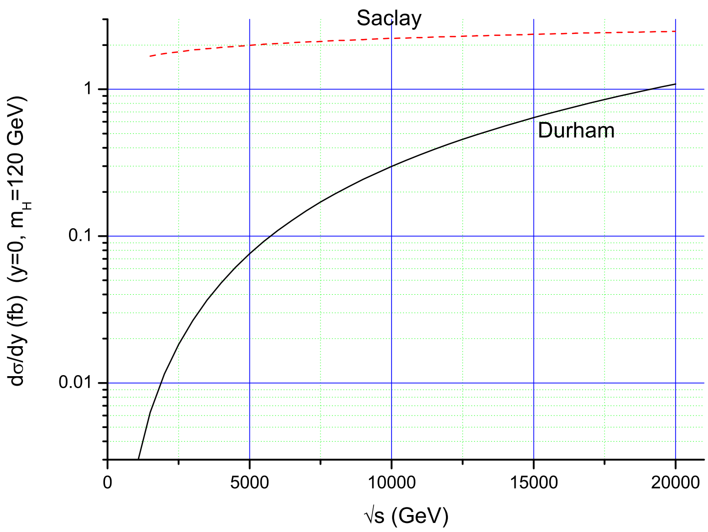


Fig. 10: Comparing dependence upon $\sqrt{s}$ of the Saclay and Durham predictions for $m_H = 120$ GeV.

4 Concluding remarks

We hope to have provided a detailed introduction to the Durham model for central exclusive Higgs production. The underlying theory has been explained and the various sources of uncertainty highlighted with particular emphasis on the sensitivity of the predictions to gluon dynamics in the infra-red region. We also made some attempt to mention other approaches which can be found in the literature.

The focus has been on the production of a Standard Model Higgs boson but it should be clear that the formalism can readily be applied to the central production of any system X which has a coupling to gluons and invariant mass much smaller than the beam energy. There are many very interesting possibilities for system X which have been explored in the literature and we have not made any attempt to explore them here [2, 3, 8, 11, 17, 16, 15]. Nor have we paid any attention to the crucial challenge of separating signal events from background [5, 9]. The inclusion of theoretical models into Monte Carlo event generators and a discussion of the experimental issues relating to central exclusive particle production have not been considered here but can be found in other contributions to these proceedings [50, 51].

It seems that perturbative QCD can be used to compute cross-sections for processes of the type $pp \to p + X + p$. The calculations are uncertain but indicate that rates ought to be high enough to be interesting at the LHC. In the case that the system X is a pair of jets there ought to be the possibility to explore this physics at the Tevatron [52]. Information gained from such an analysis would help pin down theoretical uncertainties, as would information on the rarer but cleaner channel where X is a pair of photons [53]. Of greatest interest is when X contains “new physics” whence this central exclusive production mechanism offers new possibilities for its exploration.

5 Acknowledgments

Special thanks to Hannes Jung for all his efforts in making the workshop go so well. Thanks also to Brian Cox, Markus Diehl, Valery Khoze, Peter Landshoff, Leif Lönnblad, James Monk, Leszek Motyka, Andy Pilkington and Misha Ryskin for very helpful discussions.

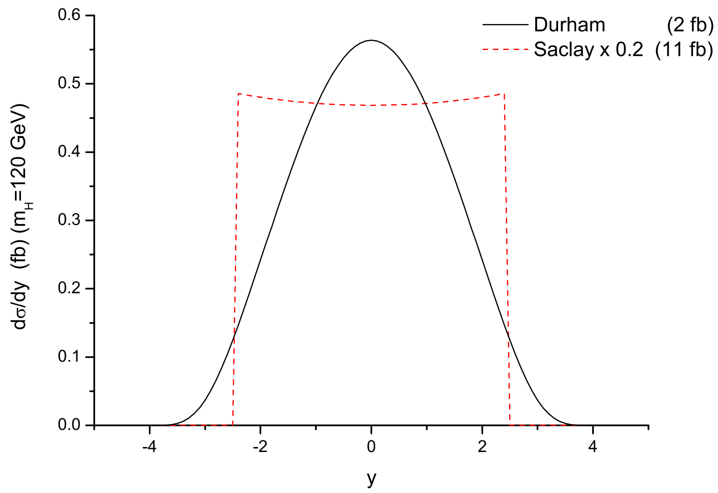


Fig. 11: Comparing the y dependence of the Saclay and Durham predictions for $m_H = 120$ GeV. Note that the Saclay prediction has been reduced by a factor 5 to make the plot easier to read. The numbers in parenthesis are the total cross-sections, i.e. integrated over rapidity.

References

[1] M.G. Albrow and A. Rostovtsev, *Searching for the Higgs at hadron colliders using the missing mass method*. E-Print Archive: hep-ph/0009336.

[2] M.G. Albrow et al, *FP420: An R&D proposal to investigate the feasibility of installing proton tagging detectors in the 420-m region at LHC*. CERN-LHCC-2005-025.

[3] B.E. Cox, AIP Conference Proceedings **753**, 103 (2005). E-Print Archive: hep-ph/0409144.

[4] J. Kalliopuska, T. Mäki, N. Marola, R. Orava, K. Österberg and M. Ottela, *Resolution studies of the leading proton measurement in exclusive central diffraction at LHC*.

[5] V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C19**, 477 (2001). Erratum-ibid. **C20**, 599 (2001). E-Print Archive: hep-ph/0011393.

[6] V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C24**, 581 (2002). E-Print Archive: hep-ph/0203122.

[7] V.A. Khoze, A.D. Martin and M.G. Ryskin, Phys. Lett. **B401**, 330 (1997). E-Print Archive: hep-ph/9701419;
V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C14**, 525 (2000). E-Print Archive: hep-ph/0002072.

[8] V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C23**, 311 (2002). E-Print Archive: hep-ph/0111078.

[9] A. De Roeck, V.A. Khoze, A.D. Martin, R. Orava and M.G. Ryskin, Eur. Phys. J. **C25**, 391 (2002). E-Print Archive: hep-ph/0207042.

[10] V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C26**, 229 (2002). E-Print Archive: hep-ph/0207313;

A.B. Kaidalov, V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C31**, 387 (2003). E-Print Archive: hep-ph/0307064.

[11] A.B. Kaidalov, V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C33**, 261 (2004). E-Print Archive: hep-ph/0311023;
V.A. Khoze, A.D. Martin and M.G.Ryskin, Eur. Phys. J. **C34**, 327 (2004). E-Print Archive: hep-ph/0401078;
K. Belotsky, V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C36**, 503 (2004). E-Print Archive: hep-ph/0406037.

[12] B.E. Cox et al, *Detecting the standard model Higgs boson in the WW decay channel using forward proton tagging at the LHC*. E-Print Archive: hep-ph/0505240.

[13] M. Boonekamp, R. Peschanski and C. Royon, Nucl. Phys. **B669**, 277 (2003). Erratum-ibid. **B676**, 493 (2004). E-Print Archive: hep-ph/0301244.

[14] M. Boonekamp, R. Peschanski and C. Royon, Phys. Lett. **B598**, 243 (2004). E-Print Archive: hep-ph/0406061;
M. Boonekamp, J. Cammin, R. Peschanski and C. Royon. E-Print Archive: hep-ph/0504199.

[15] M. Boonekamp, J. Cammin, S. Lavignac, R. Peschanski and C. Royon. E-Print Archive: hep-ph/0506275.

[16] J.R. Ellis, J.S. Lee and A. Pilaftsis, Phys. Rev. **D71**, 075007 (2005). E-Print Archive: hep-ph/0502251.

[17] B.E. Cox, J.R. Forshaw, J.S. Lee, J. Monk and A. Pilaftsis, Phys. Rev. **D68**, 075004 (2003). E-Print Archive: hep-ph/0303206.

[18] A. Bzdak, *Exclusive Higgs and dijet production by double pomeron exchange: The CDF upper limits*. E-print Archive: hep-ph/0504086.

[19] V.A. Petrov and R.A. Ryutin, Eur. Phys. J. **C36**, 509 (2004). E-Print Archive: hep-ph/0311024.

[20] A. Bialas and P.V. Landshoff, Phys. Lett. **B256**, 540 (1991).

[21] B.E. Cox and J.R. Forshaw, Comput. Phys. Commun. **144**, 104 (2002). E-Print Archive: hep-ph/0010303.

[22] M. Boonekamp and T. Kucs, Comput. Phys. Commun. **167**, 217 (2005). E-Print Archive: hep-ph/0312273.

[23] J.C. Collins et al, *Unintegrated parton density functions*. These proceedings.

[24] A.G. Shuvaev, K.J. Golec-Biernat, A.D. Martin and M.G. Ryskin, Phys. Rev. **D60**, 014015 (1999). E-Print Archive: hep-ph/9902410.

[25] M.A. Kimber, A.D. Martin, M.G. Ryskin, Phys. Rev. **D63**, 114027 (1999). E-Print Archive: hep-ph/0101348;
M.A. Kimber, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C12**, 655 (2000). E-Print Archive: hep-ph/9911379;
A.D. Martin and M.G. Ryskin, Phys. Rev. **D64**, 094017 (2001). E-Print Archive: hep-ph/0107149.

[26] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Eur. Phys. J. **C23**, 73 (2002). E-print Archive: hep-ph/0110215.

[27] `http://durpdg.dur.ac.uk/lhapdf4/`.

[28] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky and W.K. Tung, JHEP **0207**, 012 (2002). E-print Archive: hep-ph/0201195.

[29] S. Alekhin, Phys. Rev. **D68**, 014002 (2003). E-print Archive: hep-ph/0104052.

[30] W.T. Giele, S.A. Keller and D.A. Kosower. E-print Archive: hep-ph/0104052.

[31] L. Lönnblad and M. Sjödahl, JHEP **02**, 042 (2004). E-Print Archive: hep-ph/0311252.

[32] L. Lönnblad and M. Sjödahl, JHEP **05**, 038 (2005). E-Print Archive: hep-ph/0412111.

[33] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Eur. Phys. J. **C4**, 463 (1998). E-Print Archive: hep-ph/9803445.

[34] E. Thome, *Perturbative and nonperturbative effects in transverse momentum generation*. E-Print Archive: hep-ph/0401121.

[35] E.L. Nurse, *A measurement of the inclusive $z/\gamma^* \to \mu^+\mu^-$ cross section and study of W and Z events in $p\bar{p}$ collisions at D0*. FERMILAB-THESIS-2005-05.

[36] M. Ciafaloni, Nucl. Phys. **B296**, 49 (1988);
S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. **B234**, 339 (1990);
S. Catani, F. Fiorani and G. Marchesini, Nucl. Phys. **B336**, 18 (1990);
G. Marchesini, Nucl. Phys. **B445**, 49 (1995). E-Print Archive: hep-ph/9412327.

[37] B. Andersson, G. Gustafson and J. Samuelsson, Nucl. Phys. **B467**, 443 (1996);
B. Andersson, G. Gustafson and H. Kharraziha, Phys. Rev. **D57**, 5543 (1998). E-Print Archive: hep-ph/9711403.

[38] E. Gotsman et al, *Survival probabilities of large rapidity gaps*. These proceedings.

[39] E. Gotsman, E. Levin and U. Maor, Phys. Lett. **B438**, 229 (1998). E-Print Archive: hep-ph/9804404;
E. Gotsman, E. Levin and U. Maor, Phys. Lett. **B452**, 387 (1999). E-Print Archive: hep-ph/9901416;
E. Gotsman, E. Levin and U. Maor, Phys. Rev. **D60**, 094011 (1999). E-Print Archive: hep-ph/9902294;
U. Maor, *LRG production of di-jets as a probe of s-channel unitarity*. E-Print Archive: hep-ph/0406303.

[40] V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C18**, 167 (2000). E-Print Archive: hep-ph/0007359;
A.B. Kaidalov, V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C21**, 521 (2001). E-Print Archive: hep-ph/0105145.

[41] M.M. Block and F. Halzen, Phys. Rev. **D63**, 114004 (2001). E-Print Archive: hep-ph/0101022.

[42] T. Sjöstrand and M. van Zijl, Phys. Rev. **D36**, 2019 (1987).

[43] J.M. Butterworth and J.R. Forshaw, J. Phys. **G19**, 1657 (1993).

[44] J.M. Butterworth, J.R. Forshaw and M.H. Seymour, Z. Phys. **C72**, 637 (1996). E-Print Archive: hep-ph/9601371.

[45] I. Borozan and M.H. Seymour, JHEP **0209**, 015 (2002). E-Print Archive: hep-ph/0207283.

[46] T. Sjöstrand and P.Z. Skands, JHEP **0403**, 053 (2004). E-Print Archive: hep-ph/0402078.

[47] K. Odagiri, JHEP **0408**, 019 (2004). E-Print Archive: hep-ph/0407008.

[48] P.V. Landshoff and O. Nachtmann, Z. Phys. **C35**, 405 (1987).

[49] A. Donnachie and P.V. Landshoff, Nucl. Phys. **B311**, 509 (1989).

[50] M. Boonekamp et al, *Monte Carlo generators for central exclusive diffraction*. These proceedings.

[51] B.E. Cox et al, *Experimental aspects of central exclusive diffraction*. These proceedings.

[52] B.E. Cox and A. Pilkington, *Central exclusive dijet production at the Tevatron*. E-Print Archive: hep-ph/0508249.

[53] V.A. Khoze, A.D. Martin, M.G. Ryskin and W.J. Stirling, Eur. Phys. J. **C38**, 475 (2005). E-Print Archive: hep-ph/0409037.

Monte Carlo generators for central exclusive diffraction

Maarten Boonekamp, Creighton Hogg, James Monk, Andrew Pilkington & Marek Tasevsky
Particle Physics Group, School of Physics & Astronomy,
University of Manchester, Manchester, M13 9PL. United Kingdom.

Abstract
We review the three Monte Carlo generators that are available for simulating the central exclusive reaction, $pp \to p + X + p$.

1 Introduction

The central exclusive mechanism is defined as $pp \to p + X + p$ with no radiation emitted between the intact outgoing beam hadrons and the central system X. The study of central exclusive Higgs boson production has been aided with the recent development of Monte Carlo simulations to enable parton, hadron and detector level simulation. The three generators that we shall examine here are DPEMC [1], EDDE [2] and ExHuME [3]. From an experimental perspective, it is important to examine both the similarities and differences between the models in order to assess the physics potential in terms of forward proton tagging at the LHC [4].

Each of the Monte Carlos implements a different model of central exclusive production that is either perturbative or non-perturbative. ExHuME is an implementation of the perturbative calculation of Khoze, Martin and Ryskin [5], the so-called "Durham Model". In this calculation (depicted in fig 1(a)), the two gluons couple perturbatively to the off-diagonal unintegrated gluon distribution in the proton. The Durham approach includes a Sudakov factor to suppress radiation into the rapidity gap between the central system and the outgoing protons and which renders the loop diagram infra-red safe. The bare cross section is suppressed by a soft-survival probability, $\mathcal{S}^2$, that accounts for additional momentum transfer between the proton lines that lead to particle production that could fill in the gap. The current ExHuME default takes $\mathcal{S}^2$ to be 0.03 at the LHC.

In contrast, DPEMC and EDDE treat the proton vertices non-perturbatively. This is acheived in the context of Regge theory, by pomeron exchange from each of the proton lines. DPEMC follows the Bialas-Landshoff approach [6] of parameterising the pomeron flux within the proton. DPEMC also sets the default value of $\mathcal{S}^2$ to 0.03 at the LHC. EDDE uses an improved Regge-eikonal approach [7] to calculate the soft proton vertices and includes a Sudakov suppression factor to prohibit real gluon emission. There is no explicit soft-survival factor present in EDDE: it is assumed that the Regge parameterisation includes the effect of additional interactions between the proton lines. For further details of the calculations underlying both DPEMC and ExHuME please refer to [8].

The connection between the parton level process and the hadronic final state is not the same in the three Monte Carlos. Both ExHuME and EDDE are linked to Pythia [9, 10] for final state parton showering and hadronisation. DPEMC however, overrides the HERWIG [11] internal $\gamma\gamma$ interactions in e^+e^- collisions to simulate double pomeron exchange.

The processes available are similar in each Monte Carlo. Perhaps the most interesting is Higgs boson production with all subsequent decays. In addition, di-jet production is included in all three generators. None of the Monte Carlos yet includes the next-to-leading order 3 jet process, which could be an important, or even the dominant, background to the central exclusive $H \to b\bar{b}$ search channel.

Finally, inclusive double pomeron exchange (shown in figure 1(b)) will also act as a backgound to the exclusive process as there are 2 protons in the final state. These processes are always accompanied by pomeron remnants in the central system and it may be a challenge experimentally to separate these from the system of interest. Two models for these processes are the Cox-Forshaw model (CF), implemented in POMWIG [12], and the Boonekamp-Peschanski-Royon model (BPR) [13] that is included in DPEMC.

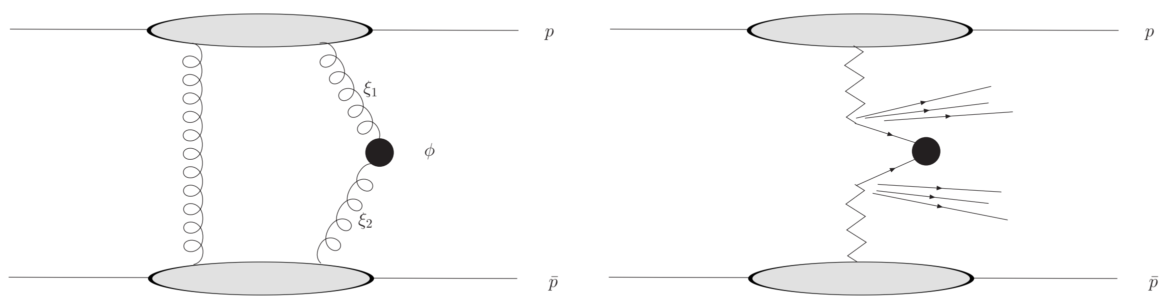


Fig. 1: The exclusive production process (a) and the inclusive (double pomeron) production process (b).

2 Results

Unless otherwise stated, all plots shown here were produced by using each of the Monte Carlos as they are distributed. In the case of ExHuME, where a parton distribution set must be chosen, the default is the 2002 MRST set, usually supplied via the LHAPDF library.

Using the default settings at the LHC energy of 14 TeV the total cross sections for production of a 120 GeV Higgs boson are 3.0 fb, 1.94 fb and 2.8 fb for DPEMC, EDDE and ExHuME respectively. However, despite these similar cross section predictions, the physics reach of the central exclusive process is predicted to differ significantly between the Monte Carlos. Figure 2(a) shows that ExHuME and EDDE predict that the cross section for exclusive Higgs boson production will fall much faster than DPEMC with an increase in Higgs boson mass. This is a direct effect of the Sudakov suppression factors growing as the available phase space for gluon emission increases with the mass of the central system. The different gluon momentum fraction, ξ, dependences lead to the differences in figure 2(b). With a fixed central mass an increase in collision energy is identical to a decrease in ξ, and the flatter ξ distributions of DPEMC and EDDE are reflected in the flatter $\sqrt{s}$ dependence compared to ExHuME.

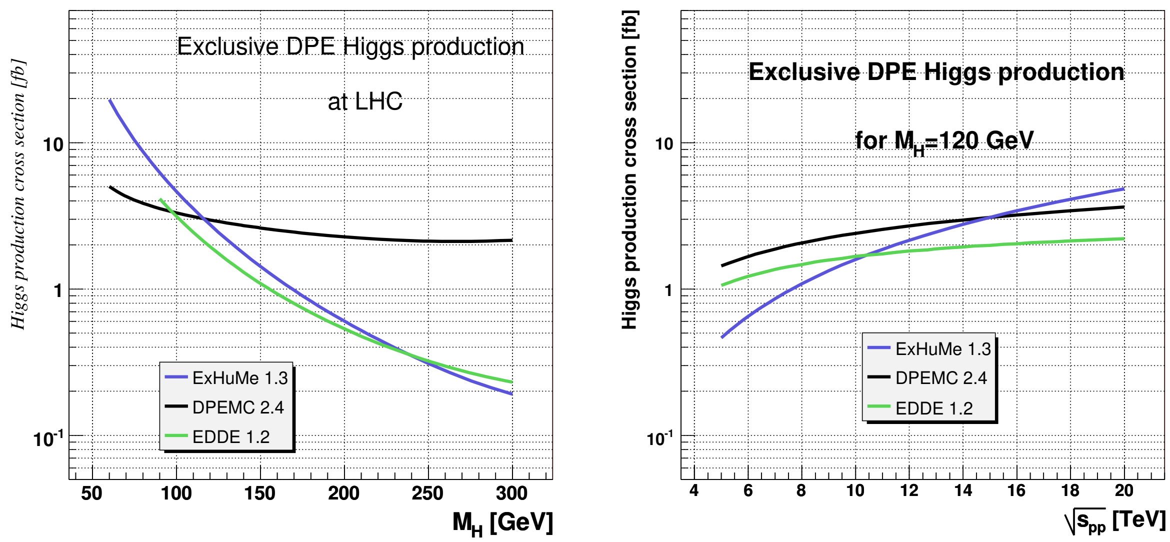


Fig. 2: (a) The left hand plot shows the higgs cross section as a function of higgs mass. (b) The right hand plot shows the increase in cross section with the collision energy (fixed gap survival factor).

The physics potential is dependent not only on the total cross section, but also on the rapidity distribution of the central system, which is shown in figure 3(b) together with the ξ distribution for the gluons. The more central rapidity distribution of ExHuME is due to the gluon distributions falling more sharply than the pomeron parameterisation present in DPEMC.

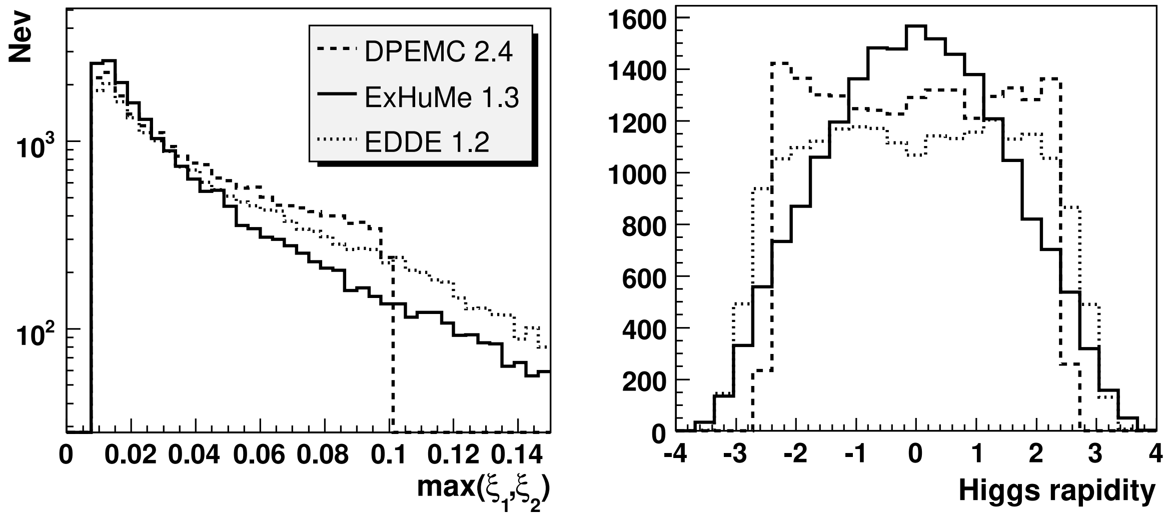


Fig. 3: The $max(\xi_1, \xi_2)$(a) and rapidity(b) of the 120 GeV Higgs. ExHuME predicts a steeper fall off in the number of events at high ξ and hence favours a less broad rapidity distribution compared to the soft non-perturbative models. Note that a cut is applied in DPEMC at $\xi = 0.1$, as required by the Bialas-Landshoff approach.

The acceptances of any forward proton taggers that might be installed at the LHC are sensitive to the rapidity distributions of the central system. The differences seen in figure 3(b) are reflected in different acceptance curves shown in figure 4. The predicted acceptances using taggers at 420 and 220 metres as a function of the mass of the central system were obtained using a fast simulation of the CMS detector. The fast simulation includes a parameterisation of the responses of the forward taggers based on a detailed simulation of the detectors [14]. As seen in figure 4, as the mass of the central system increases the combined acceptance using detectors at *both* 220 and 420 metres increases, with the relative difference between the predictions from the three generators decreasing (from about 40% down to 15% for the most extreme relative differences). For a Higgs boson of mass 120 GeV the acceptances are predicted to be 46, 50 and 57% for EDDE, DPEMC and ExHuME respectively.

Changes from the default generator settings can have an effect on all of these distributions. As an example figure 5(a) shows the rapidity distribution from ExHuME using the CTEQ6M set compared to the MRST 2002 set of parton distribution functions. The CTEQ pdf has a flatter ξ dependence in the sensitive region of $Q_\perp \simeq 3$ GeV, which leads to a broader peak and sharper fall in the rapidity distribution and a larger cross section of 3.75 fb. This in turn should improve the efficiency of the forward proton taggers because not only are there more events, but there are more events at low rapidity. It is also possible to change the DPEMC code to add a harder ξ dependence of the form $(1-\xi)^\alpha$ to the pomeron flux parameterisation. This would favour a more central rapidity distribution, thus increasing the acceptance in the forward pots.

In di-jet production the di-jet mass fraction, R_{jj} is defined as $R_{jj} = M_{jj}/\sqrt{\hat{s}}$, where M_{jj} is the mass of the di-jet system and $\sqrt{\hat{s}}$ is the total invariant mass of the central system. R_{jj} should be large in a

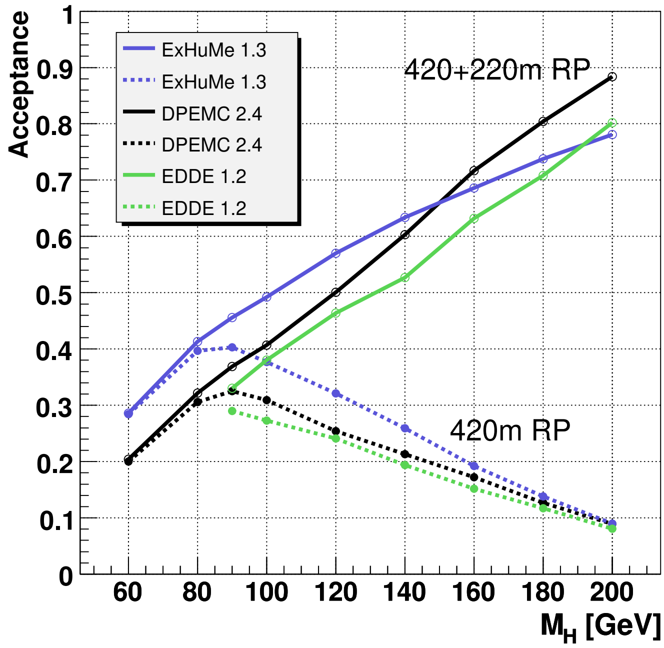


Fig. 4: The predicted acceptances for the proposed forward taggers at 420 metres and for a combination of taggers at 220 and 420 metres from the central detector.

central exclusive event. In current searches for central exclusive di-jet production at the Tevatron [15], the CDF collaboration have experimentally defined exclusive events to be those where $R_{jj} > 0.8$. It should be noted that M_{jj} depends on the particular jet reconstruction algorithm used and the $\sqrt{\hat{s}}$ measurement is dependent either on tagging the outgoing protons or on reconstructing the missing mass using the calorimeter. In figure 6(a) we show the prediction for the R_{jj} fraction in exclusive events at the LHC, whilst in figure 6(b) we show two examples of the inclusive background with pomeron remnants from Pomwig and DPEMC. It is clear that the $R_{jj} > 0.8$ definition for central exclusive production leads to an overlap between the exclusive and inclusive regions for all of the Monte Carlo predictions.

3 Summary

At a Higgs boson mass of 120GeV, all three Monte Carlo simulations give similar predictions for the cross section. However, the physics potential decreases for models that include Sudakov suppression, which will limit Higgs boson searches. The differing rapidity distributions of the central system result in different efficiencies for a forward proton tagging programme. In order to fully study the background to the $H \to b\bar{b}$ channel, future additions to the Monte Carlo programs should include the next to leading order three jet process.

Acknowledgments

We would like to thank Brian Cox, Jeff Forshaw, Valery Khoze, Misha Ryskin and Roman Ryutin for discussion throughout the proceedings.

References

[1] M. Boonekamp and T. Kucs, Comput. Phys. Commun. **167** (2005). E-Print Archive: hep-ph/0312273.

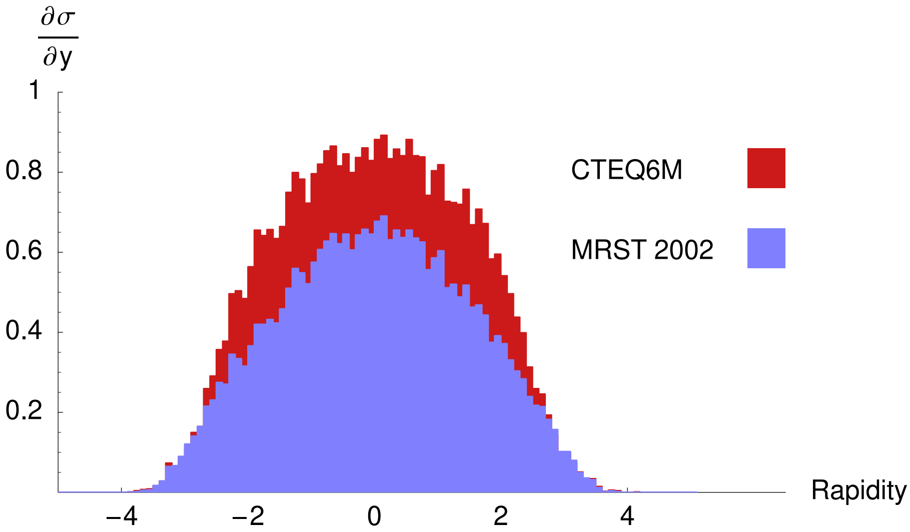


Fig. 5: The ExHuME rapidity distributions for production of a 120 GeV Higgs using MRST 2002 pdfs and the CTEQ6M pdfs.

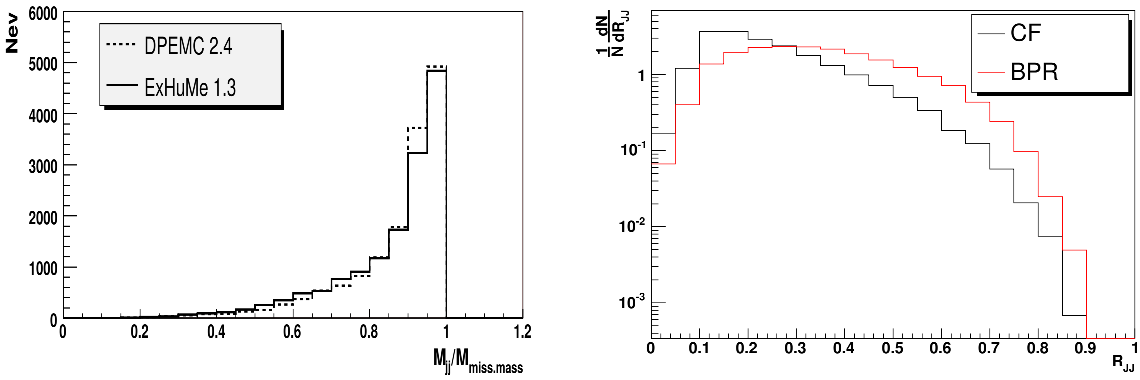


Fig. 6: R_{jj} distributions for (a) exclusive di-jet production using DPEMC and ExHuME and (b) the background inclusive di-jet production as predicted by the Cox-Forshaw (CF) and Boonekamp-Peschanski-Royon (BPR) models.

[2] R. Ryutin, *Edde monte carlo event generator*. E-Print Archive: hep-ph/0409180.

[3] J. Monk and A. Pilkington, *Exhume: A monte carlo event generator for exclusive diffraction*. E-Print Archive: hep-ph/0502077.

[4] M. G. Albrow et al., CERN-LHCC **25** (2005).

[5] V. A. Khoze, A. D. Martin and M. G. Ryskin, Eur. Phys. J **C23**, 311 (2002). E-print Archive: hep-ph/0111078.

[6] A. Bialas and P. V. Landshoff, Phys. Lett. B **B256**, 540 (1991).

[7] V. A. Petrov and R. Ryutin, Eur. Phys. Journ **C36**, 509 (2004).

[8] J. Forshaw, *Diffractive higgs production: theory*. These proceedings.

[9] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun **135**, 238 (2001). E-Print Archive: hep-ph/0010017.

[10] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands, *Pythia 6.3: Physics and manual*. E-Print Archive: hep-ph/0308153.

[11] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101**, 010 (2001).

[12] B. E. Cox and J. R. Forshaw, Comput. Phys. Commun **144**, 104 (2002). E-Print Archive: hep-ph/0010303.

[13] M. Boonekamp, R. Peschanski and C. Royon, Phys. Rev. Lett **87**, 251806 (2001).

[14] V. Avati and K. Osterberg, *Roman pot acceptances*. These proceedings.

[15] T. Affolder et al., Phys. Rev. Lett. **85**, 4215 (2000).

Diffractive Structure Functions and Diffractive PDFs

Editors: M. Arneodo[a] *and P. van Mechelen*[b]
[a]Università del Piemonte Orientale, Novara, Italy
[b]University of Antwerpen, Antwerpen, Belgium

Abstract

This section of the Proceedings contains papers summarising the current status of the F_2^D measurements at HERA, the extraction of the diffractive parton distribution functions and the relevance of a direct measurement of F_L^D.

The selection of a pure sample of inclusive diffractive events, $ep \rightarrow eXp$, is a challenging task. Three alternative approaches have been used so far by the H1 and ZEUS collaborations at HERA:

1. a fast proton in the final state is required;
2. a rapidity gap in the forward direction is required;
3. the different shape of the M_X distribution for diffractive and non-diffractive events is exploited.

The results obtained with these approaches exhibit a level of agreement which varies from tolerable to poor. This is not surprising since different final states are selected, in which the reaction $ep \rightarrow eXp$ appears with different degrees of purity. The paper by Newman and Schilling presents a systematic comparison of the results available, quantifies the differences and discusses their origins, when understood.

NLO QCD fits to the diffractive structure function F_2^D are used to extract the diffractive parton distribution functions (dPDFs) in the proton. They can be interpreted as conditional probabilities to find a parton in the proton when the final state of the process contains a fast proton of given four-momentum. They are essential to determine the cross sections of less inclusive processes in ep diffractive scattering, such as dijet or charm production. They are also a non-negotiable ingredient for the prediction of the cross sections for inclusive diffractive processes at the LHC.

Several groups have so far performed such fits to the available data. The results of these fits are presented in the papers by Newman and Schilling, Groys et al. and Watt et al. All fits give diffractive PDFs largely dominated by gluons. However, significant differences are apparent, reflecting the differences in the data, but also in the fitting procedure. Newman and Schilling and Groys et al. assume the so-called Regge factorisation hypothesis, i.e. take $F_2^D = f_{I\!P}(x_{I\!P}, t) \cdot F_2^{I\!P}(\beta, Q^2)$. This assumption has no basis in QCD and is critically discussed by Groys et al. and by Watt et al. The latter also argue that the leading-twist formula used by Newman and Schilling and by Groys et al. is inadequate in large parts of the measured kinematics, and use a modified expression which includes an estimate of power-suppressed effects.

The parametrisations of the dPDFs discussed in these three papers are available in a code library discussed in the paper by Schilling.

Finally, the paper by Newman addreses the importance of measuring the longitudinal diffractive structure function F_L^D. A measurement of F_L^D to even modest precision would provide a very powerful independent tool to verify our understanding of diffraction and to test the gluon density extracted indirectly in QCD fits from the scaling violations of F_2^D.

HERA Diffractive Structure Function Data and Parton Distributions

Paul Newman[a], *Frank-Peter Schilling*[b]
[a] School of Physics and Astronomy, University of Birmingham, B15 2TT, United Kingdom
[b] CERN/PH, CH-1211 Geneva 23, Switzerland

Abstract
Recent diffractive structure function measurements by the H1 and ZEUS experiments at HERA are reviewed. Various data sets, obtained using systematically different selection and reconstruction methods, are compared. NLO DGLAP QCD fits are performed to the most precise H1 and ZEUS data and diffractive parton densities are obtained in each case. Differences between the Q^2 dependences of the H1 and ZEUS data are reflected as differences between the diffractive gluon densities.

1 Introduction

In recent years, several new measurements of the semi-inclusive 'diffractive' deep inelastic scattering (DIS) cross section for the process $ep \to eXY$ at HERA have been released by the H1 and ZEUS experiments [1, 2, 3, 4, 5, 6]. The data are often presented in the form of a t-integrated reduced diffractive neutral current cross section $\sigma_r^{D(3)}$, defined through[1]

$$\frac{d^3\sigma^{ep\to eXY}}{\mathrm{d}x_{I\!P}\,\mathrm{d}x\,\mathrm{d}Q^2} = \frac{4\pi\alpha^2}{xQ^4}\left(1-y+\frac{y^2}{2}\right)\sigma_r^{D(3)}(x_{I\!P},x,Q^2)\ , \tag{1}$$

or in terms of a diffractive structure function $F_2^{D(3)}(x_{I\!P},\beta,Q^2)$. Neglecting any contributions from Z^0 exchange,

$$\sigma_r^{D(3)} = F_2^{D(3)} - \frac{y^2}{1+(1-y)^2}F_L^{D(3)}\ , \tag{2}$$

such that $\sigma_r^{D(3)} = F_2^{D(3)}$ is a good approximation except at very large y. The new data span a wide kinematic range, covering several orders of magnitude in Q^2, β and $x_{I\!P}$.

Within the framework of QCD hard scattering collinear factorisation in diffractive DIS [7], these data provide important constraints on the diffractive parton distribution functions (dpdf's) of the proton. These dpdf's are a crucial input for calculations of the cross sections for less inclusive diffractive processes in DIS, such as dijet or charm production [8, 9]. In contrast to the case of inclusive scattering, the dpdf's extracted in DIS are not expected to be directly applicable to hadron-hadron scattering [10, 11, 12, 7]. Indeed, diffractive factorisation breaks down spectacularly when HERA dpdf's are applied to diffractive proton-proton interactions at the TEVATRON [13]. It may, however, be possible to recover good agreement by applying an additional 'rapidity gap survival probability' factor to account for secondary scattering between the beam remnants [14, 15, 16, 17]. The HERA dpdf's thus remain an essential ingredient in the prediction of diffractive cross sections at the LHC, notably the diffractive Higgs cross section [18]. Although the poorly known rapidity gap survival probability leads to the largest uncertainty in such calculations, the uncertainty due to the input dpdf's also plays a significant role. In [3], the H1 collaboration made a first attempt to assess the uncertainty from this source, propagating the experimental errors from the data points to the 'H1 2002 NLO fit' parton densities and assessing the theoretical uncertainties from various sources.

In this contribution, we investigate the compatibility between various different measurements of F_2^D by H1 and ZEUS. We also apply the techniques developed in [3] to ZEUS data in order to explore the consequences of differences between the H1 and ZEUS measurements in terms of dpdf's.

[1] For a full definition of all terms and variables used, see for example [3].

2 Diffractive Selection Methods and Data Sets Considered

One of the biggest challenges in measuring diffractive cross sections, and often the source of large systematic uncertainties, is the separation of diffractive events in which the proton remains intact from non-diffractive events and from proton-dissociation processes in which the proton is excited to form a system with a large mass, $M_{\rm Y}$. Three distinct methods have been employed by the HERA experiments, which select diffractive events of the type $ep \to eXY$, where Y is a proton or at worst a low mass proton excitation. These methods are complimentary in that their systematics due to the rejection of proton dissociative and non diffractive contributions are almost independent of one another. They are explained in detail below.

- **Roman Pot Spectrometer Method.** Protons scattered through very small angles are detected directly in detectors housed in 'Roman Pot' insertions to the beampipe well downstream the interaction point. The proton 4-momentum at the interaction point is reconstructed from the position and slope of the tracks in these detectors, given a knowledge of the beam optics in the intervening region. The Roman Pot devices are known as the Leading Proton Spectrometer (LPS) in the case of ZEUS and the Forward Proton Spectrometer (FPS) in H1. The Roman pot method provides the cleanest separation between elastic, proton dissociative and non-diffractive events. However, acceptances are rather poor, such that statistical uncertainties are large in the data sets obtained so far.
- **Rapidity Gap Method.** This method is used by H1 for diffractive structure function measurements and by both H1 and ZEUS for the investigation of final state observables. The outgoing proton is not observed, but the diffractive nature of the event is inferred from the presence of a large gap in the rapidity distribution of the final state hadrons, separating the X system from the unobserved Y system. The diffractive kinematics are reconstructed from the mass of the X system, which is well measured in the main detector components. The rapidity gap must span the acceptance regions of various forward[2] detector components. For the H1 data presented here, these detectors efficiently identify activity in the pseudorapidity range $3.3 < \eta \lesssim 7.5$. The presence of a gap extending to such large pseudorapidities is sufficient to ensure that $M_{\rm Y} \lesssim 1.6$ GeV. In light of the poor knowledge of the $M_{\rm Y}$ spectrum at low masses, no attempt is made to correct the data for the small remaining proton dissociation contribution, but rather the cross sections are quoted integrated over $M_{\rm Y} < 1.6$ GeV.
- **$M_{\rm X}$ Method.** Again the outgoing proton is not observed, but rather than requiring a large rapidity gap, diffractive events are selected on the basis of the inclusive $\ln M_{\rm X}^2$ distribution. Diffractive events are responsible for a plateau in this distribution at low $\ln M_{\rm X}^2$, such that they can be selected cleanly for the lowest $M_{\rm X}$ values. At intermediate $M_{\rm X}$, non-diffractive contributions are subtracted on the basis of a two component fit in which the non-diffractive component rises exponentially. This method is used for diffractive structure function measurements by ZEUS. It does not discriminate between elastic and low $M_{\rm Y}$ proton-dissociative contributions. Results are quoted for $M_{\rm Y} < 2.3$ GeV.

Four recent data sets are considered, for which full details of luminosities and kinematic ranges can be found in table 1.

- Published data from ZEUS taken in 1998 and 1999, using the $M_{\rm X}$ method and taking advantage of the increased forward acceptance offered by a new plug calorimeter ('ZEUS-$M_{\rm X}$') [1].
- Published ZEUS data obtained with the LPS using data taken in 1997 ('ZEUS-LPS') [2].
- Preliminary H1 data obtained using the rapidity gap method, combining three measurements using different data sets from the period 1997-2000 for different regions in Q^2 ('H1-LRG') [3, 4, 5].
- Preliminary H1 data obtained using the FPS, based on data taken in 1999 and 2000 ('H1-FPS') [6].

[2]The forward hemisphere corresponds to that of the outgoing proton beam, where the pseudorapidity $\eta = -\ln\tan\theta/2$ is positive.

Label	Ref.	Reconstruction Method	Lumi $\mathcal{L}[\mathrm{pb}^{-1}]$	Kinematic range			
				$M_Y\,[\mathrm{GeV}]$	$Q^2[\mathrm{GeV}^2]$	β	$x_{I\!P}$
ZEUS-M_X	[1]	M_X method	4.2	< 2.3	$2.7 \ldots 55$	$0.003 \ldots 0.975$	$0.0001 \ldots 0.03$
ZEUS-LPS	[2]	Roman Pot	12.8	M_p	$2.4 \ldots 39$	$0.007 \ldots 0.48$	$0.0005 \ldots 0.06$
H1-LRG	[3, 4, 5]	Rapidity Gap	$3.4 \ldots 63$	< 1.6	$1.5 \ldots 1600$	$0.01 \ldots 0.9$	$0.0001 \ldots 0.05$
H1-FPS	[6]	Roman Pot	25	M_p	$2.6 \ldots 20$	$0.01 \ldots 0.7$	$0.002 \ldots 0.05$

Table 1: Overview of the data sets discussed here. The quoted kinematic ranges in Q^2, β and $x_{I\!P}$ correspond to the bin centres.

3 Comparisons between Data Sets

In this section, the $x_{I\!P}$ dependences of the data from the different measurements are compared at fixed values of Q^2 and β. Since the various measurements are generally presented at different Q^2 and β values, it is necessary to transport the data to the same values. The β and Q^2 values of the H1-LRG data are chosen as the reference points. The factors applied to data points from the other measurements are evaluated using two different parameterisations, corresponding to the results of QCD fits to 1994 H1 data [19] and to a subset of the present H1-LRG data at intermediate Q^2 [3] (see also section 4). In order to avoid any significant bias arising from this procedure, data points are only considered further here if the correction applied is smaller than 50% in total and if the correction factors obtained from the two parameterisations are in agreement to better than 25%. In practice, these criteria only lead to the rejection of data points in the ZEUS-M_X data set at $Q^2 = 55\ \mathrm{GeV}^2$ and $\beta = 0.975$, where the poorly known high β dependence of the diffractive cross section implies a large uncertainty on the factors required to transport them to $\beta = 0.9$. Elsewhere, there is reasonable agreement between the factors obtained from the two parameterisations and no additional uncertainties are assigned as a consequence of this procedure.

Since the various data sets correspond to different ranges in the outgoing proton system mass, M_Y, additional factors are required before comparisons can be made. For all data and fit comparisons, all data are transported to the H1 measurement range of $M_Y < 1.6$ GeV and $|t| < 1\ \mathrm{GeV}^2$. The leading proton data are scaled by a factor 1.1 [20] to correspond to the range $M_Y < 1.6$ GeV and the ZEUS-M_X data are scaled to the same range by a further factor of 0.7 [1], such that the overall factor is 0.77. The uncertainties on these factors are large, giving rise to normalisation uncertainties of perhaps 15% between the different data sets.

The ZEUS-LPS and H1-FPS data are compared in figure 1. Within the experimental uncertainties, the two data sets are in good agreement. Both data sets are also consistent with a parameterisation of the H1-LRG data [3] based on the H1 2002 NLO QCD fit, which is also shown. This good agreement between the H1-LRG and the Roman Pot data is also shown explicitly in figure 3.

In figure 2, a comparison is made between the H1-LRG and the ZEUS-M_X data after all factors have been applied. For much of the kinematic range, there is tolerable agreement between the two data sets. However, there are clear regions of disagreement. One is at the largest β (smallest M_X), where the H1 data lie significantly above the ZEUS data for $Q^2 \lesssim 20\ \mathrm{GeV}^2$. Another is at intermediate and low β, where the two data sets show significantly different dependences on Q^2. With the factor of 0.77 applied to the ZEUS data, there is good agreement at low Q^2, but the ZEUS data lie below the H1 data at large Q^2. If the factor of 0.77 is replaced with a value closer to unity, the agreement improves at large Q^2, but the H1 data lie above the ZEUS data at low Q^2. These inconsistencies between the different data sets are discussed further in section 4.

For completeness, figure 3 shows a comparison between all four data sets considered.

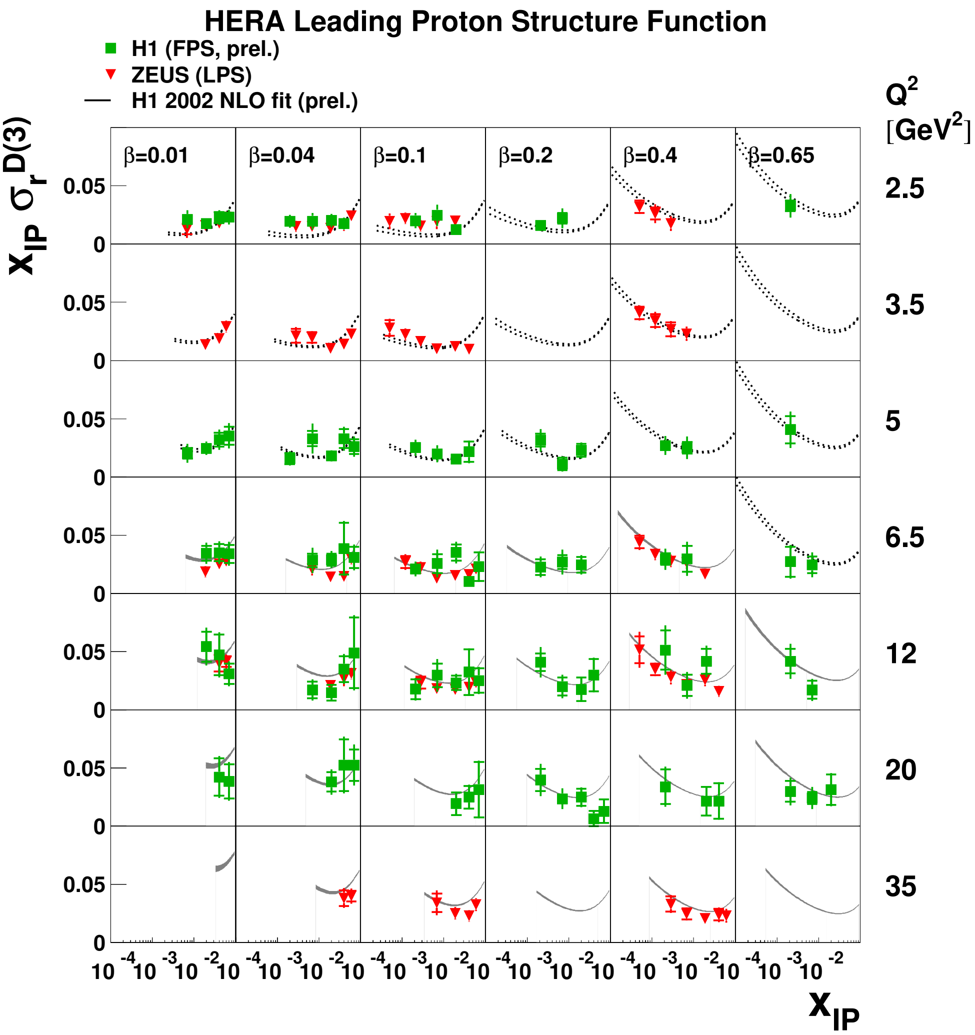


Fig. 1: Comparison of the Roman Pot data from H1 and ZEUS, scaled by a factor 1.1 such that they correspond to $M_Y < 1.6$ GeV. The Q^2 and β values have been shifted to the H1-LRG bin centres using small translation factors. The upper and lower curves form an error band on the predictions from the H1 2002 NLO QCD fit to the H1-LRG data (experimental errors only). Dotted lines are used for kinematic regions which were not included in the fit. Normalisation uncertainties of $^{+12\%}_{-10\%}$ on the ZEUS LPS data and 15% on the factor applied to shift the datasets to $M_Y < 1.6$ GeV are not shown.

4 Diffractive Parton Distributions

4.1 Theoretical Framework and Fit to H1-LRG Data

In this contribution, we adopt the fitting procedure used by H1 in [3], where next-to-leading order (NLO) QCD fits are performed to diffractive reduced cross section, $\sigma_r^{D(3)}$, data [3, 21] with $6.5 \leq Q^2 \leq 800$ GeV^2 and the β and $x_{I\!P}$ ranges given in table 1.

The proof that QCD hard scattering collinear factorisation can be applied to diffractive DIS [7] implies that in the leading $\log(Q^2)$ approximation, the cross section for the diffractive process $ep \to eXY$ can be written in terms of convolutions of universal partonic cross sections $\hat{\sigma}^{ei}$ with diffractive parton distribution functions (dpdf's) f_i^D [22, 11, 23], representing probability distributions for a parton i in the proton under the constraint that the proton is scattered with a particular 4 momentum. Thus, at

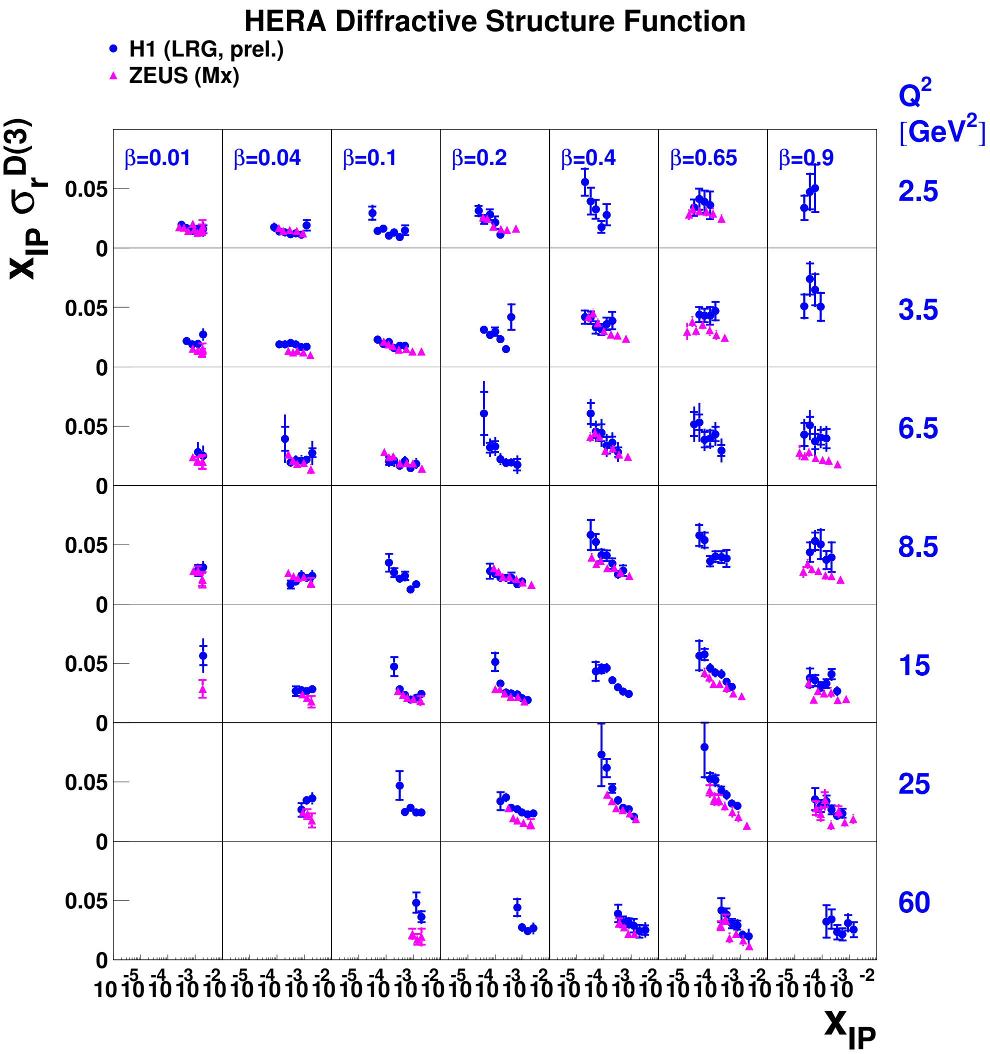


Fig. 2: Comparison of the ZEUS-M_X data with a subset of the H1-LRG data. The Q^2 and β values of the ZEUS data have been shifted to the H1 bin centres using small translation factors. The ZEUS data have also been multiplied by a universal factor of 0.77, such that both data sets correspond to $M_Y < 1.6$ GeV. Normalisation uncertainties of 15% on this factor and of $\pm 6.7\%$ on the H1 data are not shown.

leading twist,[3]

$$\frac{\mathrm{d}^2\sigma(x, Q^2, x_{I\!P}, t)^{ep\to eXp'}}{\mathrm{d}x_{I\!P}\,\mathrm{d}t} = \sum_i \int_x^{x_{I\!P}} \mathrm{d}\xi\ \hat{\sigma}^{ei}(x, Q^2, \xi)\ f_i^D(\xi, Q^2, x_{I\!P}, t)\ . \tag{3}$$

This factorisation formula is valid for sufficiently large Q^2 and fixed $x_{I\!P}$ and t. It also applies to the case of proton dissociation into a system of fixed mass M_Y and thus to any cross section which is integrated over a fixed range in M_Y. The partonic cross sections $\hat{\sigma}^{ei}$ are the same as those for inclusive DIS and the dpdf's f_i^D, which are not known from first principles, should obey the DGLAP evolution equations [25].

In addition to the rigorous theoretical prescription represented by equation (3), an additional assumption is necessary for the H1 fits in [3], that the shape of the dpdf's is independent of $x_{I\!P}$ and t and that their normalisation is controlled by Regge asymptotics [26]. Although this assumption has no

[3] A framework also exists to include higher order operators [24].

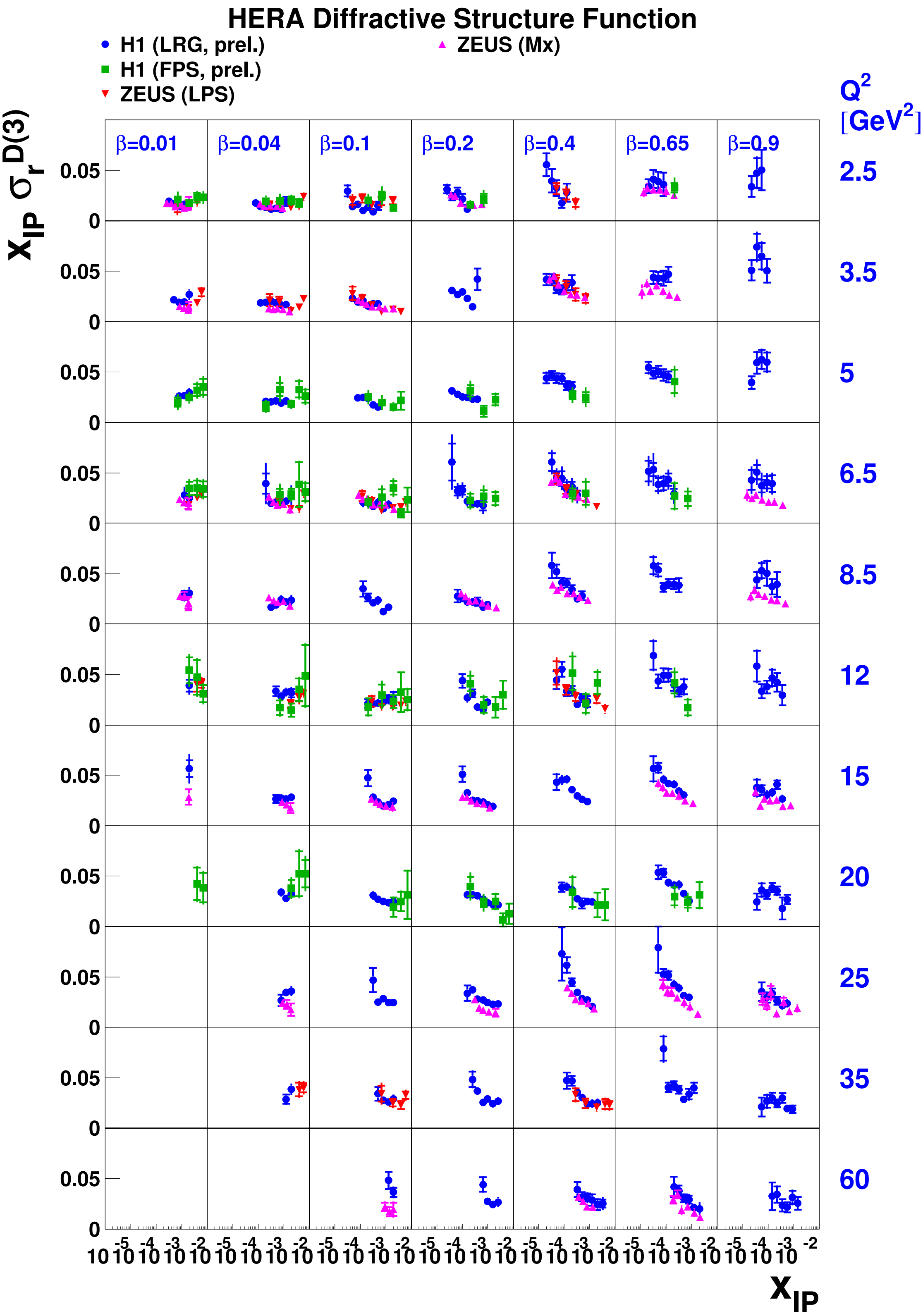


Fig. 3: Summary plot of all diffractive DIS data sets considered here. Additional H1-LRG data with $Q^2 < 2.5$ GeV^2, $Q^2 = 45$ GeV^2 and $Q^2 > 60$ GeV^2 are not shown. The Q^2 and β values for all data sets have been shifted to the H1 bin centres using small translation factors. The ZEUS data have been multiplied by a universal factor of 0.77 and the LPS and FPS data by factors of 1.1, such that all data sets correspond to $M_Y < 1.6$ GeV. Relative normalisation uncertainties of 15% due to these factors and further normalisation uncertainties of $\pm 6.7\%$ (H1-LRG) and $^{+12\%}_{-10\%}$ (ZEUS-LPS) data are not shown.

solid basis in QCD, it is compatible with the data fitted. The diffractive parton distributions can then be factorised into a term depending only on $x_{I\!P}$ and t and a term depending only on x (or β) and Q^2:

$$f_i^D(x_{I\!P}, t, x, Q^2) = f_{I\!P/p}(x_{I\!P}, t) \cdot f_i^{I\!P}(\beta = x/x_{I\!P}, Q^2) \; . \tag{4}$$

Under this 'Regge' factorisation assumption, the diffractive exchange can be treated as an object (a 'pomeron', $I\!P$) with a partonic structure given by parton distributions $f_i^{I\!P}(\beta, Q^2)$. The variable β then corresponds to the fraction of the pomeron longitudinal momentum carried by the struck parton. The 'pomeron flux factor' $f_{I\!P/p}(x_{I\!P}, t)$ represents the probability that a pomeron with particular values of $x_{I\!P}$ and t couples to the proton.

In the fit, the $x_{I\!P}$ dependence is parameterised using a Regge flux factor

$$f_{I\!P/p}(x_{I\!P}, t) = A \cdot \int_{t_{cut}}^{t_{min}} \frac{e^{B_{I\!P} t}}{x_{I\!P}^{2\alpha_{I\!P}(t)-1}} \, \mathrm{d}t \; , \tag{5}$$

where $t_{cut} = -1.0$ GeV^2, $|t_{min}|$ is the minimum kinematically allowed value of $|t|$ and the pomeron trajectory is assumed to be linear, $\alpha_{I\!P}(t) = \alpha_{I\!P}(0) + \alpha'_{I\!P} t$. The parameters $B_{I\!P}$ and α' and their uncertainties are fixed as described in [3]. The value of A is chosen such that the flux factor is normalised to unity at $x_{I\!P} = 0.003$. The pomeron intercept is then obtained from the $x_{I\!P}$ dependence of the data and takes the value $\alpha_{I\!P}(0) = 1.173 \; \pm 0.018 \text{ (stat.)} \; \pm 0.017 \text{ (syst.)} \; ^{+0.063}_{-0.035} \text{ (model)}$.

The description of the data is improved with the inclusion of an additional separately factorisable sub-leading exchange with a trajectory intercept of $\alpha_{I\!R}(0) = 0.50$ and parton densities taken from a parameterisation of the pion [27]. This exchange contributes significantly only at low β and large $x_{I\!P}$.

The dpdf's are modelled in terms of a light flavour singlet

$$\Sigma(z) = u(z) + d(z) + s(z) + \bar{u}(z) + \bar{d}(z) + \bar{s}(z) \; , \tag{6}$$

with $u = d = s = \bar{u} = \bar{d} = \bar{s}$ and a gluon distribution $g(z)$ at a starting scale $Q_0^2 = 3$ GeV^2. Here, z is the momentum fraction of the parton entering the hard sub-process with respect to the diffractive exchange, such that $z = \beta$ for the lowest-order quark parton model process, whereas $0 < \beta < z$ for higher order processes. The singlet quark and gluon distributions are parameterised using the form

$$z p_i(z, Q_0^2) = \left[\sum_{j=1}^{n} C_j^i P_j(2z - 1) \right]^2 e^{\frac{0.01}{z-1}} \; , \tag{7}$$

where $P_j(\xi)$ is the j^{th} member of a set of Chebychev polynomials[4]. The series is squared to ensure positivity. The exponential term is added to guarantee that the dpdf's tend to zero in the limit of $z \to 1$. It has negligible influence on the extracted partons at low to moderate z. The numbers of terms in the polynomial parameterisations are optimised to the precision of the data, with the first three terms in the series used for both the quark singlet and the gluon distributions, yielding 3 free parameters (C_j^Σ and C_j^g) for each. The normalisation of the sub-leading exchange contribution at high $x_{I\!P}$ is also determined by the fit such that the total number of free parameters is 7. The data used in the fit are restricted to $M_{\mathrm{X}} > 2$ GeV to suppress non-leading twist contributions. The effects of F_L^D are considered through its relation to the NLO gluon density, such that no explicit cut on y is required.

The NLO DGLAP equations are used to evolve the dpdfs to $Q^2 > Q_0^2$ using the method of [28], extended for diffraction. No momentum sum rule is imposed. Charm quarks are treated in the massive scheme (appearing via boson gluon fusion processes) with $m_c = 1.5 \pm 0.1$ GeV. The strong coupling is set via[5] $\Lambda_{\mathrm{QCD}}^{\overline{\mathrm{MS}}} = 200 \pm 30$ MeV. The statistical and experimental systematic errors on the data points

[4] $P_1 = 1$, $P_2 = \xi$ and $P_{j+1}(\xi) = 2\xi P_j(\xi) - P_{j-1}(\xi)$.

[5] Although this value is rather different from the world average, we retain it here for consistency with previous H1 preliminary results, where it has been used consistently for QCD fits [3] and final state comparisons [8, 9].

and their correlations are propagated to obtain error bands for the resulting dpdfs, which correspond to increases in the χ^2 by one unit [29]. A theoretical error on the dpdfs is estimated by variations of $\Lambda_{\rm QCD}$, m_c and the parameterisation of the $x_{I\!P}$ dependences as described in [3]. No theoretical uncertainty is assigned for the choice of parton parameterisation, though the results are consistent within the quoted uncertainties if alternative approaches [30] are used. No inhomogeneous term of the type included in [31] is considered here. The presence of such a term would lead to a reduction in the gluon density extracted.

The central fit gives a good description of the data, with a χ^2 of 308.7 for 306 degrees of freedom. The resulting diffractive quark singlet and gluon distributions are shown in figure 4. Both extend to large fractional momenta z. Whereas the singlet distribution is well constrained by the fit, there is a substantial uncertainty in the gluon distribution, particularly for $z \gtrsim 0.5$. The fraction of the exchanged momentum carried by gluons integrated over the range $0.01 < z < 1$ is $75 \pm 15\%$ (total error), confirming the conclusion from earlier work [19] that diffraction is a gluon-induced phenomenon. These dpdf's have been astonishingly successful in describing diffractive final state data in DIS such as charm [9] and jet [8] production, which, being induced by boson-gluon fusion-type processes, are roughly proportional to the diffractive gluon density.

4.2 Fit to ZEUS Data

A very similar fit to that described in section 4.1 is performed to the ZEUS-$M_{\rm X}$ data and the implications of the differences between the data sets to the dpdf's are investigated. The data are fitted in their original binning scheme, but are scaled to $M_{\rm Y} < 1.6$ GeV using the factor of 0.77. As for the fit to the H1 data, the first 3 terms are included in the polynomial expansions for the quark and gluon densities at the starting scale for QCD evolution. The same fit program, prescription and parameters are used as was the case for the H1 2002 NLO fit, with the following exceptions.

- ZEUS-$M_{\rm X}$ data with $Q^2 > 4$ GeV^2 are included in the fit, whereas only H1 data with $Q^2 > 6.5$ GeV^2 are included. It has been checked that the result for ZEUS is not altered significantly if the minimum Q^2 value is increased to 6 GeV^2.
- The quadratic sum of the statistical and systematic error is considered, i.e. there is no treatment of correlations between the data points through the systematics.
- No sub-leading Reggeon exchange component is included in the parameterisation. Including one does not improve or alter the fit significantly.
- The Pomeron intercept is fitted together with the dpdf's, in contrast to the two stage process of [3]. This does not influence the results significantly, though it does decrease the uncertainty on $\alpha_{I\!P}(0)$.

The fit describes the ZEUS-$M_{\rm X}$ data well ($\chi^2 = 90$ for 131 degrees of freedom) and yields a value for the Pomeron intercept of $\alpha_{I\!P}(0) = 1.132 \pm 0.006$ (experimental error only). This value is in agreement with the H1 result if the full experimental and theoretical errors are taken into account. A good fit is thus obtained without any variation of $\alpha_{I\!P}(0)$ with Q^2 or other deviation from Regge factorisation.

The diffractive parton densities from the fit to the ZEUS-$M_{\rm X}$ data are compared with the results from H1 in figure 4. The differences observed between the H1 and the ZEUS data are directly reflected in the parton densities. The quark singlet densities are closely related to the measurements of F_2^D themselves. They are similar at low Q^2 where the H1 and ZEUS data are in good agreement, but become different at larger Q^2, where discrepancies between the two data sets are observed. This difference between the Q^2 dependences of the H1 and ZEUS data is further reflected in a difference of around a factor of 2 between the gluon densities, which are roughly proportional to the logarithmic Q^2 derivative $\partial F_2^D / \partial \ln Q^2$ [32].

The H1-LRG and ZEUS-$M_{\rm X}$ data are shown together with the results from both QCD fits in figure 5. Both fits give good descriptions of the data from which they are obtained. The differences between the two data sets are clearly reflected in the fit predictions, most notably in the Q^2 dependence.

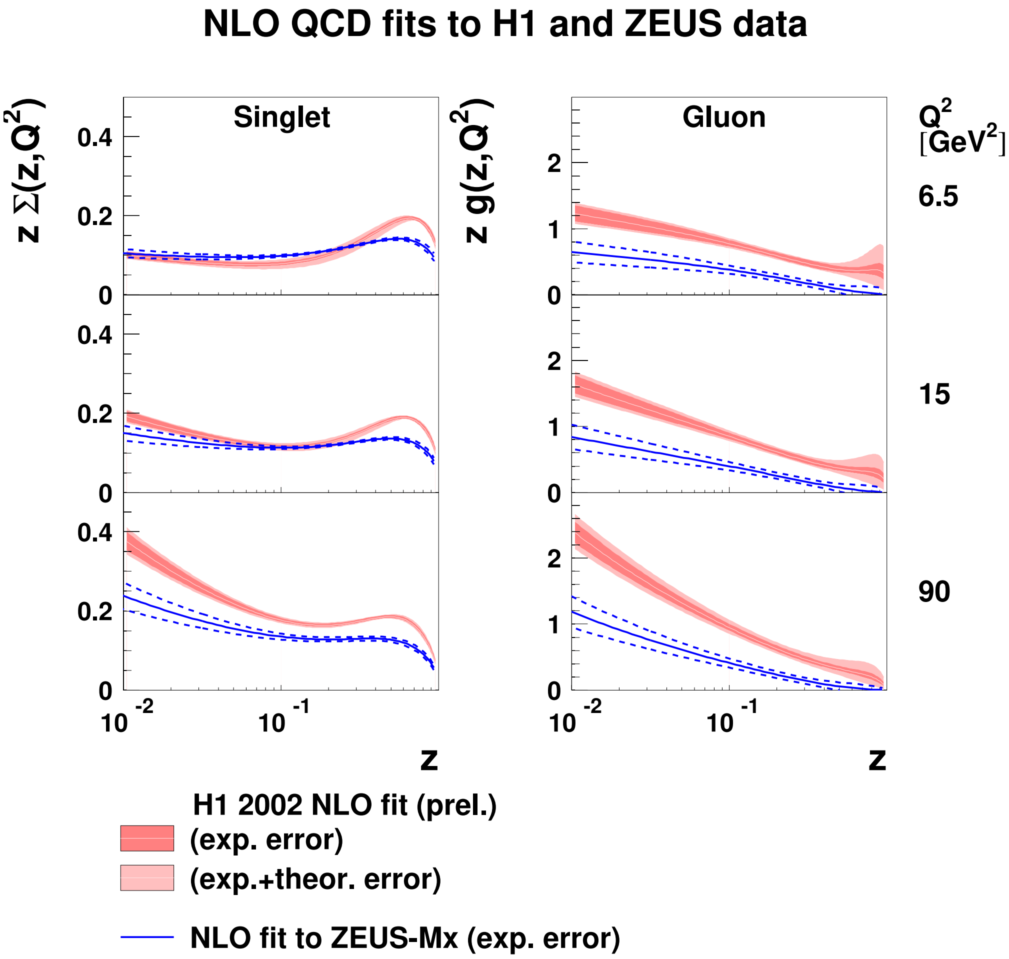


Fig. 4: Diffractive quark singlet and gluon pdf's for various Q^2 values, as obtained from the NLO DGLAP fits to the H1-LRG and ZEUS-M_X data. The bands around the H1 result indicate the experimental and theoretical uncertainties. The dotted lines around the result for ZEUS indicate the experimental uncertainty. The ZEUS data used in the fit are scaled by a normalisation factor of 0.77 to match the H1-LRG range of $M_Y < 1.6$ GeV. This factor is reflected in the normalisations of the quark and gluon densities. An uncertainty of 15% on this factor is not included in the error bands shown.

5 Summary

Recent diffractive structure function data from H1 and ZEUS have been compared directly. The leading proton data from both experiments (H1-FPS and ZEUS-LPS) are in good agreement with one other and with the H1 large rapidity gap data (H1-LRG). There is reasonable agreement between the H1-LRG and the ZEUS-M_X data over much of the kinematic range. However, differences are observed at the highest β (smallest M_X) and the Q^2 dependence at intermediate to low β is weaker for the ZEUS-M_X data than is the case for the H1-LRG data.

An NLO DGLAP QCD fit has been performed to the ZEUS-M_X data, using the same theoretical framework, assumptions and parameterisations as have been employed previously for the H1-2002-prelim NLO QCD fit to a subset of the H1-LRG data. As a consequence of the differences between the Q^2 dependences of the H1-LRG and ZEUS-M_X data, the gluon density obtained from the ZEUS data is significantly smaller than that for H1.

References

[1] S. Chekanov et al. [ZEUS Collaboration], Nucl. Phys. **B713**, 3 (2005).

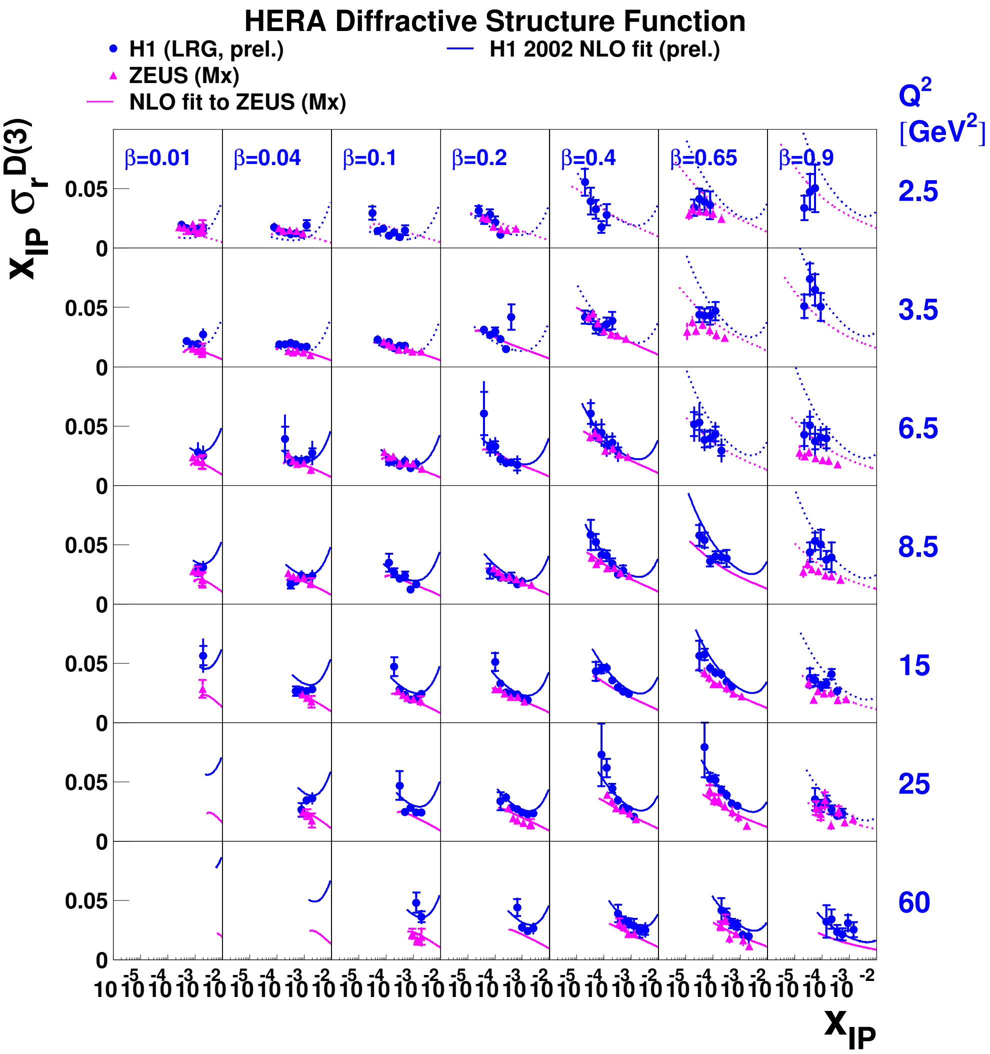


Fig. 5: As figure 2, but also showing the predictions using the NLO QCD fits to the H1-LRG and ZEUS-M_X data (uncertainties not shown).

[2] S. Chekanov et al. [ZEUS Collaboration], Eur. Phys. J. **C38**, 43 (2004).

[3] H1 Collaboration, paper 980 contributed to ICHEP 2002, Amsterdam. H1prelim-02-012.

[4] H1 Collaboration, paper 981 contributed to ICHEP 2002, Amsterdam. H1prelim-02-112.

[5] H1 Collaboration, paper 5-090 contributed to EPS 2003, Aachen. H1prelim-03-011.

[6] H1 Collaboration, paper 984 contributed to ICHEP 2002, Amsterdam. H1prelim-01-112.

[7] J. Collins, Phys. Rev. **D57**, 3051 (1998). Erratum ibid **D61** (2000) 019902.

[8] H1 Collaboration, paper 6-0177 contributed to ICHEP 2004, Beijing. H1prelim-04-113.

[9] H1 Collaboration, paper 6-0178 contributed to ICHEP 2004, Beijing. H1prelim-04-111.

[10] J. Collins, L. Frankfurt, M. Strikman, Phys. Lett. **B307**, 161 (1993).

[11] A. Berera, D. Soper, Phys. Rev. **D50**, 4328 (1994).

[12] A. Berera, D. Soper, Phys. Rev. **D53**, 6162 (1996).

[13] T. Affolder et al. [CDF Collaboration], Phys. Rev. Lett. **84**, 5043 (2000).

[14] J. Bjorken, Phys. Rev. **D47**, 101 (1993).

[15] E. Gotsman, E. Levin, U. Maor, Phys. Lett. **B438**, 229 (1998).

[16] B. Cox, J. Forshaw, L. Lonnblad, JHEP **10**, 023 (1999).

[17] A. Kaidalov, V. Khoze, A. Martin, M. Ryskin, Phys. Lett. **B567**, 61 (2003).

[18] J. Forshaw, these proceedings.

[19] C. Adloff et al. [H1 Collaboration], Z. Phys. **C76**, 613 (1997).

[20] C. Adloff et al. [H1 Collaboration], Z. Phys. **C74**, 221 (1997).

[21] H1 Collaboration, paper 571 contributed to ICHEP 1998, Vancouver.

[22] L. Trentadue, G. Veneziano, Phys. Lett. **B323**, 201 (1994).

[23] M. Grazzini, L. Trentadue, G. Veneziano, Nucl. Phys. **B519**, 394 (1998).

[24] J. Blumlein, D. Robaschik, Phys. Rev. **D65**, 096002 (2002).

[25] A. Hebecker, Nucl. Phys. **B505**, 349 (1997).

[26] G. Ingelman, P. Schlein, Phys. Lett. **B152**, 256 (1985).

[27] J. Owens, Phys. Rev. **D30**, 943 (1984).

[28] C. Adloff et al. [H1 Collaboration], Eur. Phys. J. **C21**, 33 (2001).

[29] C. Pascaud, F. Zomer, Preprint LAL-95-05.

[30] A. Martin, R. Roberts, J. Stirling, R. Thorne, Phys. Lett. **B531**, 216 (2002).

[31] G. Watt, A. Martin, M. Ryskin, these proceedings.

[32] K. Prytz, Phys. Lett. **B311**, 286 (1993).

Diffractive parton distributions from the HERA data

Michael Groys[a], *Aharon Levy*[a] *and Alexander Proskuryakov*[b]
[a]Raymond and Beverly Sackler Faculty of Exact Sciences, School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
[b]Institute of Nuclear Physics, Moscow State University, Moscow, Russia

Abstract

Measurements of the diffractive structure function, F_2^D, of the proton at HERA are used to extract the partonic structure of the Pomeron. Regge Factorization is tested and is found to describe well the existing data within the selected kinematic range. The analysis is based on the next to leading order QCD evolution equations. The results obtained from various data sets are compared.

1 INTRODUCTION

In the last 10 years a large amount of diffractive data was accumulated at the HERA collider [1, 2, 3]. There are three methods used at HERA to select diffractive events. One uses the Leading Proton Spectrometer (LPS) [3] to detect the scattered proton and by choosing the kinematic region where the scattered proton looses very little of its initial longitudinal energy, it ensures that the event was diffractive. A second method [2] simply requests a large rapidity gap (LRG) in the event and fits the data to contributions coming from Pomeron and Reggeon exchange. The third method [1] relies on the distribution of the mass of the hadronic system seen in the detector, M_X, to isolate diffractive events and makes use of the Forward Plug Calorimeter (FPC) to maximize the phase space coverage. We will refer to these three as ZEUS LPS, H1 and ZEUS FPC methods.

The experiments [4, 5, 6] provide sets of results for inclusive diffractive structure function, $x_{I\!P} F_2^{D(3)}$, in different regions of phase space. In extracting the initial Pomeron parton distribution functions (pdfs), the data are fitted assuming the validity of Regge factorization.

In the present study, Regge factorization is tested. New fits, based on a NLO QCD analysis, are provided and include the contribution of the longitudinal structure function. The obtained PDFs are systematically analyzed. A comparison of the different experimental data sets is provided. Additional quantities derived from the fit results are also presented.

In order to make sure that diffractive processes are selected, a cut of $x_{I\!P} < 0.01$ was performed, where $x_{I\!P}$ is the fraction of the proton momentum carried by the Pomeron. It was shown [7] that this cuts ensures the dominance of Pomeron exchange. In addition, a cut of $Q^2 > 3$ GeV^2 was performed on the exchanged photon virtuality for applying the NLO analysis. Finally, a cut on $M_X > 2$ GeV was used so as to exclude the light vector meson production.

2 Regge factorization

The Regge Factorization assumption can be reduced to the following,

$$F_2^{D(4)}(x_{I\!P},\, t\,, \beta,\, Q^2) \;\; = \;\; f(x_{I\!P}, t) \cdot F(\beta, Q^2), \tag{1}$$

where $f(x_{I\!P}, t)$ represents the Pomeron flux which is assumed to be independent of β and Q^2 and $F(\beta, Q^2)$ represents the Pomeron structure and is β and Q^2 dependent. In order to test this assumption, we check whether the flux $f(x_{I\!P}, t)$ is indeed independent of β and Q^2 on the basis of the available experimental data.

The flux is assumed to have a form $\sim x_{I\!P}^{-A}$ (after integrating over t which is not measured in the data) . A fit of this form to the data was performed in different Q^2 intervals, for the whole β range, and for different β intervals for the whole Q^2 range.

Figure 1 shows the Q^2 dependence of the exponent A for all three data sets, with the $x_{I\!P}$ and M_X cuts as described in the introduction. The H1 and the LPS data show no Q^2 dependence. The ZEUS FPC data show a small increase in A at the higher Q^2 region. It should be noted that while for the H1 and LPS data, releasing the $x_{I\!P}$ cut to 0.03 seems to have no effect, the deviation of the ZEUS FPC data from a flat dependence increases from a 2.4 standard deviation (s.d.) to a 4.2 s.d. effect (not shown).

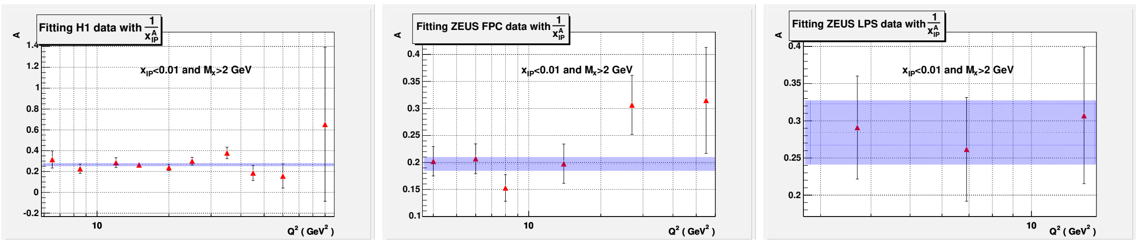


Fig. 1: A as a function of Q^2 for $x_{I\!P} < 0.01$ and $M_X > 2$ GeV, for the three data sets, as indicated in the figure. The line corresponds to a fit over the whole Q^2 region

The β dependence of A is shown in figure 2. All three data sets seem to show no β dependence, within the errors of the data. Note however, that by releasing the $x_{I\!P}$ cut to higher values, a strong dependence of the flux on β is observed (not shown).

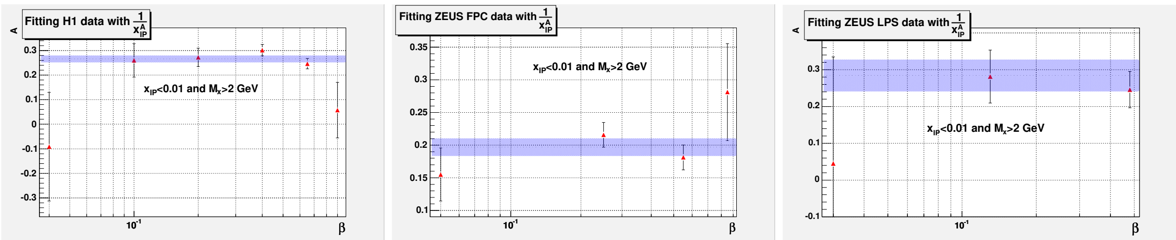


Fig. 2: A as a function of β for $x_{I\!P} < 0.01$ and $M_X > 2$ GeV, for the three data sets, as indicated in the figure. The line corresponds to a fit over the whole β region

We thus conclude that for $x_{I\!P} < 0.01$, the Pomeron flux seems to be independent of Q^2 and of β and thus the Regge factorization hypothesis holds.

3 NLO QCD fits

We parameterized the parton distribution functions of the Pomeron at Q_0^2 = 3 GeV^2 in a simple form of $Ax^b(1-x)^c$ for u and d quarks (and anti-quarks) and set all other quarks to zero at the initial scale. The gluon distribution was also assumed to have the same mathematical form. We thus had 3 parameters for quarks, 3 for gluons and an additional parameter for the flux, expressed in terms of the Pomeron intercept $\alpha_{I\!P}(0)$. Each data set was fitted to 7 parameters and a good fit was achieved for each. The H1 and ZEUS FPC had χ^2/df ≈ 1, while for the LPS data, the obtained value was 0.5. The data together with the results of the fits are shown in figure 3. The following values were obtained for $\alpha_{I\!P}(0)$, for each of the three data sets: $\alpha_{I\!P}(0) = 1.138 \pm 0.011$, for the ZEUS FPC data, $\alpha_{I\!P}(0) = 1.189 \pm 0.020$, for the ZEUS LPS data, $\alpha_{I\!P}(0) = 1.178 \pm 0.007$, for the H1 data.

The parton distribution functions are shown in figure 4 for the H1 and the ZEUS FPC data points. Because of the limited β range covered by the LPS data, the resulting pdfs uncertainties are large and are

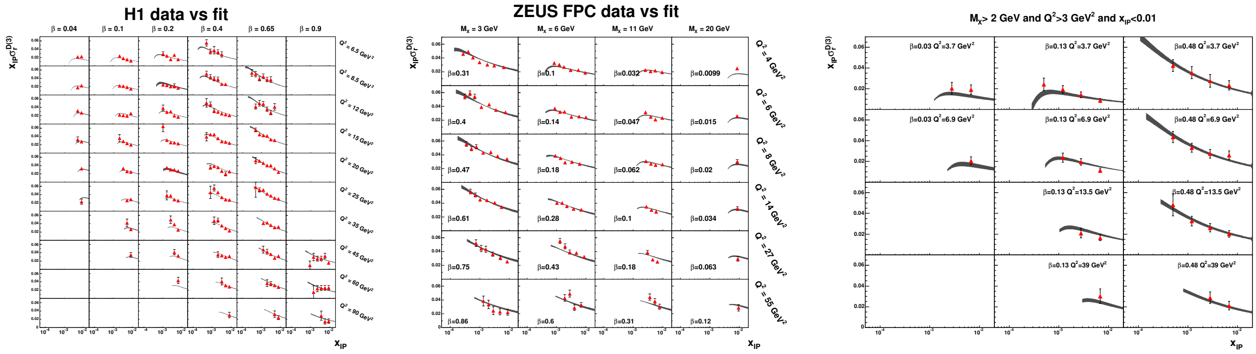


Fig. 3: The diffractive reduced cross section of the proton multiplies by $x_{I\!P}$, as a function of $x_{I\!P}$ for the different data sets (the most right plot is for the LPS data) in different bins of Q^2 and β, as indicated in the figure. The bands are the results of the fits including uncertainties.

not shown here. In fact one gets two solutions; one where the gluon contribution is dominant and another one where the gluons and the quarks contribute about equally. Note however that once the diffractive charm structure function data [8] are included in the fit, the gluon dominant solution is chosen (see below). For the H1 fit one sees the dominance of the gluons in all the β range. For the ZEUS FPC data, the quark constituent of the Pomeron dominates at high β while gluons dominate at low β. We can

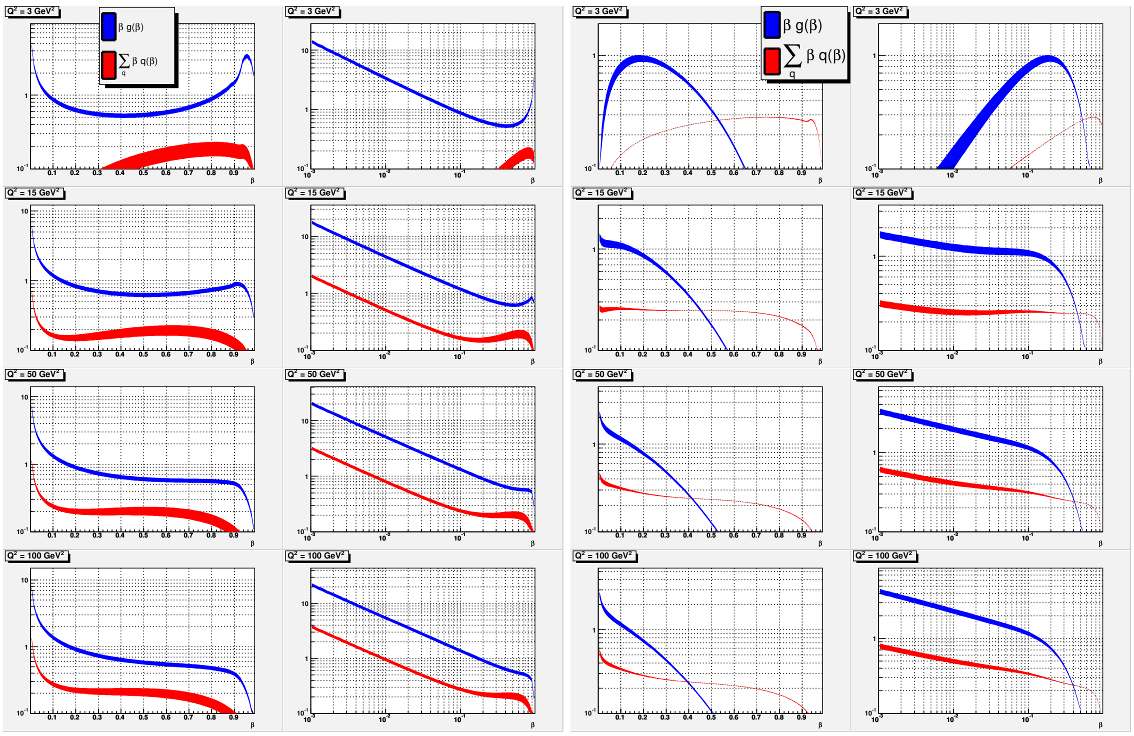


Fig. 4: Quark and gluon pdfs of the Pomeron as obtained from the H1 data fit (left two figures) and from the ZEUS FPC data fit (two rightmost figures) as a function of β, at different values of Q^2.

quantify this by calculating the Pomeron momentum carried by the gluons. Using the fit results one gets for the H1 data 80-90%. while for the ZEUS FPC data, 55-65%.

4 Comparison of the data sets

One way of checking the compatibility of all three data sets is to make an overall fit for the whole data sample. Since the coverage of the β range in the LPS data is limited, we compare only the H1 and the ZEUS FPC data. A fit with a relative overall scaling factor of the two data sets failed. Using the fit results of one data sets superimposed on the other shows that the fit can describe some kinematic regions, while failing in other bins. This leads to the conclusion that there seems to be some incompatibility between the two data sets.

5 Comparison to $F_2^{D(3)}$(charm)

The ZEUS collaboration measured the diffractive charm structure function, $F_2^{D(3),c\bar{c}}$ [8] and these data were used together with the LPS data for a combined fit [6]. The charm data are shown in figure 5 as function of β. The full line shown the resulting best fit, where the contribution from charm was calculated as photon-gluon fusion. In the same figure one sees the prediction from the NLO QCD fit to the ZEUS

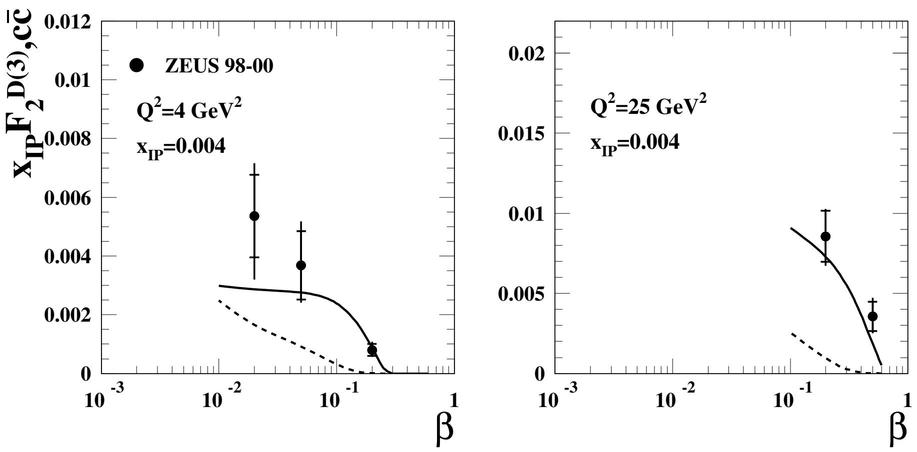


Fig. 5: Diffractive charm structure function, $F_2^{D(3),c\bar{c}}$, as a function of β for values of Q^2 and $x_{I\!P}$ as indicated in the figure. The full line is the result of a combined fit to the LPS and the diffractive charm data. The dashed line is the prediction using the gluons from the ZEUS FPC fit.

FPC data (dashed line). Clearly, the gluons from the ZEUS FPC fit can not describe the diffractive charm data.

Acknowledgments

We would like to thank Prof. Halina Abramowicz for her useful and clarifying comments during this analysis. We would also like to thank Prof. John Collins for providing the program to calculate the NLO QCD equations for the diffractive data. This work was supported in part by the Israel Science Foundation (ISF).

References

[1] ZEUS Collab., M. Derrick et al., *Phys. Lett.* **315** (1993) 481; J. Breitweg et al., *Eur. Phys. J.* **C6** (1999) 43.

[2] H1 Collab., C. Adloff et al., *Zeit. Phys.* **C76** (1997) 613.

[3] ZEUS Collab., S. Chekanov et al., *Eur. Phys. J.* **C25** (2002) 169.

[4] ZEUS Collaboration, S. Chekanov et al., *Nucl. Phys.* **B713** (2005) 3.

[5] H1 Collaboration, “Measurement and NLO DGLAP QCD Interpretation of Diffractive Deep-Inelastic Scattering at HERA,” paper 089 submitted to EPS 2003, Aachen.

[6] ZEUS Collaboration, S. Chekanov et al., *Eur. Phys. J.* **C38** (2004) 43.

[7] K. Golec-Biernat, J. Kwiecinski and A. Szczurek, *Phys. Rev.* **D56** (1997) 3955.

[8] ZEUS Collaboration, S. Chekanov et al., *Nucl. Phys.* **B672** (2003) 3.

Diffractive parton distributions

G. Watt[a], A.D. Martin[b], M.G. Ryskin[b,c]
[a] Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany
[b] Institute for Particle Physics Phenomenology, University of Durham, DH1 3LE, UK
[c] Petersburg Nuclear Physics Institute, Gatchina, St. Petersburg, 188300, Russia

Abstract

We discuss the perturbative QCD description of diffractive deep-inelastic scattering, and extract diffractive parton distributions from recent HERA data. The asymptotic collinear factorisation theorem has important modifications in the sub-asymptotic HERA regime. In addition to the usual *resolved* Pomeron contribution, the *direct* interaction of the Pomeron must also be accounted for. The diffractive parton distributions are shown to satisfy an *inhomogeneous* evolution equation, analogous to the parton distributions of the photon.

1 Diffractive parton distributions from Regge factorisation

It is conventional to extract diffractive parton distribution functions (DPDFs) from diffractive deep-inelastic scattering (DDIS) data using two levels of factorisation. Firstly, collinear factorisation means that the diffractive structure function can be written as [1]

$$F_2^{\mathrm{D}(3)}(x_{I\!P},\beta,Q^2) = \sum_{a=q,g} C_{2,a}\otimes a^{\mathrm{D}}, \tag{1}$$

where the DPDFs $a^{\mathrm{D}} = zq^{\mathrm{D}}$ or zg^{D} satisfy DGLAP evolution:

$$\frac{\partial a^{\mathrm{D}}}{\partial \ln Q^2} = \sum_{a'=q,g} P_{aa'}\otimes {a'}^{\mathrm{D}}, \tag{2}$$

and where $C_{2,a}$ and $P_{aa'}$ are the *same* hard-scattering coefficients and splitting functions as in inclusive DIS. The factorisation theorem (1) applies when Q is made large, therefore it is correct up to power-suppressed corrections. It says nothing about the mechanism for diffraction. What *is* the exchanged object with vacuum quantum numbers ('Pomeron') which *causes* the large rapidity gap (LRG) characterising diffractive interactions?

In a second stage [2] Regge factorisation is usually assumed, such that

$$a^{\mathrm{D}}(x_{I\!P},z,Q^2) = f_{I\!P}(x_{I\!P})\, a^{I\!P}(z,Q^2), \tag{3}$$

where the Pomeron PDFs $a^{I\!P} = zq^{I\!P}$ or $zg^{I\!P}$. The Pomeron flux factor $f_{I\!P}$ is taken from Regge phenomenology,

$$f_{I\!P}(x_{I\!P}) = \int_{t_{\mathrm{cut}}}^{t_{\mathrm{min}}} \mathrm{d}t \quad \mathrm{e}^{B_{I\!P}\,t}\, x_{I\!P}^{1-2\alpha_{I\!P}(t)}. \tag{4}$$

Here, $\alpha_{I\!P}(t) = \alpha_{I\!P}(0) + \alpha'_{I\!P}\,t$, and the parameters $B_{I\!P}$, $\alpha_{I\!P}(0)$, and $\alpha'_{I\!P}$ should be taken from fits to soft hadron data. Although the first fits to use this approach assumed a 'soft' Pomeron, $\alpha_{I\!P}(0)\simeq 1.08$ [3], all recent fits require a substantially higher value to describe the data. In addition, a secondary Reggeon contribution is needed to describe the data for $x_{I\!P}\gtrsim 0.01$. This approach is illustrated in Fig. 1(a), where the virtualities of the t-channel partons are strongly ordered as required by DGLAP evolution. The Pomeron PDFs $a^{I\!P}$ are parameterised at some arbitrary low scale Q_0^2, then evolved up to the factorisation scale, usually taken to be the photon virtuality Q^2.

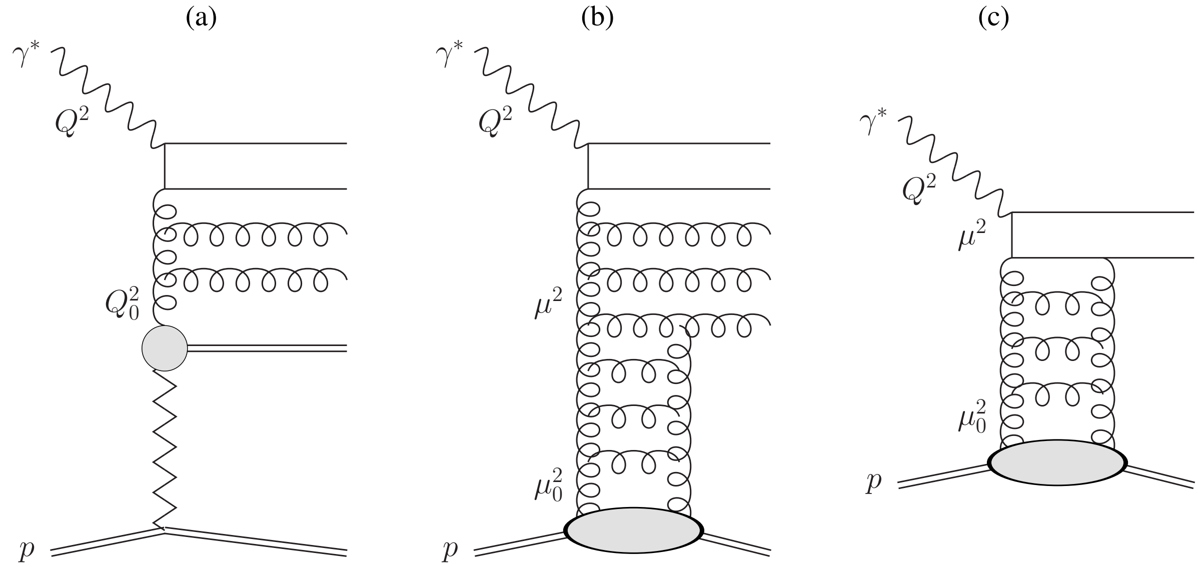


Fig. 1: (a) Resolved Pomeron contribution in the 'Regge factorisation' approach. (b) Resolved Pomeron contribution in the 'perturbative QCD' approach. (c) Direct Pomeron contribution in the 'perturbative QCD' approach.

Although this approach has been found to give a good description[1] of the DDIS data [4, 5, 6, 7], it has little theoretical justification. The 'Regge factorisation' of (3) is merely a simple way of parameterising the $x_{I\!P}$ dependence of the DPDFs. Note, however, that the effective Pomeron intercept $\alpha_{I\!P}(0)$ has been observed to depend on Q^2 [8]. The fact that the required $\alpha_{I\!P}(0)$ is greater than the 'soft' value indicates that there is a significant perturbative QCD (pQCD) contribution to DDIS.

2 Diffractive parton distributions from perturbative QCD

In pQCD, Pomeron exchange can be described by two-gluon exchange, two gluons being the minimum number needed to reproduce the quantum numbers of the vacuum. Two-gluon exchange calculations are the basis for the colour dipole model description of DDIS, in which the photon dissociates into $q\bar{q}$ or $q\bar{q}g$ final states. Such calculations have successfully been used to describe HERA data. The crucial question, therefore, is how to reconcile two-gluon exchange with collinear factorisation as given by (1) and (2). Are these two approaches compatible?

Generalising the $q\bar{q}$ or $q\bar{q}g$ final states to an arbitrary number of parton emissions from the photon dissociation, and replacing two-gluon exchange by exchange of a parton ladder, we have diagrams like that shown in Fig. 1(b) [9, 10, 11, 12]. Again, the virtualities of the t-channel partons are strongly ordered: $\mu_0^2 \ll \ldots \ll \mu^2 \ll \ldots \ll Q^2$. The scale μ^2 at which the Pomeron-to-parton splitting occurs can vary between $\mu_0^2 \sim 1\ \mathrm{GeV}^2$ and the factorisation scale Q^2. Therefore, to calculate the inclusive diffractive structure function, $F_2^{\mathrm{D}(3)}$, we need to integrate over μ^2:

$$F_2^{\mathrm{D}(3)}(x_{I\!P},\beta,Q^2) = \int_{\mu_0^2}^{Q^2} \frac{\mathrm{d}\mu^2}{\mu^2}\; f_{I\!P}(x_{I\!P};\mu^2)\; F_2^{I\!P}(\beta,Q^2;\mu^2). \tag{5}$$

Here, the perturbative Pomeron flux factor can be shown to be [12]

$$f_{I\!P}(x_{I\!P};\mu^2) = \frac{1}{x_{I\!P}B_D}\left[R_g\frac{\alpha_S(\mu^2)}{\mu}\; x_{I\!P}g(x_{I\!P},\mu^2)\right]^2. \tag{6}$$

[1]Note that the H1 2002 NLO fit [4] uses the 2-loop α_S with $\Lambda_{\mathrm{QCD}} = 200$ MeV for 4 flavours. This gives α_S values much smaller than the world average, meaning that the H1 2002 diffractive gluon density is artificially enhanced.

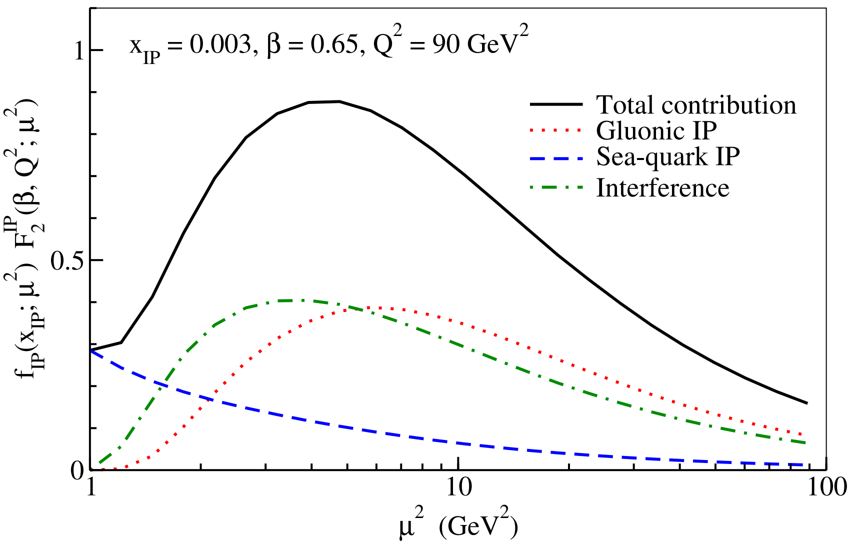


Fig. 2: Contributions to $F_2^{\mathrm{D}(3)}$ as a function of μ^2.

The diffractive slope parameter B_D comes from the t-integration, while the factor R_g accounts for the skewedness of the proton gluon distribution [13]. There are similar contributions from sea quarks, where $g(x_{I\!P}, \mu^2)$ in (6) is replaced by $S(x_{I\!P}, \mu^2)$, together with an interference term. In the fits presented here, we use the MRST2001 NLO gluon and sea-quark distributions of the proton [14]. The Pomeron structure function in (5), $F_2^{I\!P}(\beta, Q^2; \mu^2)$, is calculated from Pomeron PDFs, $a^{I\!P}(z, Q^2; \mu^2)$, evolved using NLO DGLAP from a starting scale μ^2 up to Q^2, taking the input distributions to be LO Pomeron-to-parton splitting functions, $a^{I\!P}(z, \mu^2; \mu^2) = P_{aI\!P}(z)$ [11, 12]. At first glance, it would appear that the perturbative Pomeron flux factor (6) behaves as $f_{I\!P}(x_{I\!P}; \mu^2) \sim 1/\mu^2$, so that contributions from large μ^2 are strongly suppressed. However, at large μ^2, the gluon distribution of the proton behaves as $x_{I\!P}g(x_{I\!P}, \mu^2) \sim (\mu^2)^\gamma$, where γ is the anomalous dimension. In the BFKL limit of $x_{I\!P} \to 0$, $\gamma \simeq 0.5$, so $f_{I\!P}(x_{I\!P}; \mu^2)$ would be approximately independent of μ^2. The HERA domain is in an intermediate region: γ is not small, but is less than 0.5. It is interesting to plot the integrand of (5) as a function of μ^2, as shown in Fig. 2. Notice that there is a large contribution from $\mu^2 > 3$ GeV^2, which is the value of the input scale Q_0^2 typically used in the 'Regge factorisation' fits of Sect. 1. Recall that fits using 'Regge factorisation' include contributions from $\mu^2 \leq Q_0^2$ in the input distributions, but neglect all contributions from $\mu^2 > Q_0^2$; from Fig. 2 this is clearly an unreasonable assumption.

As well as the *resolved* Pomeron contribution of Fig. 1(b), we must also account for the *direct* interaction of the Pomeron in the hard subprocess, Fig. 1(c), where there is no DGLAP evolution in the upper part of the diagram. Therefore, the diffractive structure function can be written as

$$F_2^{\mathrm{D}(3)} = \underbrace{\sum_{a=q,g} C_{2,a} \otimes a^{\mathrm{D}}}_{\text{Resolved Pomeron}} + \underbrace{C_{2,I\!P}}_{\text{Direct Pomeron}} ; \tag{7}$$

cf. (1) where there is no direct Pomeron contribution. The direct Pomeron coefficient function, $C_{2,I\!P}$, calculated from Fig. 1(c), will again depend on $f_{I\!P}(x_{I\!P}; \mu^2)$ given by (6). Therefore, it is formally suppressed by a factor $1/\mu^2$, but in practice does not behave as such; see Fig. 2.

The contribution to the DPDFs from scales $\mu > \mu_0$ is

$$a^{\mathrm{D}}(x_{I\!P}, z, Q^2) = \int_{\mu_0^2}^{Q^2} \frac{\mathrm{d}\mu^2}{\mu^2}\ f_{I\!P}(x_{I\!P}; \mu^2)\ a^{I\!P}(z, Q^2; \mu^2). \tag{8}$$

Differentiating (8), we see that the evolution equations for the DPDFs are [12]

$$\frac{\partial a^{\mathrm{D}}}{\partial \ln Q^2} = \sum_{a'=q,g} P_{aa'} \otimes a'^{\mathrm{D}} \; + \; P_{a\mathbb{P}}(z)\, f_{\mathbb{P}}(x_{\mathbb{P}}; Q^2); \tag{9}$$

cf. (2) where the second term of (9) is absent. That is, the DPDFs satisfy an *inhomogeneous* evolution equation [10, 12], with the extra inhomogeneous term in (9) leading to more rapid evolution than in the 'Regge factorisation' fits described in Sect. 1. Note that the inhomogeneous term will change the $x_{\mathbb{P}}$ dependence evolving upwards in Q^2, in accordance with the data, and unlike the 'Regge factorisation' assumption (3). Again, the inhomogeneous term in (9) is formally suppressed by a factor $1/Q^2$, but in practice does not behave as such; see Fig. 2.

Therefore, the diffractive structure function is analogous to the photon structure function, where there are both resolved and direct components and the photon PDFs satisfy an inhomogeneous evolution equation, where at LO the inhomogeneous term accounts for the splitting of the point-like photon into a $q\bar{q}$ pair. If we consider, for example, diffractive dijet photoproduction, there are four classes of contributions; see Fig. 3. The relative importance of each contribution will depend on the values of x_γ, the fraction of the photon's momentum carried by the parton entering the hard subprocess, and $z_{\mathbb{P}}$, the fraction of the Pomeron's momentum carried by the parton entering the hard subprocess.

3 Description of DDIS data

A NLO analysis of DDIS data is not yet possible. The direct Pomeron coefficient functions, $C_{2,\mathbb{P}}$, and Pomeron-to-parton splitting functions, $P_{a\mathbb{P}}$, need to be calculated at NLO within a given factorisation scheme (for example, $\overline{\mathrm{MS}}$). Here, we perform a simplified analysis where the usual coefficient functions $C_{2,a}$ and splitting functions $P_{aa'}$ $(a, a' = q, g)$ are taken at NLO, but $C_{2,\mathbb{P}}$ and $P_{a\mathbb{P}}$ are taken at LO [12]. We work in the fixed flavour number scheme, where there is no charm DPDF. Charm quarks are produced via $\gamma^* g^{\mathbb{P}} \to c\bar{c}$ at NLO [15] and $\gamma^* \mathbb{P} \to c\bar{c}$ at LO [16]. For light quarks, we include the direct Pomeron process $\gamma_L^* \mathbb{P} \to q\bar{q}$ at LO [12], which is higher-twist and known to be important at large β.

To see the effect of the direct Pomeron contribution and the inhomogeneous evolution, we make two types of fits:

"Regge" : The 'Regge factorisation' approach discussed in Sect. 1, where there is no direct Pomeron contribution and no inhomogeneous term in the evolution equation.

"pQCD" : The 'perturbative QCD' approach discussed in Sect. 2, where these effects are included.

We make separate fits to the recent H1 LRG (prel.) [4] and ZEUS M_X [8] $\sigma_r^{\mathrm{D}(3)}$ data, applying cuts $Q^2 \geq 3$ GeV^2 and $M_X \geq 2$ GeV, and allowing for overall normalisation factors of 1.10 and 1.43 to account for proton dissociation up to masses of 1.6 GeV and 2.3 GeV respectively. Statistical and systematic experimental errors are added in quadrature. The strong coupling is set via $\alpha_S(M_Z) = 0.1190$. We take the input forms of the DPDFs at a scale $Q_0^2 = 3$ GeV^2 to be

$$z\Sigma^{\mathrm{D}}(x_{\mathbb{P}}, z, Q_0^2) = f_{\mathbb{P}}(x_{\mathbb{P}})\, C_q\, z^{A_q}(1-z)^{B_q}, \tag{10}$$

$$zg^{\mathrm{D}}(x_{\mathbb{P}}, z, Q_0^2) = f_{\mathbb{P}}(x_{\mathbb{P}})\, C_g\, z^{A_g}(1-z)^{B_g}, \tag{11}$$

where $f_{\mathbb{P}}(x_{\mathbb{P}})$ is given by (4), and where $\alpha_{\mathbb{P}}(0)$, C_a, A_a, and B_a $(a = q, g)$ are free parameters. The secondary Reggeon contribution to the H1 data is treated in a similar way as in the H1 2002 fit [4], using the GRV pionic parton distributions [17]. Good fits are obtained in all cases, with $\chi^2/\mathrm{d.o.f.} = 0.75$, 0.71, 0.76, and 0.84 for the "Regge" fit to H1 data, "pQCD" fit to H1 data, "Regge" fit to ZEUS M_X data, and "pQCD" fit to ZEUS M_X data respectively. The "pQCD" fits are shown in Fig. 4, including a breakdown of the different contributions. The DPDFs are shown in Fig. 5. Note that the "pQCD" DPDFs are smaller than the corresponding "Regge" DPDFs at large z due to the inclusion of the higher-twist $\gamma_L^* \mathbb{P} \to q\bar{q}$ contribution. Also note that the "pQCD" DPDFs have slightly more rapid evolution

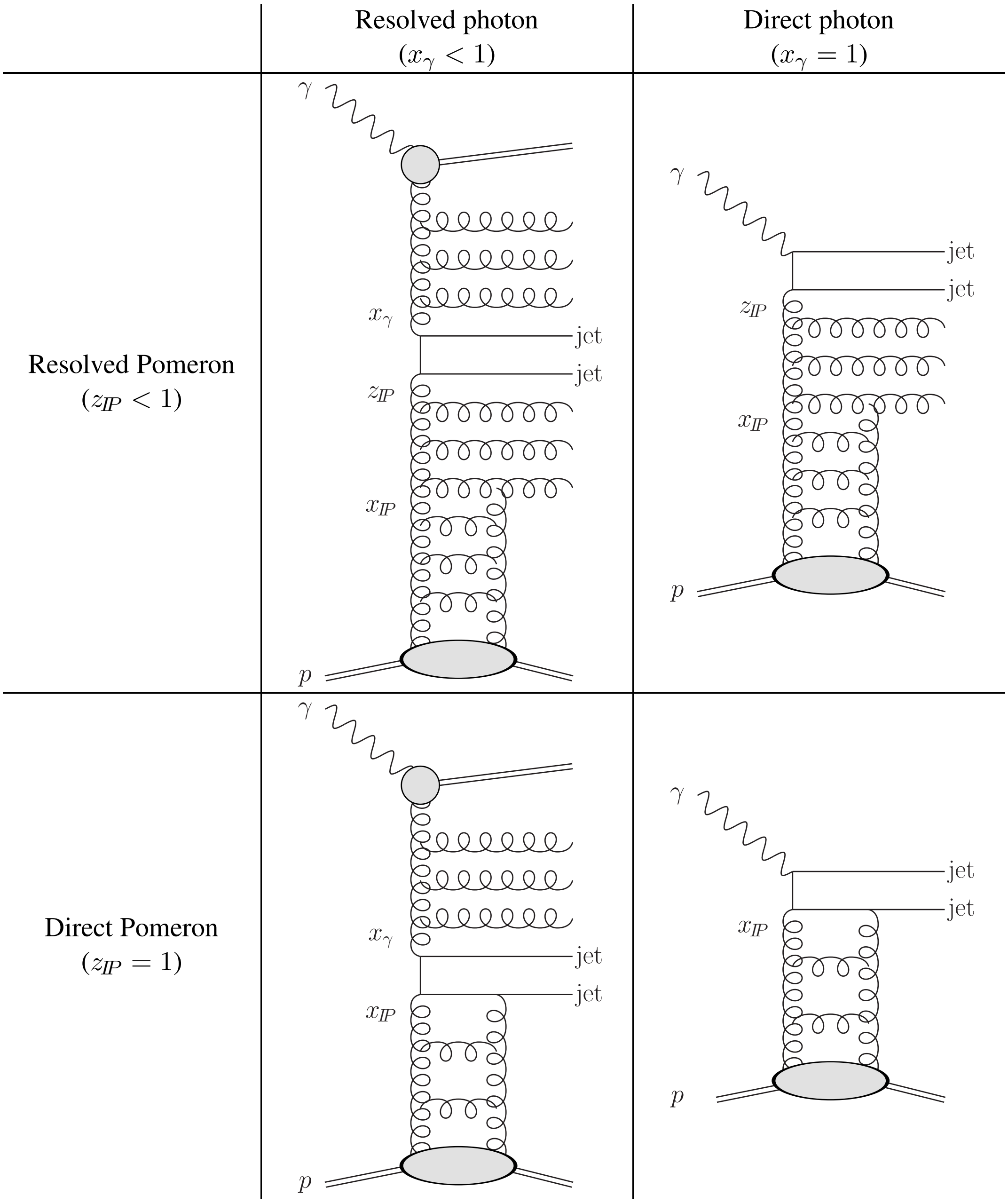


Fig. 3: The four classes of contributions to diffractive dijet photoproduction at LO. Both the photon and the Pomeron can be either ‘resolved’ or ‘direct’.

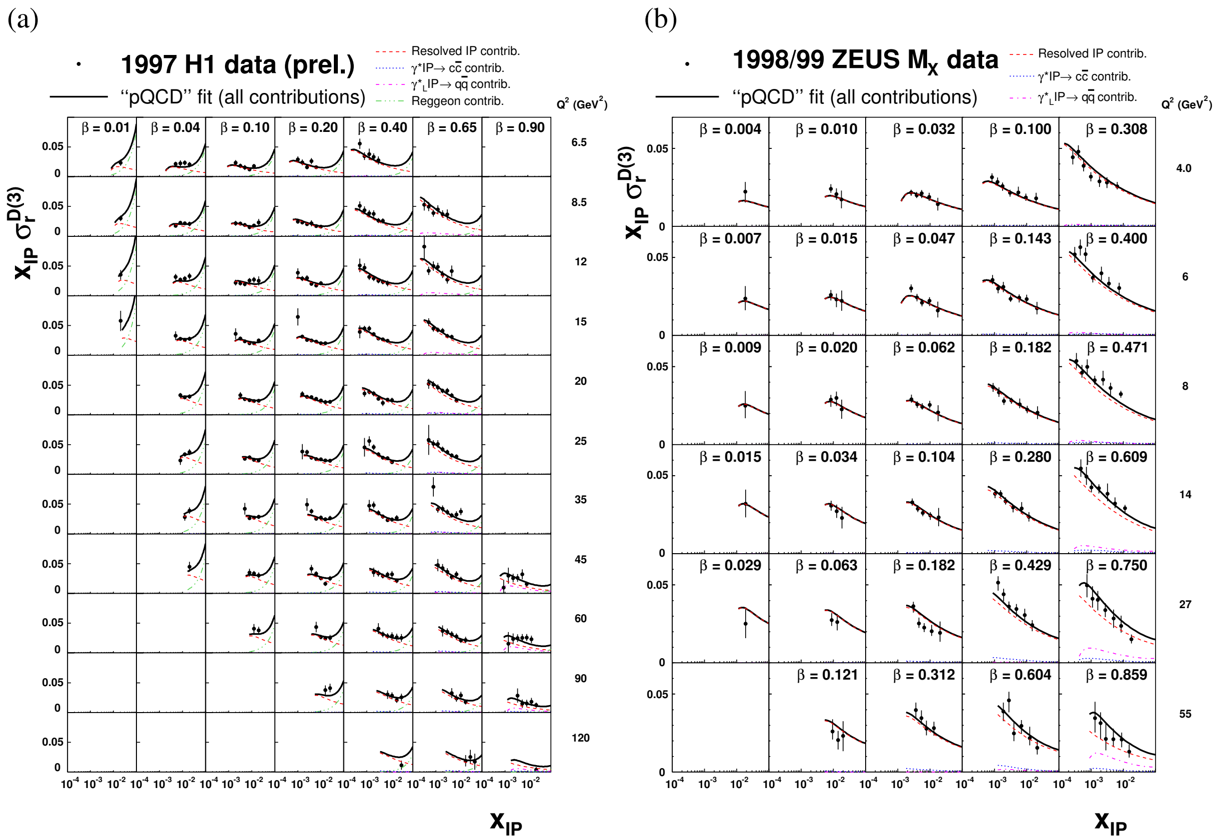


Fig. 4: "pQCD" fits to (a) H1 LRG and (b) ZEUS M_X data.

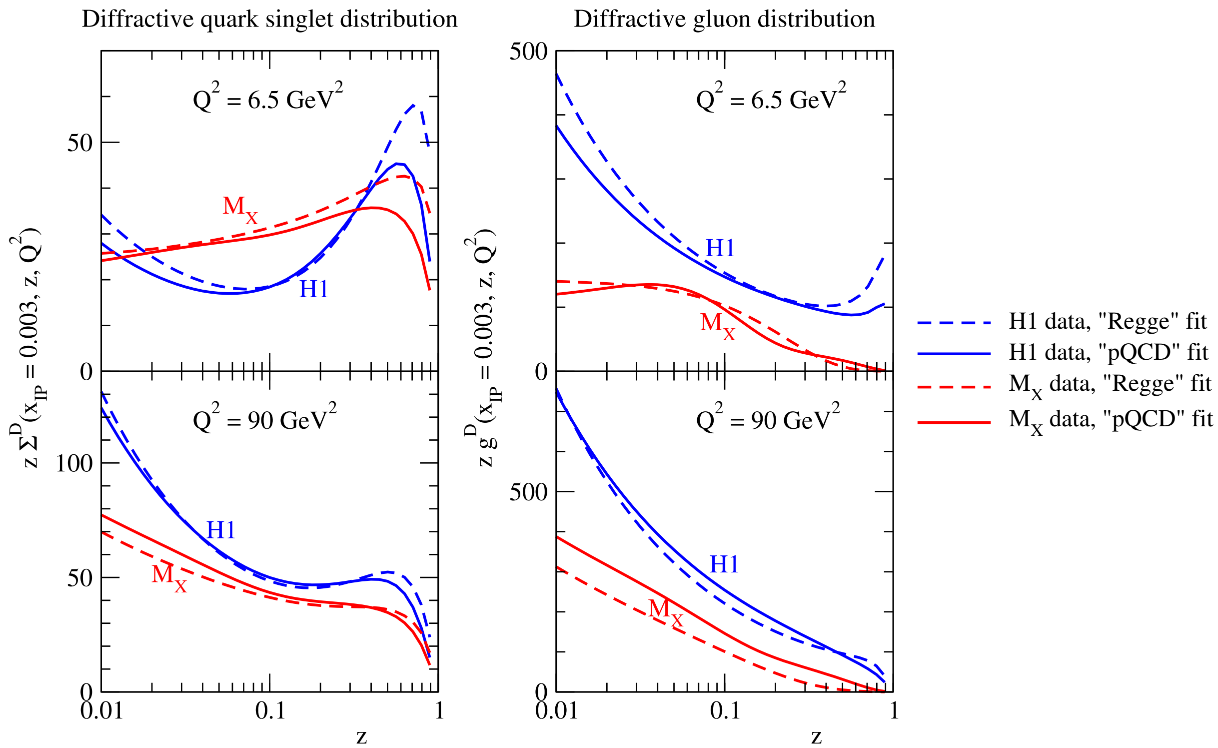


Fig. 5: DPDFs obtained from separate fits to H1 LRG and ZEUS M_X data using the "Regge" and "pQCD" approaches.

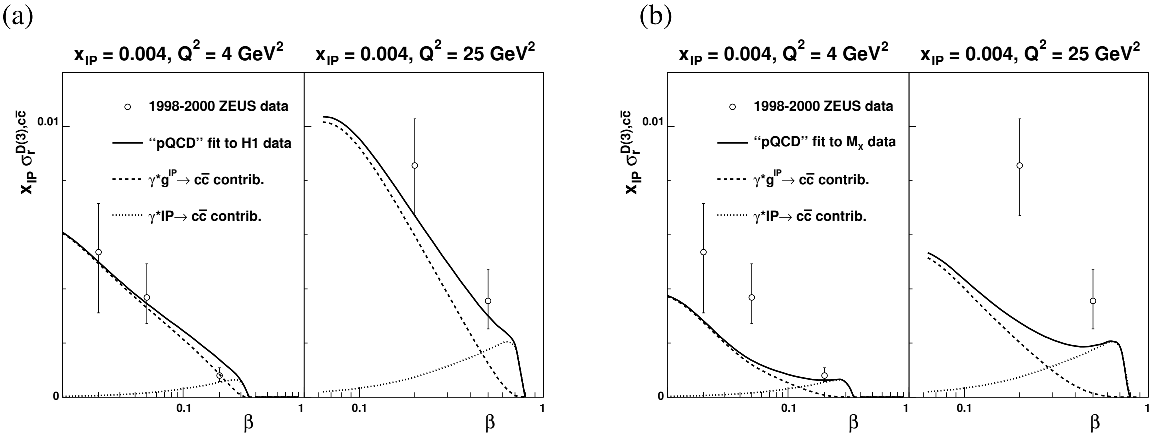


Fig. 6: Predictions for ZEUS LRG diffractive charm production data using DPDFs from the "pQCD" fits to (a) H1 LRG and (b) ZEUS M_X data. Note the large direct Pomeron ($\gamma^* I\!P \to c\bar{c}$) contribution at moderate β.

than the "Regge" DPDFs due to the extra inhomogeneous term in the evolution equation (9). There is a large difference between the DPDFs obtained from the H1 LRG and ZEUS M_X data due to the different Q^2 dependence of these data sets; see also [6, 7].

The predictions from the two "pQCD" fits for the charm contribution to the diffractive structure function as measured by ZEUS using the LRG method [18] are shown in Fig. 6. Our H1 LRG fit gives a good description, while our ZEUS M_X fit is too small at low β. Note that the direct Pomeron contribution is significant at moderate β. These charm data points were included in the determination of DPDFs from ZEUS LPS data [5], but only the resolved Pomeron ($\gamma^* g^{I\!P} \to c\bar{c}$) contribution was included and not the direct Pomeron ($\gamma^* I\!P \to c\bar{c}$) contribution. Therefore, the diffractive gluon distribution from the ZEUS LPS fit [5] needed to be artificially large to fit the charm data at moderate β.

4 Conclusions and outlook

To summarise, diffractive DIS is more complicated than inclusive DIS. Collinear factorisation holds, but we need to account for the direct Pomeron coupling, leading to an inhomogeneous evolution equation (9). Therefore, the treatment of DPDFs has more in common with photon PDFs than with proton PDFs. The H1 LRG and ZEUS M_X data seem to have a different Q^2 dependence, leading to different DPDFs. This issue needs further attention.[2] For a NLO analysis of DDIS data, the direct Pomeron coefficient functions, $C_{2,I\!P}$, and Pomeron-to-parton splitting functions, $P_{aI\!P}$, need to be calculated at NLO. There are indications [16] that there are large π^2-enhanced virtual loop corrections ('K-factors') similar to those found in the Drell–Yan process. As with all PDF determinations, the sensitivity to the form of the input parameterisation, (10) and (11), and input scale Q_0^2 needs to be studied.[3] The inclusion of jet and heavy quark DDIS data, and possibly $F_L^{\rm D(3)}$ if it is measured [19], would help to constrain the DPDFs further. The extraction of DPDFs from HERA data will provide an important input for calculations of diffraction at the LHC.

References

[1] J. C. Collins, Phys. Rev. **D57**, 3051 (1998).

[2] In particular, the main assumption of the M_X method, that the diffractive contribution to the $\ln M_X^2$ distribution is constant, while the non-diffractive contribution rises exponentially, is motivated by Regge theory in the limit that $t = 0$, $\alpha_{I\!P}(0) \equiv 1$, and $Q^2 \ll M_X^2$, and it is not clear that this should be true in general.

[3] The ZEUS LPS fit [5], where the data have large statistical uncertainties, found that the shape of the DPDFs had a significant dependence on the functional form of the initial parameterisation. Even for the other data sets, where the statistical uncertainties are smaller, there seems to be a fairly strong dependence on the value of the input scale Q_0^2.

[2] G. Ingelman and P. E. Schlein, Phys. Lett. **B152**, 256 (1985).

[3] A. Donnachie and P. V. Landshoff, Phys. Lett. **B296**, 227 (1992).

[4] H1 Collaboration, paper 089 submitted to EPS 2003, Aachen.

[5] S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. **C38**, 43 (2004).

[6] P. Newman and F.-P. Schilling, these proceedings.

[7] M. Groys, A. Levy and A. Proskuryakov, these proceedings.

[8] S. Chekanov *et al.* [ZEUS Collaboration], Nucl. Phys. **B713**, 3 (2005).

[9] M. G. Ryskin, Sov. J. Nucl. Phys. **52**, 529 (1990).

[10] E. Levin and M. Wusthoff, Phys. Rev. **D50**, 4306 (1994).

[11] A. D. Martin, M. G. Ryskin and G. Watt, Eur. Phys. J. **C37**, 285 (2004).

[12] A. D. Martin, M. G. Ryskin and G. Watt, Eur. Phys. J. **C44**, 69 (2005).

[13] A. G. Shuvaev, K. J. Golec-Biernat, A. D. Martin and M. G. Ryskin, Phys. Rev. **D60**, 014015 (1999).

[14] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C23**, 73 (2002).

[15] S. Riemersma, J. Smith and W. L. van Neerven, Phys. Lett. **B347**, 143 (1995).

[16] E. M. Levin, A. D. Martin, M. G. Ryskin and T. Teubner, Z. Phys. **C74**, 671 (1997).

[17] M. Gluck, E. Reya and A. Vogt, Z. Phys. **C53**, 651 (1992).

[18] S. Chekanov *et al.* [ZEUS Collaboration], Nucl. Phys. **B672**, 3 (2003).

[19] P. Newman, these proceedings.

DPDF: A Library for Diffractive Parton Distributions

Frank-Peter Schilling
CERN/PH, CH-1211 Geneva 23, Switzerland

Abstract
A code library is presented which provides a common interface to available parameterizations of diffractive parton distribution functions determined from QCD fits to HERA diffractive structure function data.

1 Introduction

In recent years, various precise measurements of the diffractive reduced cross section[1] $\sigma_r^{D(3)}(x_{I\!P}, \beta, Q^2)$ have been made by the HERA experiments H1 and ZEUS. Within the framework of QCD factorization in diffractive DIS [1], several sets of *diffractive parton distributions* (dpdf's) have been obtained from leading or next-to-leading order DGLAP QCD fits to these data[2]. The extracted dpdf's are a crucial input for the calculations of the cross sections of less inclusive diffractive processes such as diffractive jet, heavy quark or even Higgs production.

Since these diffractive pdf's are used in many different Monte-Carlo generators as well as in fixed order QCD calculations, it is desirable to provide them through a common software interface, similar in spirit to the common PDFLIB [2] and LHAPDF [3] packages for non-diffractive pdf's. To achieve this, the DPDF library has been developed. When a new dpdf set becomes available, it then needs to be implemented only in one place. Furthermore, additional features such as custom QCD evolution, structure function calculation and error information are available. Thus, the DPDF library provides a useful way to make the knowledge from HERA available to the TEVATRON, LHC and theory communities.

2 Theoretical Framework

The concept of QCD factorization in diffractive DIS implies that the diffractive $\gamma^* p$ cross section can be expressed as a convolution of universal diffractive parton distributions f_i^D with process-dependent coefficient functions:

$$\frac{\mathrm{d}^2\sigma(x, Q^2, x_{I\!P}, t)^{\gamma^* p \to p' X}}{\mathrm{d}x_{I\!P}\,\mathrm{d}t} = \sum_i \int_x^{x_{I\!P}} \mathrm{d}\xi\, \hat{\sigma}^{\gamma^* i}(x, Q^2, \xi)\, f_i^D(x_{I\!P}, t, \xi, Q^2)\ . \tag{1}$$

The diffractive pdf's $f_i^D(x_{I\!P}, t, \beta, Q^2)$ can be extracted from a DGLAP QCD analysis of the diffractive reduced cross section σ_r^D.

For many (but not all) of the included parameterizations the $(x_{I\!P}, t)$ dependence factorizes ("Regge factorization") so that a *flux factor* $f_{I\!P/p}(x_{I\!P}, t)$ and dpdf's $f_i^{I\!P}(\beta, Q^2)$ are defined separately:

$$f_i^D(x_{I\!P}, t, \beta, Q^2) = f_{I\!P/p}(x_{I\!P}, t) \cdot f_i^{I\!P}(\beta, Q^2)\ . \tag{2}$$

For those parameterizations which include a secondary Reggeon exchange contribution (often using a pion structure function) in order to describe the data at high $x_{I\!P}$, such a possibility is also included. The dpdf's are typically parameterized in terms of a light quark flavor singlet and a gluon distribution, which are evolved using the (N)LO DGLAP equations[3].

[1]The reduced cross section σ_r^D corresponds to the structure function F_2^D if contributions from F_L^D and xF_3^D are neglected.
[2]In some cases, final state data were used in addition in order to better constrain the diffractive gluon density.
[3]For details of the parameterizations, see the original publications.

Set	Fit	Var	Name	Ref.	Order	Q^2 (GeV^2)	β	$x_{\mathbb{P}}$
1	4	-	H1-1997-LO-Fit-1	H1 Coll. [5]	LO	$4.5\ldots75$	$0.04\ldots0.9$	< 0.05
1	5	-	H1-1997-LO-Fit-2		LO			
1	6	-	H1-1997-LO-Fit-3		LO			
2	1	-	H1-2002-NLO	H1 Coll. (prel.) [6]	NLO	$6.5\ldots800$	$0.01\ldots0.9$	< 0.05
2	2	-	H1-2002-LO		LO			
3	1	1…3	ACTW-NLO-A	Alvero et al. [7]	NLO	$6.0\ldots75$	$0.20\ldots0.7$	< 0.01
3	2	1…3	ACTW-NLO-B		NLO			
3	3	1…3	ACTW-NLO-C		NLO			
3	4	1…3	ACTW-NLO-D		NLO			
3	5	1…3	ACTW-NLO-SG		NLO			
4	-	-	BGH-LO	Buchmueller et al. [8]	LO	$4.5\ldots75$	$0.04\ldots0.9$	< 0.01
5	-	-	HS-NLO	Hautmann and Soper [9]	NLO			
6	-	-	ZEUS-LPS	ZEUS Coll. [11, 12]	NLO	$2.4\ldots39$	$0.01\ldots0.5$	< 0.01
7	1	-	MRW-NLO-Lambda	Martin et al. [10]	NLO	$2.4\ldots90$	$0.01\ldots0.9$	< 0.05
7	2	-	MRW-NLO-MRST		NLO			
8	-	-	ZEUS-MX	Groys et al. [13]	NLO	$4.0\ldots55$	$0.01\ldots0.9$	< 0.01

Table 1: Overview of the diffractive pdf sets implemented in the DPDF package. The Q^2, β and $x_{\mathbb{P}}$ ranges correspond to the approximate kinematic range of the data used in the fit.

3 Implementation

DPDF is a FORTRAN 77 package. A C++ wrapper will be provided. There is an external dependency on the QCDNUM [4] package, which can be disabled.

3.1 Available Parameterizations

Currently the following dpdf sets are implemented: the fits performed by the H1 collaboration in [5], the preliminary H1 fits presented in [6], the fits by Alvero et al. [7], a parameterization of the semi-classical model by Buchmueller et al. [8], the fits by Hautmann and Soper [9] and by Martin et al. [10], the ZEUS fit from [11, 12] and a fit to recent ZEUS data presented at this workshop [13, 12].

Details of the available dpdf sets are presented in table 1, including the kinematic ranges of the data which were included in the fits. This information can be used as a guideline for the range of validity of the fits. Note in particular that typically only data for $x_{\mathbb{P}} < 0.05$ or < 0.01 are included in the fits, which introduces an additional uncertainty when these fits are used for comparisons with experimental data at higher $x_{\mathbb{P}}$.

3.2 Interface to QCDNUM

DPDF provides an interface to the NLO DGLAP QCD evolution package QCDNUM [4]. It is possible to perform a QCD evolution of the given pdf set from its starting scale Q_0^2 using either the original evolution scheme and parameters, or by providing modified parameters. The benefits are:

- QCDNUM calculates the full (N)LO structure functions F_2 and F_L for light and heavy quarks, which can be used for consistent comparisons with experimental data;
- The QCD evolution parameters such as α_s can be varied for systematic studies;
- The dpdf's can be evolved to Q^2 or β values beyond the grid on which the original parameterization is provided, which is particularly interesting for LHC applications.

3.3 Usage

The DPDF package can be obtained from [14]. In the following the principal user subroutines of the library are listed.

- The package is initialized for a given dpdf by calling `dpdf_init(iset,ifit,ivar)` where `iset`, `ifit` and `ivar` are the parameters as given in table 1.
- The diffractive proton pdf's for either Pomeron or sub-leading Reggeon exchange or their sum (if provided) are returned at given values of $(x_{I\!P}, t, \beta, Q^2)$ in an array `xpq(-6:6)` using `dpdf_ppdf`. The result may also be integrated over t.
- If provided, the flux factors $f_{I\!P}(x_{I\!P}, t)$ and the parton densities $f_i(\beta, Q^2)$ can be obtained separately from `dpdf_flux` and `dpdf_pdf`.
- The diffractive structure function can be obtained from `dpdf_f2d`.
- QCD evolution using QCDNUM can be performed using default parameters for the given set with `dpdf_evolve_std` and using modified evolution parameters with `dpdf_evolve`.

Note that the details of the user interface may change in the future. For details refer to the user manual available from [14].

4 Outlook

It is planned to update DPDF if new dpdf sets become available. Additional features such as the possibility of error dpdf's (as for LHAPDF) are foreseen. The code and manual are available from [14].

References

[1] J.C. Collins, Phys. Rev. **D57**, 3051 (1998).

[2] H. Plothow-Besch, Comput. Phys. Commun. **75**, 396 (1993).

[3] S. Alekhin, W. Giele and J. Pumplin, *The lhapdf interface*, in *The QCD/SM working group: Summary report, Physics at TeV Colliders II Workshop, Les Houches, France*. 2002. `http://durpdg.dur.ac.uk/lhapdf`.

[4] M. Botje, *QCDNUM version 16.12*. ZEUS-97-066. `http://www.nikhef.nl/~h24/qcdnum`.

[5] C. Adloff et al. [H1 Collaboration], Z. Phys. **C76**, 613 (1997).

[6] H1 Collaboration, *Measurement and NLO DGLAP QCD Interpretation of Diffractive Deep-Inelastic Scattering at HERA*. Paper 980 contributed to ICHEP 2002, Amsterdam.

[7] L. Alvero, J.C. Collins, J. Terron and J.J. Whitmore, Phys. Rev. **D59**, 074022 (1999).

[8] W. Buchmuller, T. Gehrmann and A. Hebecker, Nucl. Phys. **B537**, 477 (1999).

[9] F. Hautmann and D.E. Soper, Phys. Rev. **D63**, 011501 (2000).

[10] A.D. Martin, M.G. Ryskin and G. Watt, Eur. Phys. J. **C37**, 285 (2004).

[11] S. Chekanov et al. [ZEUS Collaboration], Eur. Phys. J. **C38**, 43 (2004).

[12] A. Proskuryakov. Private communucation.

[13] M. Groys, A. Levy and A. Proskuryakov, *Diffractive Parton Distributions from the HERA Data*. These proceedings.

[14] F.-P. Schilling, *DPDF web page*. `http://fpschill.home.cern.ch/fpschill/dpdf`.

Prospects for F_L^D Measurements at HERA-II

Paul Newman
School of Physics and Astronomy, University of Birmingham, B15 2TT, United Kingdom

Abstract
The theoretical interest in the longitudinal diffractive structure function F_L^D is briefly motivated and possible measurement methods are surveyed. A simulation based on realistic scenarios with a reduced proton beam energy at HERA-II using the H1 apparatus shows that measurements are possible with up to 4σ significance, limited by systematic errors.

1 Introduction

In order to understand inclusive diffraction fully, it is necessary to separate out the contributions from transversely and longitudinally polarised exchange photons. Here, the formalism of [1] is adopted, where by analogy with inclusive scattering and neglecting weak interactions, a reduced cross section σ_r^D is defined,[1] related to the experimentally measured cross section by

$$\frac{d^3\sigma^{ep\to eXY}}{\mathrm{d}x_{I\!P}\,\mathrm{d}\beta\,\mathrm{d}Q^2} = \frac{2\pi\alpha^2}{\beta\,Q^4}\cdot Y_+ \cdot \sigma_r^D(x_{I\!P},\beta,Q^2)\,, \quad \text{where} \quad \sigma_r^D = F_2^D - \frac{y^2}{Y_+}F_L^D \tag{1}$$

and $Y_+ = 1 + (1-y)^2$. The structure function F_L^D, is closely related to the longitudinal photon contribution, whereas the more familiar F_2^D contains information on the sum of transverse and longitudinal photon contributions.

It is generally understood [2] that at high β and low-to-moderate Q^2, σ_r^D receives a significant, perhaps dominant, higher twist contribution due to longitudinally polarised photons. Definite predictions [3] exist for this contribution, obtained by assuming 2-gluon exchange, with a similar phenomenology to that successfully applied to vector meson cross sections at HERA. The dominant role played by gluons in the diffractive parton densities [1] implies that the leading twist F_L^D must also be relatively large. Assuming the validity of QCD hard scattering collinear factorisation [4], this gluon dominance results in a leading twist F_L^D which is approximately proportional to the diffractive gluon density. A measurement of F_L^D to even modest precision would provide a very powerful independent tool to verify our understanding of the underlying dynamics and to test the gluon density extracted indirectly in QCD fits from the scaling violations of F_2^D. This is particularly important at the lowest x values, where direct information on the gluon density cannot be obtained from jet or D^* data due to kinematic limitations and where novel effects such as parton saturation or non-DGLAP dynamics are most likely to become important.

Several different methods have been proposed to extract information on F_L^D. It is possible in principle to follow the procedure adopted by H1 in the inclusive case [5, 6], exploiting the decrease in σ_r^D at large y relative to expectations for F_2^D alone (see equation 1). This method may yield significant results if sufficient precision and y range can be achieved [7], though assumptions are required on the $x_{I\!P}$ dependence of F_2^D, which is currently not well constrained by theory. An alternative method, exploiting the azimuthal decorrelation between the proton and electron scattering planes caused by interference between the transverse and longitudinal photon contributions [8], has already been used with the scattered proton measured in the ZEUS LPS [9]. However, due to the relatively poor statistical precision achievable with Roman pots at HERA-I, the current results are consistent with zero. If the potential of the H1 VFPS is fully realised, this method may yet yield significant results in the HERA-II data [10]. However, if

[1] It is assumed here that all results are integrated over t. The superscript (3) usually included for $F_2^{D(3)}$ and other quantities is dropped for convenience.

the necessary data are taken, the most promising possibility is to extract F_L^D by comparing data at the same Q^2, β and $x_{I\!P}$, but from different centre of mass energies $\sqrt{s}$ and hence from different y values. The longitudinal structure function can then be extracted directly and model-independently from the measured data using equation 1. In this contribution, one possible scenario is investigated, based on modified beam energies and luminosities which are currently under discussion as a possible part of the HERA-II programme.

2 Simulated F_L^D Measurement

Given the need to obtain a large integrated luminosity at the highest possible beam energy for the remainder of the HERA programme and the fixed end-point in mid 2007, it is likely that only a relatively small amount of data can be taken with reduced beam energies. A possible scenario is investigated here in which 10 pb^{-1} are taken at just one reduced proton beam energy of $E_p = 400$ GeV, the electron beam energy being unchanged at 27.5 GeV. Since the maximum achievable instantaneous luminosity at HERA scales like the proton beam energy squared [11], this data sample could be obtained in around 2-3 months at the current level of HERA performance. It is assumed that a larger data volume of $100\ \mathrm{pb}^{-1}$ is available at $E_p = 920$ GeV, which allows for downscaling of high rate low Q^2 inclusive triggers.[2] The results presented here can be used to infer those from other scenarios given that the statistical uncertainty scales like $\sigma_r^{D\,400}/\sqrt{\mathcal{L}_{400}} + \sigma_r^{D\,920}/\sqrt{\mathcal{L}_{920}}$, where $\sigma_r^{D\,E_p}$ and $\mathcal{L}_{E_p}$ are the reduced cross section and the luminosity at a proton beam energy of E_p, respectively.

The longitudinal structure function can be extracted from the data at the two beam energies using

$$F_L^D = \frac{Y_+^{400}\,Y_+^{920}}{y_{400}^2 Y_+^{920} - y_{920}^2 Y_+^{400}} \left(\sigma_r^{D\,920} - \sigma_r^{D\,400}\right) , \tag{2}$$

where y_{E_p} and $Y_+^{E_p}$ denote y and Y_+ at a beam energy E_p. It is clear from equation 2 that the best sensitivity to F_L^D requires the maximum difference between the reduced cross sections at the two beam energies, which (equation 1) implies the maximum possible y at $E_p = 400$ GeV. By measuring scattered electrons with energies E_e' as low as 3 GeV [5], the H1 collaboration has obtained data at $y = 0.9$. This is possible with the use of the SPACAL calorimeter in combination with a measurement of the electron track in either the backward silicon tracker (BST) or the central jet chamber (CJC). For HERA-II running, the corresponding available range of scattered electron polar angle is $155^\circ < \theta_e' < 173^\circ$, which is used in the current study.[3] Three intervals in y are considered, corresponding at $E_p = 400$ GeV to $0.5 < y_{400} < 0.7$, $0.7 < y_{400} < 0.8$ and $0.8 < y_{400} < 0.9$. It is ensured that identical ranges in β, $x_{I\!P}$ and Q^2 are studied at $E_p = 920$ GeV by choosing the bin edges such that $y_{920} = y_{400} \cdot 400/920$. Since the highest possible precision is required in this measurement, the restriction $x_{I\!P} < 0.02$ is imposed, which leads to negligible acceptance losses with a typical cut on the forwardmost extent of the diffractive system $\eta_{\max} < 3.3$. The kinematic restrictions on E_e', θ_e' and $x_{I\!P}$ lead to almost no change in the mean Q^2, M_{X}^2 or $\beta \simeq Q^2/(Q^2 + M_{\mathrm{X}}^2)$ as either y or E_p are varied. In contrast, $x_{I\!P} = Q^2/(s\,y\,\beta)$ varies approximately as $1/y$. As is shown in figure 1, at the average $\beta = 0.23$, there is at least partial acceptance for all y bins in the range $7 < Q^2 < 30\ \mathrm{GeV}^2$, which is chosen for this study, leading to an average value of Q^2 close to 12 GeV^2.

The simulation is performed using the RAPGAP [13] Monte Carlo generator to extract the number of events per unit luminosity in each bin at each centre of mass energy. The values of F_2^D and F_L^D, and hence $\sigma_r^{D\,920}$ and $\sigma_r^{D\,400}$ are obtained using an updated version of the preliminary H1 2002 NLO QCD fit [1].

[2]Alternative scenarios in which a smaller data volume at large E_p is taken in a short, dedicated run, could potentially lead to better controlled systematics at the expense of increased statistical errors.

[3]One interesting alternative running scenario [12] is to obtain data at $E_p = 920$ GeV with the vertex shifted by 20 cm in the outgoing proton direction, which would allow measurements up to $\theta_e' = 175^\circ$, giving a low Q^2 acceptance range which closely matches that for the $E_p = 400$ GeV data at the normal vertex position.

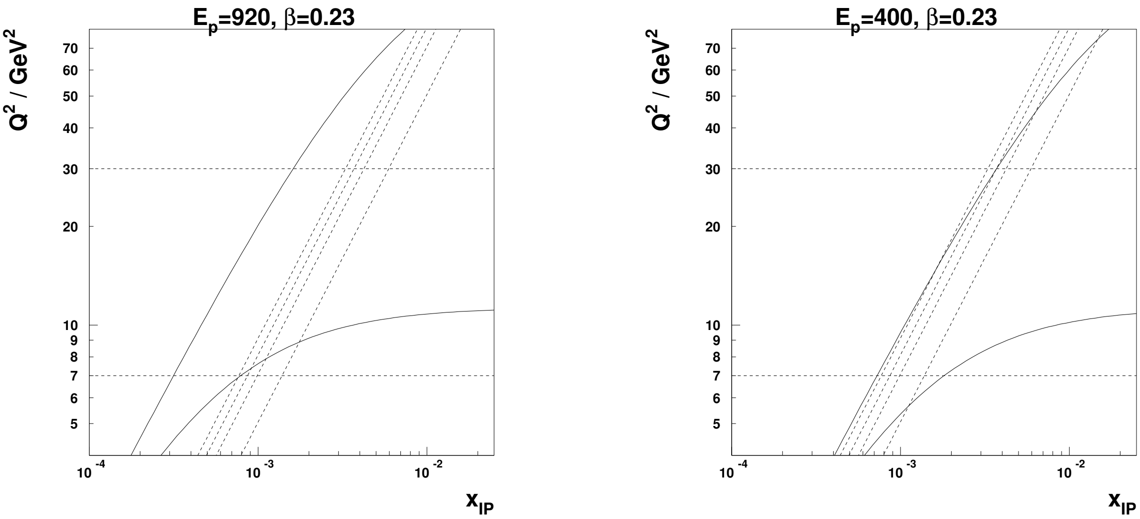


Fig. 1: Illustration of the kinematic plane in Q^2 and $x_{I\!P}$ at proton energies of 920 GeV and 400 GeV, with fixed $\beta = x/x_{I\!P} = 0.23$. The solid lines illustrate the experimental limits of $155^\circ < \theta'_e < 173^\circ$. The horizontal dashed lines illustrate the Q^2 range used for the simulation. The diagonal dashed lines illustrate the binning in y, corresponding at $E_p = 400$ GeV to $y = 0.9$ (leftmost line), $y = 0.8$, $y = 0.7$ and $y = 0.5$ (rightmost line).

The expected precision on F_L^D is obtained by error propagation through equation 2. The systematic uncertainties are estimated on the basis of previous experience with the H1 detector [1, 5]. At the large y values involved, the kinematic variables are most accurately reconstructed using the electron energy and angle alone. The systematic uncertainties on the measurements of these quantities are assumed to be correlated between the two beam energies. With the use of the BST and CJC, the possible bias in the measurement of θ'_e is at the level of $0.2\ \mathrm{mrad}$. The energy scale of the SPACAL calorimeter is known with a precision varying linearly from 2% at $E'_e = 3$ GeV to 0.2% at $E'_e = 27.5$ GeV. Other uncertainties which are correlated between the two beam energies arise from the photoproduction background subtraction (important at large y and assumed to be known with a precision of 25%) and the energy scale for the hadronic final state used in the reconstruction of M_X and hence $x_{I\!P}$ (taken to be known to 4%, as currently). Sources of uncertainty which are assumed to be uncorrelated between the low and high E_p measurements are the luminosity measurement (taken to be $\pm1\%$), the trigger and electron track efficiencies ($\pm1\%$ combined) and the acceptance corrections, obtained using RAPGAP ($\pm2\%$). The combined uncorrelated error is thus 2.4%. Finally, a normalisation uncertainty of $\pm6\%$ due to corrections for proton dissociation contributions is taken to act simultaneously in the two measurements. Other sources of uncertainty currently considered in H1 measurements of diffraction are negligible in the kinematic region studied here.

y_{400}	y_{920}	$x_{I\!P}$	F_2^D	F_L^D	$\delta_{\rm unc}$	$\delta_{\rm norm}$	$\delta E'_e$	$\delta\theta'_e$	δM_X	$\delta\gamma p$	$\delta_{\rm syst}$	$\delta_{\rm stat}$	$\delta_{\rm tot}$
$0.5-0.7$	0.217 - 0.304	0.0020	15.72	3.94	34	6	8	2	7	0	36	20	41
$0.7-0.8$	0.304 - 0.348	0.0016	20.87	5.25	19	6	3	2	5	6	22	17	28
$0.8-0.9$	0.348 - 0.391	0.0014	24.47	6.16	14	6	6	1	2	13	21	13	25

Table 1: Summary of the simulation at $Q^2 = 12$ GeV and $\beta = 0.23$. The first three columns contain the y ranges used at $E_p = 400$ GeV and $E_p = 920$ GeV and the $x_{I\!P}$ values. The next two columns contain the values of the diffractive structure functions. These are followed by the uncorrelated ($\delta_{\rm unc}$) and proton dissociation ($\delta_{\rm norm}$) uncertainties and the correlated systematics due to the electron energy ($\delta E'_e$) and angle ($\delta\theta'_e$) measurements, the hadronic energy scale (δM_X) and the photoproduction background ($\delta\gamma p$), all in percent. The last three columns summarise the systematic, statistical and total uncertainties.

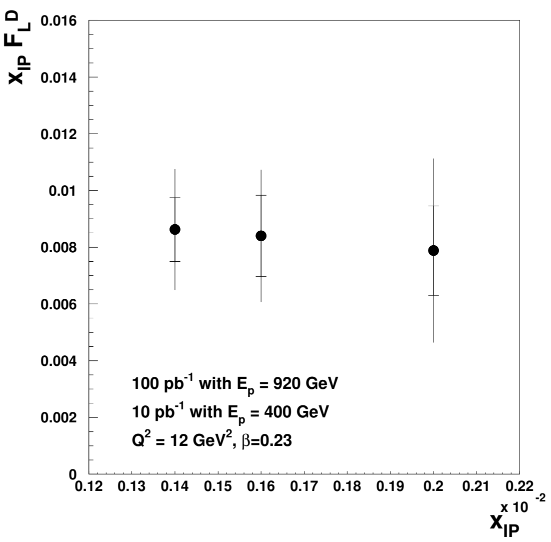


Fig. 2: Illustration of the simulated result for F_L^D, showing the three data points with statistical (inner bars) and total (outer bars) errors.

Full details of the simulated uncertainties on the F_L^D measurements are given in table 1. An illustration of the corresponding expected measurement, based on the F_L^D from the H1 2002 fit is shown in figure 2. The most precise measurement is obtained at the highest y, where F_L^D would be determined to be unambiguously different from its maximum value of F_2^D and to be non-zero at the 4σ level. Two further measurements are obtained at lower y values. The dominant errors arise from statistical uncertainties and from uncertainties which are uncorrelated between the two beam energies. Minimising the latter is a major experimental challenge to be addressed in the coming years.

Only one possible scenario has been investigated here, leading to a highly encouraging result at relatively low β, which would provide a very good test of the leading twist F_L^D and thus of the gluon density extracted in QCD fits to F_2^D. It may also be possible to obtain results at high β, giving information on the higher twist contributions in that region, for example by restricting the analysis to lower $x_{I\!P}$.

Acknowledgements

For comments, corrections, cross-checks and code, thanks to Markus Diehl, Joel Feltesse, Max Klein and Frank-Peter Schilling!

References

[1] H1 Collaboration, paper 980 contributed to ICHEP 2002, Amsterdam (H1prelim-02-012).

[2] J. Bartels, J. Ellis, H. Kowalski, M. Wusthoff, Eur. Phys. J. **C7**, 443 (1999).

[3] A. Hebecker, T. Teubner, Phys. Lett. **B498**, 16 (2001).

[4] J. Collins, Phys. Rev. **D57**, 3051 (1998). Erratum ibid **D61** (2000) 019902.

[5] C. Adloff et al. [H1 Collaboration], Eur. Phys. J. **C21**, 33 (2001).

[6] H1 Collaboration, paper 4-083 contributed to EPS 2003, Aachen (H1prelim-03-043).

[7] A. Mehta, J. Phillips, B. Waugh, proceedings of the workshop ‘Future Physics at HERA’, DESY, 1996, ed. Ingelman, de Roeck, Klanner, pg 704.

[8] M. Diehl, proceedings of the XIth International Conference on Elastic and Diffractive Scattering, Blois, France, 2005 [hep-ph/0509107].

[9] S. Chekanov et al. [ZEUS Collaboration], Eur. Phys. J. **C38**, 43 (2004).

[10] L. Favart et al., ‘Proposal for Installation of a Very Forward Proton Spectrometer in H1 after 2000’, DESY-PRC 01/00.

[11] F. Willeke, ‘Issues for HERA Future Operation’, talk at the Workshop on the Future of Lepton-Nucleon Scattering, Durham, UK, 2001.

[12] M. Klein. Proceedings of the 12th International Workshop on Deep Inelastic Scattering (DIS 2004), Strbske Pleso, Slovakia, 2004.

[13] H. Jung, Comp. Phys. Commun. **86**, 147 (1995). See also http://www.desy.de/ jung/rapgap.html.

Diffractive Dijet Production at HERA

A. Bruni[1], *M. Klasen*[2,3], *G. Kramer*[3] *and S. Schätzel*[4]
[1] INFN Bologna, Via Irnerio 46, 40156 Bologna, Italy
[2] Laboratoire de Physique Subatomique et de Cosmologie, Université Joseph Fourier/CNRS-IN2P3, 53 Avenue des Martyrs, 38026 Grenoble, France
[3] II. Inst. für Theoret. Physik, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany
[4] DESY FLC, Notkestr. 85, 22607 Hamburg, Germany

Abstract
We present recent experimental data from the H1 and ZEUS Collaborations at HERA for diffractive dijet production in deep-inelastic scattering (DIS) and photoproduction and compare them with next-to-leading order (NLO) QCD predictions using diffractive parton densities. While good agreement is found for DIS, the dijet photoproduction data are overestimated by the NLO theory, showing that factorization breaking occurs at this order. While this is expected theoretically for resolved photoproduction, the fact that the data are better described by a global suppression of direct *and* resolved contribution by about a factor of two comes as a surprise. We therefore discuss in some detail the factorization scheme and scale dependence between direct and resolved contributions and propose a new factorization scheme for diffractive dijet photoproduction.

1 Introduction

It is well known that in high-energy deep-inelastic ep-collisions a large fraction of the observed events are diffractive. These events are defined experimentally by the presence of a forward-going system Y with four-momentum p_Y, low mass M_Y (in most cases a single proton and/or low-lying nucleon resonances), small momentum transfer squared $t = (p - p_Y)^2$, and small longitudinal momentum transfer fraction $x_{I\!P} = q(p - p_Y)/qp$ from the incoming proton with four-momentum p to the system X (see Fig. 1). The presence of a hard scale, as for example the photon virtuality $Q^2 = -q^2$ in deep-inelastic scattering (DIS) or the large transverse jet momentum p_T^* in the photon-proton centre-of-momentum frame, should then allow for calculations of the production cross section for the central system X with the known methods of perturbative QCD. Under this assumption, the cross section for the inclusive production of two jets, $e + p \to e + 2\text{ jets} + X' + Y$, can be predicted from the well-known formulæ for jet production in

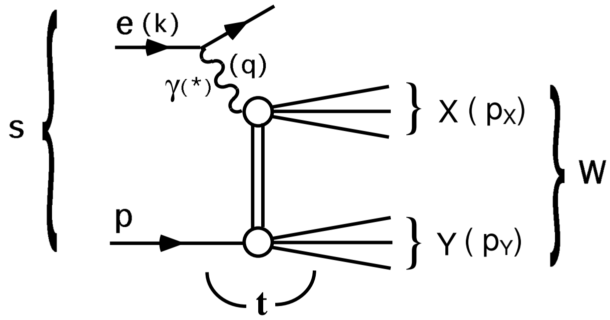


Fig. 1: Diffractive scattering process $ep \to eXY$, where the hadronic systems X and Y are separated by the largest rapidity gap in the final state.

non-diffractive ep collisions, where in the convolution of the partonic cross section with the parton distribution functions (PDFs) of the proton the latter ones are replaced by the diffractive PDFs. In the simplest approximation, they are described by the exchange of a single, factorizable pomeron/Regge-pole.

The diffractive PDFs have been determined by the H1 Collaboration at HERA from high-precision inclusive measurements of the DIS process $ep \to eXY$ using the usual DGLAP evolution equations in leading order (LO) and next-to-leading order (NLO) and the well-known formula for the inclusive cross section as a convolution of the inclusive parton-level cross section with the diffractive PDFs [1]. For a similar analysis of the inclusive measurements of the ZEUS Collaboration see [2, 3]. A longer discussion of the extraction of diffractive PDFs can also be found in these proceedings [4] and in [5]. For inclusive diffractive DIS it has been proven by Collins that the formula referred to above is applicable without additional corrections and that the inclusive jet production cross section for large Q^2 can be calculated in terms of the same diffractive PDFs [6]. The proof of this factorization formula, usually referred to as the validity of QCD factorization in hard diffraction, may be expected to hold for the direct part of photoproduction ($Q^2 \simeq 0$) or low-Q^2 electroproduction of jets [6]. However, factorization does not hold for hard processes in diffractive hadron-hadron scattering. The problem is that soft interactions between the ingoing two hadrons and their remnants occur in both the initial and final state. This agrees with experimental measurements at the Tevatron [7]. Predictions of diffractive dijet cross sections for $p\bar{p}$ collisions as measured by CDF using the same PDFs as determined by H1 [1] overestimate the measured cross section by up to an order of magnitude [7]. This suppression of the CDF cross section can be explained by considering the rescattering of the two incoming hadron beams which, by creating additional hadrons, destroy the rapidity gap [8].

Processes with real photons ($Q^2 \simeq 0$) or virtual photons with fixed, but low Q^2 involve direct interactions of the photon with quarks from the proton as well as resolved photon contributions, leading to parton-parton interactions and an additional remnant jet coming from the photon (for a review see [9]). As already said, factorization should be valid for direct interactions as in the case of DIS, whereas it is expected to fail for the resolved process similar as in the hadron-hadron scattering process. In a two-channel eikonal model similar to the one used to calculate the suppression factor in hadron-hadron processes [8], introducing vector-meson dominated photon fluctuations, a suppression by about a factor of three for resolved photoproduction at HERA is predicted [10]. Such a suppression factor has recently been applied to diffractive dijet photoproduction [11, 12] and compared to preliminary data from H1 [13] and ZEUS [14]. While at LO no suppression of the resolved contribution seemed to be necessary, the NLO corrections increase the cross section significantly, showing that factorization breaking occurs at this order at least for resolved photoproduction and that a suppression factor R must be applied to give a reasonable description of the experimental data.

As already mentioned elsewhere [11, 12], describing the factorization breaking in hard photoproduction as well as in electroproduction at very low Q^2 [15] by suppressing the resolved contribution only may be problematic. An indication for this is the fact that the separation between the direct and the resolved process is uniquely defined only in LO. In NLO these two processes are related. The separation depends on the factorization scheme and the factorization scale M_γ. The sum of both cross sections is the only physically relevant cross section, which is approximately independent of the factorization scheme and scale [16]. As demonstrated in Refs. [11, 12] multiplying the resolved cross section with the suppression factor $R = 0.34$ destroys the correlation of the M_γ-dependence between the direct and resolved part, and the sum of both parts has a stronger M_γ-dependence than for the unsuppressed case ($R = 1$), where the M_γ-dependence of the NLO direct cross section is compensated to a high degree against the M_γ-dependence of the LO resolved part.

In the second Section of this contribution, we present the current experimental data from the H1 and ZEUS Collaborations on diffractive dijet production in DIS and photoproduction and compare these data to theoretical predictions at NLO for two different scenarios: suppression of only the resolved part by a factor $R = 0.34$ as expected from LO theory and proposed in [8], and equal suppression of all direct and resolved contributions by a factor $R = 0.5$, which appears to describe the data better phenomenologically. This motivates us to investigate in the third Section the question whether certain parts of the direct contribution might break factorization as well and therefore need a suppression factor.

The introduction of the resolved cross section is dictated by perturbation theory. At NLO, collinear singularities arise from the photon initial state, which are absorbed at the factorization scale into the photon PDFs. This way the photon PDFs become M_γ-dependent. The equivalent M_γ-dependence, just with the opposite sign, is left in the NLO corrections to the direct contribution. With this knowledge, it is obvious that we can obtain a physical cross section at NLO, *i.e.* the superposition of the NLO direct and LO resolved cross section, with a suppression factor $R < 1$ and no M_γ-dependence left, if we also multiply the $\ln M_\gamma$-dependent term of the NLO correction to the direct contribution with the same suppression factor as the resolved cross section. We are thus led to the theoretical conclusion that, contrary to what one may expect, not *all* parts of the direct contribution factorize. Instead, the *initial state* singular part appearing beyond LO breaks factorization even in direct photoproduction, presumably through soft gluon attachments between the proton and the collinear quark-antiquark pair emerging from the photon splitting. This would be in agreement with the general remarks about initial state singularities in Ref. [6].

In the third Section of this contribution, we present the special form of the $\ln M_\gamma$-term in the NLO direct contribution and demonstrate that the M_γ-dependence of the physical cross section cancels to a large extent in the same way as in the unsuppressed case ($R = 1$). These studies can be done for photoproduction ($Q^2 \simeq 0$) as well as for electroproduction with fixed, small Q^2. Since in electroproduction the initial-state singularity in the limit $Q^2 \to 0$ is more directly apparent than for the photoproduction case, we shall consider in this contribution the low-Q^2 electroproduction case just for demonstration. This diffractive dijet cross section has been calculated recently [15]. A consistent factorization scheme for low-Q^2 virtual photoproduction has been defined and the full (direct and resolved) NLO corrections for inclusive dijet production have been calculated in [17]. In this work we adapt this inclusive NLO calculational framework to diffractive dijet production at low-Q^2 in the same way as in [15], except that we multiply the $\ln M_\gamma$-dependent terms as well as the resolved contributions with the same suppression factor $R = 0.34$, as an example, as in our earlier work [11, 12, 15]. The exact value of this suppression factor may change in the future, when better data for photoproduction and low-Q^2 electroproduction have been analyzed. We present the $\ln M_\gamma$-dependence of the partly suppressed NLO direct and the fully suppressed NLO resolved cross section $\mathrm{d}\sigma/\mathrm{d}Q^2$ and their sum for the lowest Q^2 bin, before we give a short summary in section 4.

2 Comparison of H1 and ZEUS Data with NLO Theory Predictions

In this Section, diffractive PDFs [1, 2, 3] extracted from diffractive structure function data are used in NLO calculations to test factorisation in diffractive dijet production. Dijet production is directly sensitive to the diffractive gluon (Fig. 2) whereas in inclusive measurements the gluon is determined from scaling violations.

2.1 Diffractive Dijet Production in DIS

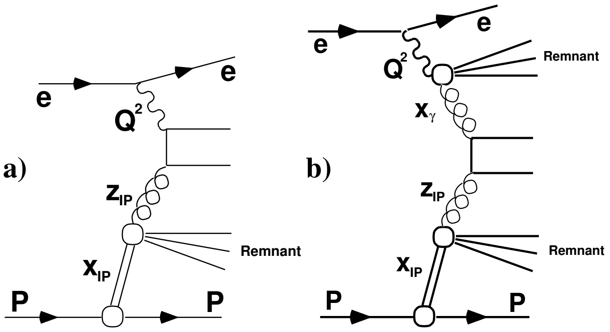


Fig. 2: Example processes for a) direct photon and b) resolved photon interactions.

H1 has measured the cross sections for dijet production [13] in the kinematic range $Q^2 > 4$ GeV^2, $165 < W < 242$ GeV (photon-proton centre-of-mass energy) and $x_{I\!P} < 0.03$. Jets are identified using the inclusive k_T cluster algorithm and selected by requiring $E_T^{*,\text{jet}}(1,2) > 5,4$ GeV and $-3 < \eta^*_{\text{jet}} < 0$.[1] NLO predictions have been obtained by interfacing the H1 diffractive PDFs with the DISENT program [18]. The renormalisation and factorisation scales were set to the transverse energy of the leading parton jet. The NLO parton jet cross sections have been corrected for hadronisation effects using the leading order (LO) generator RAPGAP [19] with parton showers and the Lund string fragmentation model. Comparisons of the DISENT and RAPGAP predictions with the measured cross section differential in $z_{I\!P}^{\text{jets}}$, an estimator for the fraction of the momentum of the diffractive exchange entering the hard scatter, are shown in Fig. 3a. The inner band around the NLO calculation indicates

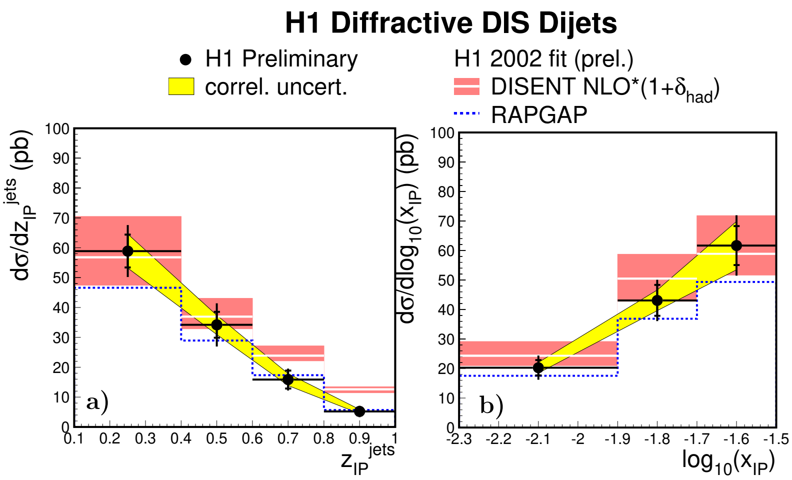


Fig. 3: Diffractive DIS dijet cross sections compared with a NLO prediction based on diffractive PDFs and with RAPGAP.

the $\approx 20\%$ uncertainty resulting from a variation of the renormalisation scale by factors 0.5 and 2. The uncertainty in the diffractive PDFs is not shown. Within this additional uncertainty, which is large at high $z_{I\!P}^{\text{jets}}$, the cross section is well described. The cross section differential in $\log_{10}(x_{I\!P})$, p_T^{jet1}, and Q^2 is shown in Fig. 3b and Fig. 4. All distributions are well described and QCD factorisation is therefore in good agreement with dijet production in diffractive DIS.

Similar results are presented by ZEUS [20]; the dijet cross sections have been measured in the kinematic range $5 < Q^2 < 100$ GeV^2, $100 < W < 200$ GeV, $x_{I\!P} < 0.03$. The jets were identified using the inclusive k_T algorithm in the γp frame and required to satisfy $E_T^{*,\text{jet}}(1,2) > 5,4$ GeV and $-3.5 < \eta^*_{\text{jet}} < 0.0$. NLO predictions have been obtained with the DISENT program interfaced to three different sets of diffractive PDFs: from fit to H1 data [1], from fit to the ZEUS M_X data (GLP) [3] and from fit to ZEUS LPS and $F_2^{D,charm}$ data [2]. Comparisons of the DISENT predictions with the measured cross section differential in $E_T^{*,\text{jet}}$, η^*_{jet}, $z_{I\!P}^{\text{jets}}$ and x_γ^{obs} are shown in Fig. 5. The $20-30\%$ uncertainty in the NLO calculations resulting from a variation of the renormalisation and factorisation scales is not shown. Within the experimental and QCD scale uncertainties, the predictions based on the H1 and ZEUS-LPS PDFs give a good description of the dijet cross section. The normalisation of

[1] The '$*$' denotes variables in the photon-proton centre-of-mass system.

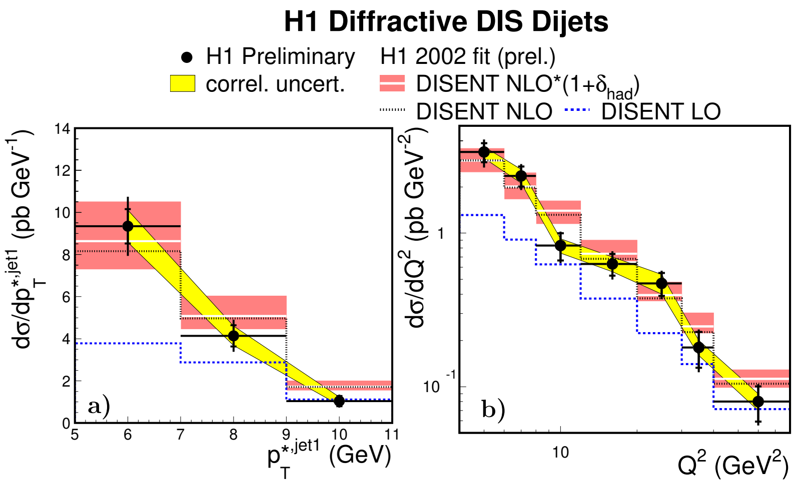


Fig. 4: Diffractive DIS dijet cross sections compared with a NLO prediction based on diffractive PDFs.

the prediction using the GLP fit is substantially lower than those from the other two sets of PDFs. For ZEUS, the difference observed between the three sets may be interpreted as an estimate of the uncertainty associated with the diffractive PDFs and with the definition of the diffractive region. The dijet data could be included in future fits in order to better constrain the diffractive gluon distribution.

Within the experimental and theoretical uncertainties and assuming the H1 diffractive PDFs, factorisation is in good agreement with diffractive D^* production [21, 22] in the DIS kinematic region.

2.2 Diffractive Photoproduction of Dijets

In photoproduction, a sizeable contribution to the cross section is given by resolved photon processes (Fig. 2b) in which only a fraction $x_\gamma < 1$ of the photon momentum enters the hard scatter. The photoproduction dijet cross section measured by H1 ($Q^2 < 0.01$ GeV^2, $165 < W < 242$ GeV, $x_{I\!P} < 0.03$, $E_T^{\rm jet}(1,2) > 5,4$ GeV, $-1 < \eta_{\rm jet} < 2$, inclusive k_T algorithm) is shown in Fig. 6 [13]. NLO predictions have been obtained with the Frixione *et al.* program [23] interfaced to the H1 diffractive PDFs. The parton jet calculation is corrected for hadronisation effects using RAPGAP. The cross section differential in $z_{I\!P}^{\rm jets}$ is shown in Fig. 6a. The calculation lies a factor ≈ 2 above the data. Fig. 6b and 6c show the cross section as a function of $p_T^{\rm jet1}$ and $x_\gamma^{\rm jets}$ and the NLO predictions have been scaled down by a factor 0.5. Good agreement is obtained for this global suppression. In Fig. 6d, only the "resolved" part for which $x_\gamma^{\rm jets} < 0.9$ at the parton level is scaled by a factor 0.34. This factor was proposed by Kaidalov *et al.* [10] for the suppression of the resolved part in LO calculations. The calculation for $x_\gamma^{\rm jets} > 0.9$ is left unscaled. This approach is clearly disfavoured.

The ZEUS measurement [24] ($Q^2 < 0.01$ GeV^2, $x_{I\!P} < 0.025$, $0.2 < y < 0.85$, $E_T^{\rm jet}(1,2) > 7.5, 6.5$ GeV, $-1.5 < \eta < 1.5$, inclusive k_T algorithm) is shown in Fig. 7 and Fig. 8 separately for samples enriched in "direct" ($x_\gamma^{\rm jets} > 0.75$) and "resolved" ($x_\gamma^{\rm jets} < 0.75$) processes, respectively. The NLO [12] prediction using the H1 diffractive PDFs is also presented corrected for hadronization effects and with the "resolved" part scaled by the factor 0.34. No evidence is observed for a suppression of resolved photon processes relative to direct photon processes in any particular kinematic region.

Diffractive dijet photoproduction is overestimated by calculations based on PDFs which give a good description of the diffractive DIS data. Factorisation is broken in photoproduction relative to DIS by a factor ≈ 0.5 with no observed dependence on x_γ or other kinematic variables.

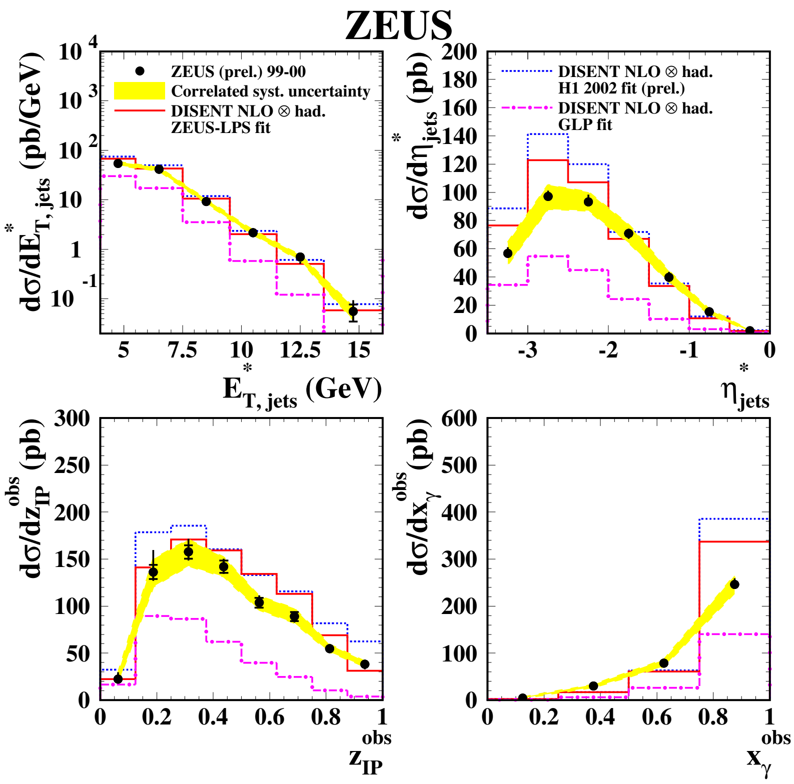


Fig. 5: Diffractive DIS dijet cross sections compared with NLO predictions based on three sets of diffractive PDFs.

3 Factorization and its Breaking in Diffractive Dijet Production

The fact that equal suppression of direct *and* resolved photoproduction by a factor $R = 0.5$ appears to describe the H1 and ZEUS data better phenomenologically motivates us to investigate in some detail the question whether certain parts of the direct contribution might break factorization as well and therefore need a suppression factor. These studies can be done for photoproduction ($Q^2 \simeq 0$) as well as for electroproduction with fixed, small Q^2. Since in electroproduction the initial-state singularity in the limit $Q^2 \to 0$ is more directly apparent than for the photoproduction case, we shall consider in this contribution the low-Q^2 electroproduction case just for demonstration.

A factorization scheme for virtual photoproduction has been defined and the full NLO corrections for inclusive dijet production have been calculated in [17]. They have been implemented in the NLO Monte Carlo program JETVIP [25] and adapted to diffractive dijet production in [15]. The subtraction term, which is absorbed into the PDFs of the virtual photon $f_{a/\gamma}(x_\gamma, M_\gamma)$, can be found in [26]. The main term is proportional to $\ln(M_\gamma^2/Q^2)$ times the splitting function

$$P_{q_i \leftarrow \gamma}(z) = 2N_c Q_i^2 \frac{z^2 + (1-z)^2}{2}, \tag{1}$$

where $z = p_1 p_2 / p_0 q \in [x; 1]$ and Q_i is the fractional charge of the quark q_i. p_1 and p_2 are the momenta of the two outgoing jets, and p_0 and q are the momenta of the ingoing parton and virtual photon, respectively. Since $Q^2 = -q^2 \ll M_\gamma^2$, the subtraction term is large and is therefore resummed by the DGLAP evolution equations for the virtual photon PDFs. After this subtraction, the finite term $M(Q^2)_{\overline{\rm MS}}$, which remains in the matrix element for the NLO correction to the direct process [17], has the same M_γ-dependence as the subtraction term, *i.e.* $\ln M_\gamma$ is multiplied with the same factor. As already mentioned, this yields the M_γ-dependence before the evolution is turned on. In the usual non-diffractive dijet photoproduction these two M_γ-dependences cancel, when the NLO correction to the direct part is added to the

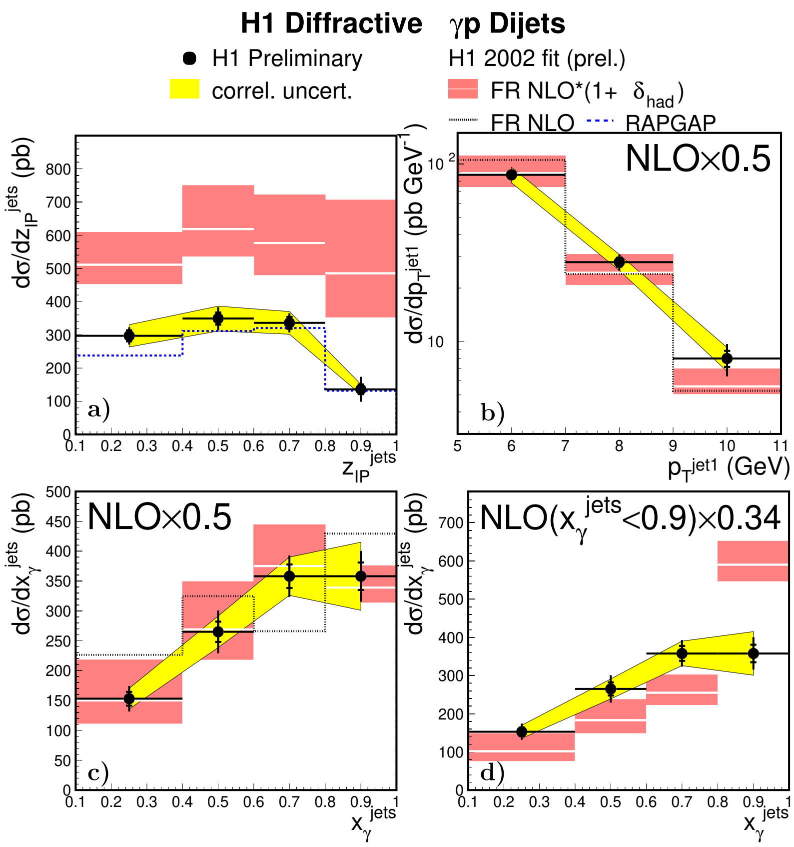


Fig. 6: a) Diffractive dijet photoproduction cross section differential in $z_{I\!P}^{\rm jets}$ compared with a NLO prediction based on diffractive PDFs and RAPGAP. b)-d): Cross section differential in $p_T^{\rm jet1}$ and $x_\gamma^{\rm jets}$, compared with the NLO prediction modified as follows: in b) and c) the calculation is scaled by a global factor 0.5 whereas in d) only the “resolved” part is scaled by 0.34.

LO resolved cross section [16]. Then it is obvious that the approximate M_γ-independence is destroyed, if the resolved cross section is multiplied by a suppression factor R to account for the factorization breaking in the experimental data. To remedy this deficiency, we propose to multiply the $\ln M_\gamma$-dependent term in $M(Q^2)_{\overline{\rm MS}}$ with the same suppression factor as the resolved cross section. This is done in the following way: we split $M(Q^2)_{\overline{\rm MS}}$ into two terms using the scale p_T^* in such a way that the term containing the slicing parameter y_s, which was used to separate the initial-state singular contribution, remains unsuppressed. In particular, we replace the finite term after the subtraction by

$$\begin{aligned} M(Q^2,R)_{\overline{\rm MS}} &= \left[-\frac{1}{2N_c}P_{q_i\leftarrow\gamma}(z)\ln\left(\frac{M_\gamma^2 z}{p_T^{*2}(1-z)}\right)+\frac{Q_i^2}{2}\right]R \\ &\quad -\frac{1}{2N_c}P_{q_i\leftarrow\gamma}(z)\ln\left(\frac{p_T^{*2}}{zQ^2+y_s s}\right), \end{aligned} \tag{2}$$

where R is the suppression factor. This expression coincides with the finite term after subtraction (see Ref. [26]) for $R=1$, as it should, and leaves the second term in Eq. (2) unsuppressed. In Eq. (2) we have suppressed in addition to $\ln(M_\gamma^2/p_T^{*2})$ also the z-dependent term $\ln(z/(1-z))$, which is specific to the $\overline{\rm MS}$ subtraction scheme as defined in [17]. The second term in Eq. (2) must be left in its original form, *i.e.* being unsuppressed, in order to achieve the cancellation of the slicing parameter (y_s) dependence of the complete NLO correction in the limit of very small Q^2 or equivalently very large s. It is clear that the suppression of this part of the NLO correction to the direct cross section will change the full cross section only very little as long as we choose $M_\gamma \simeq p_T^*$. The first term in Eq. (2), which has the suppression factor R, will be denoted by DIR_{IS} in the following.

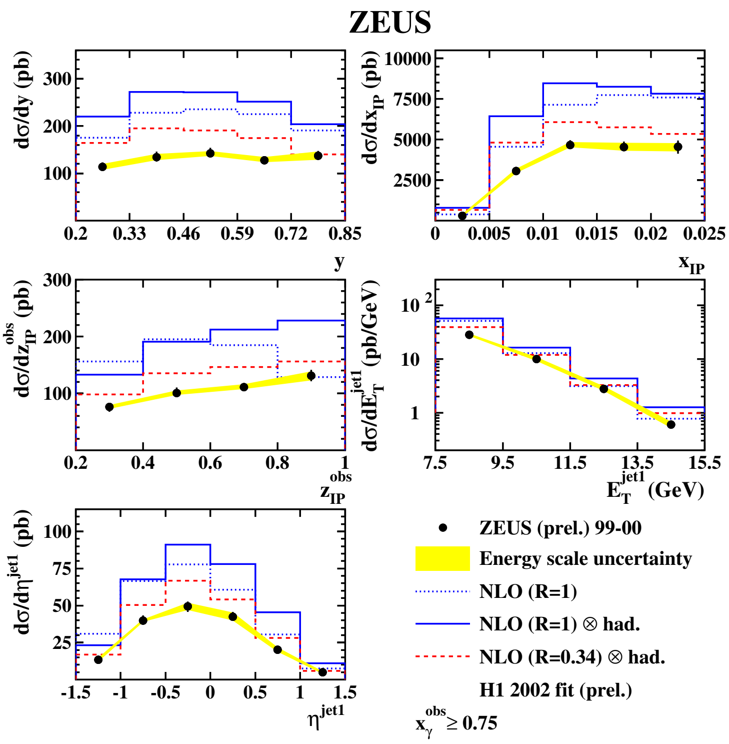


Fig. 7: Direct enriched photoproduction. Diffractive dijet photoproduction cross section differential in y, $x_{I\!P}$, $z_{I\!P}^{\rm jets}$, $E_T^{\rm jet\,1}$ and $\eta_{\rm jet\,1}$ compared with a NLO prediction based on diffractive PDFs. The NLO prediction is also presented corrected for hadronization effects and with the "resolved" part scaled by 0.34.

To study the left-over M_γ-dependence of the physical cross section, we have calculated the diffractive dijet cross section with the same kinematic constraints as in the H1 experiment [27]. Jets are defined by the CDF cone algorithm with jet radius equal to one and asymmetric cuts for the transverse momenta of the two jets required for infrared stable comparisons with the NLO calculations [28]. The original H1 analysis actually used a symmetric cut of 4 GeV on the transverse momenta of both jets [29]. The data have, however, been reanalyzed for asymmetric cuts [27].

For the NLO resolved virtual photon predictions, we have used the PDFs SaS1D [30] and transformed them from the DIS$_\gamma$ to the $\overline{\rm MS}$ scheme as in Ref. [17]. If not stated otherwise, the renormalization and factorization scales at the pomeron and the photon vertex are equal and fixed to $p_T^* = p_{T,jet1}^*$. We include four flavors, *i.e.* $n_f = 4$ in the formula for α_s and in the PDFs of the pomeron and the photon. With these assumptions we have calculated the same cross section as in our previous work [15]. First we investigated how the cross section ${\rm d}\sigma/{\rm d}Q^2$ depends on the factorization scheme of the PDFs for the virtual photon, *i.e.* ${\rm d}\sigma/{\rm d}Q^2$ is calculated for the choice SaS1D and SaS1M. Here ${\rm d}\sigma/{\rm d}Q^2$ is the full cross section (sum of direct and resolved) integrated over the momentum and rapidity ranges as in the H1 analysis. The results, shown in Fig. 2 of Ref. [26], demonstrate that the choice of the factorization scheme of the virtual photon PDFs has negligible influence on ${\rm d}\sigma/{\rm d}Q^2$ for all considered Q^2. The predictions agree reasonably well with the preliminary H1 data [27].

We now turn to the M_γ-dependence of the cross section with a suppression factor for DIR$_{\rm IS}$. To show this dependence for the two suppression mechanisms, (i) suppression of the resolved cross section only and (ii) additional suppression of the DIR$_{\rm IS}$ term as defined in Eq. (2) in the NLO correction of the direct cross section, we consider ${\rm d}\sigma/{\rm d}Q^2$ for the lowest Q^2-bin, $Q^2 \in [4,6]$ GeV^2. In Fig. 9, this cross section is plotted as a function of $\xi = M_\gamma/p_T^*$ in the range $\xi \in [0.25;4]$ for the cases (i)

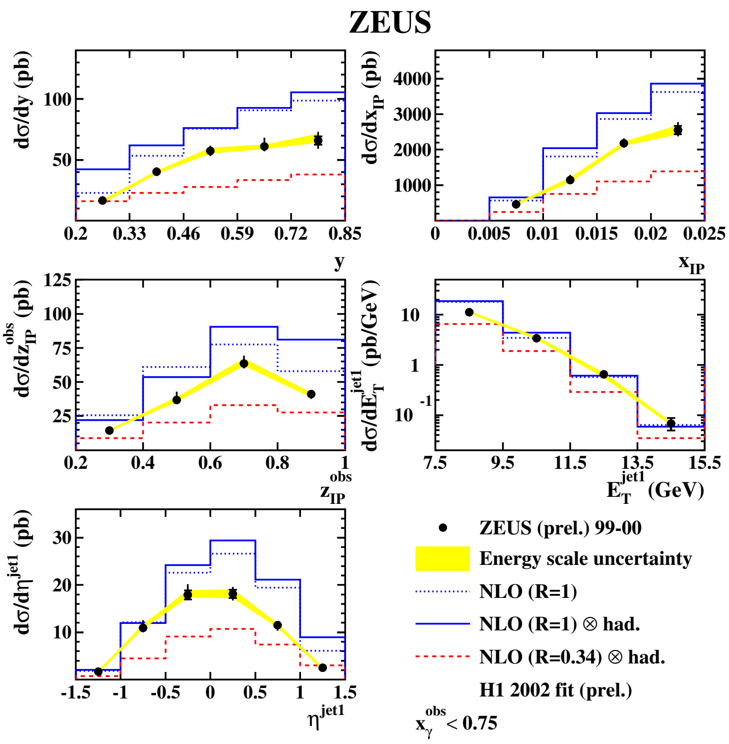


Fig. 8: Resolved enriched photoproduction. Diffractive dijet photoproduction cross section differential in y, $x_{I\!P}$, $z_{I\!P}^{\rm jets}$, $E_T^{\rm jet\,1}$ and $\eta_{\rm jet\,1}$ compared with a NLO prediction based on diffractive PDFs. The NLO prediction is also presented corrected for hadronization effects and with the "resolved" part scaled by 0.34.

(light full curve) and (ii) (full curve). We see that the cross section for case (i) has an appreciable ξ-dependence in the considered ξ range of the order of 40%, which is caused by the suppression of the resolved contribution only. With the additional suppression of the DIR$_{\rm IS}$ term in the direct NLO correction, the ξ-dependence of $\mathrm{d}\sigma/\mathrm{d}Q^2$ is reduced to approximately less than 20%, if we compare the maximal and the minimal value of $\mathrm{d}\sigma/\mathrm{d}Q^2$ in the considered ξ range. The remaining ξ-dependence is caused by the NLO corrections to the suppressed resolved cross section and the evolution of the virtual photon PDFs. How the compensation of the M_γ-dependence between the suppressed resolved contribution and the suppressed direct NLO term works in detail is exhibited by the dotted and dashed-dotted curves in Fig. 9. The suppressed resolved term increases and the suppressed direct NLO term decreases by approximately the same amount with increasing ξ. In addition we show also $\mathrm{d}\sigma/\mathrm{d}Q^2$ in the DIS theory, *i.e.* without subtraction of any $\ln Q^2$ terms (dashed line). Of course, this cross section must be independent of ξ. This prediction agrees very well with the experimental point, whereas the result for the subtracted and suppressed theory (full curve) lies slightly below. We notice, that for $M_\gamma = p_T^*$ the additional suppression of DIR$_{\rm IS}$ has only a small effect. It increases $\mathrm{d}\sigma/\mathrm{d}Q^2$ by 5% only.

4 Summary

Experimental data from the H1 and ZEUS Collaborations at HERA for diffractive dijet production in DIS and photoproduction have been compared with NLO QCD predictions using diffractive parton densities from H1 and ZEUS. While good agreement was found for DIS assuming the H1 diffractive PDFs, the dijet photoproduction data are overestimated by the NLO theory, showing that factorization breaking occurs at this order. While this is expected theoretically for resolved photoproduction, the fact that the data are better described by a global suppression of direct *and* resolved contribution by about a factor

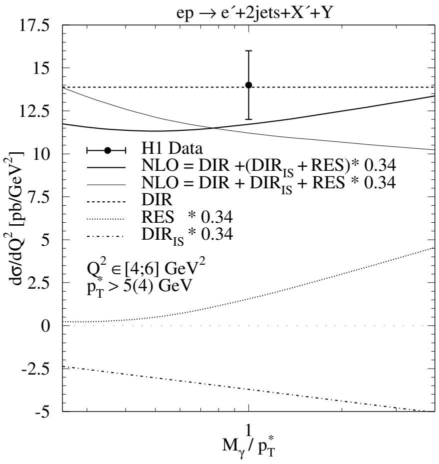


Fig. 9: Photon factorization scale dependence of resolved and direct contributions to $\mathrm{d}\sigma/\mathrm{d}Q^2$ together with their weighted sums for (i) suppression of the resolved cross section and for (ii) additional suppression of $\mathrm{DIR}_{\mathrm{IS}}$, using SaS1D virtual photon PDFs [30].

of two has come as a surprise. We have therefore discussed in some detail the factorization scheme and scale dependence between direct and resolved contributions and proposed a new factorization scheme for diffractive dijet photoproduction.

Acknowledgments

M.K. thanks the II. Institute for Theoretical Physics at the University of Hamburg for hospitality while this work was being finalized.

References

[1] H1 Collaboration, Abstract 980, contributed to the 31$^{\mathrm{st}}$ International Conference on High Energy Physics (ICHEP 2002), Amsterdam, July 2002.

[2] S. Chekanov et al. [ZEUS Collaboration], *Eur. Phys. J.* **C 38** (2004) 43, and A. Proskuryakov, private communication.

[3] M. Groys, A. Levy and A. Proskuryakov, these proceedings.

[4] F.-P. Schilling and P. Newman, these proceedings.

[5] A. D. Martin, M. G. Ryskin and G. Watt, Eur. Phys. J. C **37**, 285 (2004) and DESY 05-055, IPPP/05/07, DCPT/05/14, hep-ph/0504132.

[6] J. C. Collins, Phys. Rev. D **57**, 3051 (1998) [Erratum-ibid. D **61**, 019902 (2000)].

[7] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **84**, 5043 (2000).

[8] A. B. Kaidalov, V. A. Khoze, A. D. Martin and M. G. Ryskin, Eur. Phys. J. C **21**, 521 (2001).

[9] M. Klasen, Rev. Mod. Phys. **74**, 1221 (2002).

[10] A. B. Kaidalov, V. A. Khoze, A. D. Martin and M. G. Ryskin, Phys. Lett. B **567**, 61 (2003).

[11] M. Klasen and G. Kramer, hep-ph/0401202, Proceedings of the 12$^{\mathrm{th}}$ International Workshop on Deep Inelastic Scattering (DIS 2004), eds. D. Bruncko, J. Ferencei and P. Strizenec, Kosice, Inst. Exp. Phys. SAS, 2004, p. 492.

[12] M. Klasen and G. Kramer, Eur. Phys. J. C **38**, 93 (2004).

[13] H1 Collaboration, Abstract 6-0177, contributed to the 32nd International Conference on High Energy Physics (ICHEP 2004), Beijing, August 2004.

[14] ZEUS Collaboration, Abstract 6-0249, contributed to the 32nd International Conference on High Energy Physics (ICHEP 2004), Beijing, August 2004.

[15] M. Klasen and G. Kramer, Phys. Rev. Lett. **93**, 232002 (2004).

[16] D. Bödeker, G. Kramer and S. G. Salesch, Z. Phys. C **63**, 471 (1994).

[17] M. Klasen, G. Kramer and B. Pötter, Eur. Phys. J. C **1**, 261 (1998).

[18] S. Catani, M.H. Seymour, *Nucl. Phys.* **B 485** (1997) 291; erratum ibid. **B510** (1997) 503.

[19] H. Jung, *Comp. Phys. Commun.* **86** (1995) 147.

[20] ZEUS Collaboration, Abstract 295 and addendum, contributed to the 22nd International Symposium on Lepton-Photon Interactions oh High Energy, Uppsala, Sweden, June 2005.

[21] H1 Collaboration, Abstract 6-0178, contributed to the 32nd International Conference on High Energy Physics (ICHEP 2004), Beijing, August 2004.

[22] S. Chekanov et al. [ZEUS Collaboration], *Nucl. Phys.* **B 672** (2003) 3,

[23] S. Frixione, Z. Kunszt and A. Signer, *Nucl. Phys.* **B 467** (1996) 399;
S. Frixione, *Nucl. Phys.* **B 507** (1997) 295.

[24] ZEUS Collaboration, Abstract 293, contributed to the 22nd International Symposium on Lepton-Photon Interactions oh High Energy, Uppsala, Sweden, June 2005.

[25] B. Pötter, Comput. Phys. Commun. **133**, 105 (2000).

[26] M. Klasen and G. Kramer, hep-ph/0506121, accepted for publication in J. Phys. G.

[27] S. Schätzel, hep-ex/0408049, Proceedings of the 12th International Workshop on Deep Inelastic Scattering (DIS 2004), eds. D. Bruncko, J. Ferencei and P. Strizenec, Kosice, Inst. Exp. Phys. SAS, 2004, p. 529; H1 Collaboration, Abstract 6-0176, contributed to the 32nd International Conference on High Energy Physics (ICHEP 2004), Beijing, August 2004.

[28] M. Klasen and G. Kramer, Phys. Lett. B **366**, 385 (1996).

[29] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **20**, 29 (2001).

[30] G. A. Schuler and T. Sjöstrand, Phys. Lett. B **376**, 193 (1996).

Effect of absorptive corrections on inclusive parton distributions

G. Watt[a], A.D. Martin[b], M.G. Ryskin[b,c]
[a] Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany
[b] Institute for Particle Physics Phenomenology, University of Durham, DH1 3LE, UK
[c] Petersburg Nuclear Physics Institute, Gatchina, St. Petersburg, 188300, Russia

Abstract
We study the effect of absorptive corrections due to parton recombination on the parton distributions of the proton. A more precise version of the GLRMQ equations, which account for non-linear corrections to DGLAP evolution, is derived. An analysis of HERA F_2 data shows that the small-x gluon distribution is enhanced at low scales when the absorptive effects are included, such that there is much less need for a negative gluon distribution at 1 GeV.

1 Parton recombination at small x

At very small values of x it is expected that the number density of partons within the proton becomes so large that they begin to recombine with each other. This phenomenon of parton recombination is also referred to as absorptive corrections, non-linear effects, screening, shadowing, or unitarity corrections, all leading to saturation. The first perturbative QCD (pQCD) calculations describing the fusion of two Pomeron ladders into one were made by Gribov-Levin-Ryskin (GLR) [1] and by Mueller-Qiu (MQ) [2]. The GLRMQ equations add an extra non-linear term, quadratic in the gluon density, to the usual DGLAP equations for the gluon and sea-quark evolution. The evolution of the gluon distribution is then given by

$$\frac{\partial xg(x,Q^2)}{\partial \ln Q^2} = \frac{\alpha_S}{2\pi} \sum_{a'=q,g} P_{ga'} \otimes a' \; - \; \frac{9}{2}\,\frac{\alpha_S^2(Q^2)}{R^2\,Q^2} \int_x^1 \frac{\mathrm{d}x'}{x'} \left[x'g(x',Q^2)\right]^2, \tag{1}$$

where $R \sim 1$ fm is of the order of the proton radius. The GLRMQ equations account for all 'fan' diagrams, that is, all possible $2 \to 1$ ladder recombinations, in the double leading logarithmic approximation (DLLA) which resums all powers of the parameter $\alpha_S \ln(1/x) \ln(Q^2/Q_0^2)$.

There has been much recent theoretical activity in deriving (and studying) more precise non-linear evolution equations, such as the Balitsky-Kovchegov (BK) and Jalilian-Marian–Iancu–McLerran–Weigert–Leonidov–Kovner (JIMWLK) equations (see [3] for a review). Note that the BK and JIMWLK equations are both based on BFKL evolution. However, for the most relevant studies in the HERA and LHC domain ($x \gtrsim 10^{-4}$), the predominant theoretical framework is collinear factorisation with DGLAP-evolved parton distribution functions (PDFs). At very small values of x it might be expected that the DGLAP approximation would break down, since large $\alpha_S \ln(1/x)$ (BFKL) terms would appear in the perturbation series in addition to the $\alpha_S \ln(Q^2/Q_0^2)$ terms resummed by DGLAP evolution. However, it turns out that the resummed NLL BFKL calculations of the gluon splitting function P_{gg} [4] and the gluon transverse momentum distribution [5] are rather close to the DGLAP calculations. Moreover, the convolution $P_{gg} \otimes g(x,Q^2)$ coincides with the NNLO DGLAP result and is close to the NLO DGLAP result for $x \gtrsim 10^{-4}$ [6]. Hence, in the analysis of current data, it is reasonable to ignore BFKL effects.

If recombination effects are significant, it is therefore important that they be incorporated into the global DGLAP parton analyses which determine the PDFs from deep-inelastic scattering (DIS) and related hard-scattering data. Such a programme, based on GLRMQ evolution (which accounts for gluon-induced screening only), was implemented some years ago [7], before the advent of HERA. The input gluon and sea-quark distributions were *assumed* to have a small-x behaviour of the form $xg, xS \sim x^{-0.5}$ at an input scale of $Q_0^2 = 4$ GeV^2. The inclusion of shadowing effects, both in the form of the input

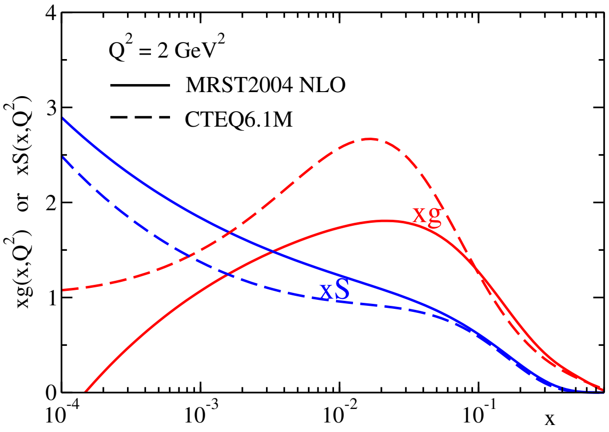


Fig. 1: The behaviour of the gluon and sea-quark distributions at $Q^2 = 2$ GeV^2 found in the MRST2004 NLO and CTEQ6.1M global analyses. The valence-like behaviour of the gluon is evident.

PDFs and in the GLRMQ evolution, was found to significantly *decrease* the size of the small-x gluon distribution in comparison with the result with no absorptive corrections. A crucial observation is that, at that time (1990), F_2 data were only available for $x_B \geq 0.07$, and so these results were largely dependent on the theoretical assumptions made for the starting distributions. However, with HERA, we now have F_2 data down to $x_B \sim 10^{-4}$ or less, and so the PDFs at small x can be determined directly from the HERA data.

In fact, the advent of HERA data has led to a puzzling behaviour of the small-x gluon and sea-quark PDFs at low scales Q^2. If we write $xg \sim x^{-\lambda_g}$ and $xS \sim x^{-\lambda_S}$, then the expectation of Regge theory is that $\lambda_g = \lambda_S = \lambda_{\rm soft}$ for low scales $Q \lesssim Q_0 \sim 1$ GeV, where $\lambda_{\rm soft} \simeq 0.08$ [8] is the power of s obtained from fitting soft hadron data. At higher $Q \gtrsim 1$ GeV, QCD evolution should take over, increasing the powers λ_g and λ_S. However, the current MRST2004 NLO [9] and CTEQ6.1M [10] PDF sets exhibit a very different behaviour at low scales from that theoretically expected; see Fig. 1. In fact, the MRST group has found that a negative input gluon distribution at $Q_0 = 1$ GeV is required in all their NLO DGLAP fits since MRST2001 [11]. The CTEQ group, who take a slightly higher input scale of $Q_0 = 1.3$ GeV, also find a negative gluon distribution when evolving backwards to 1 GeV.

Since data at small x_B now exist, the introduction of the absorptive corrections is expected to *increase* the size of the input gluon distribution at small x to maintain a satisfactory fit to the data. To understand this, note that the negative non-linear term in the GLRMQ equation (1) slows down the evolution. Therefore, it is necessary to start with a *larger* small-x gluon distribution at low scales $Q \sim Q_0$ to achieve the *same* PDFs at larger scales required to describe the data. If the non-linear term is neglected, the input small-x gluon distribution is forced to be artificially small in order to *mimic* the neglected screening corrections.

We have anticipated that the introduction of absorptive corrections will *enhance*[1] the small-x gluon at low scales, and hence could possibly avoid what appears to be anomalous behaviour at small x. Thus, here, we perform such a study using an abridged version of the MRST2001 NLO analysis [11], improving on our previous analysis [13]. First, we derive a more precise form of the GLRMQ equations.

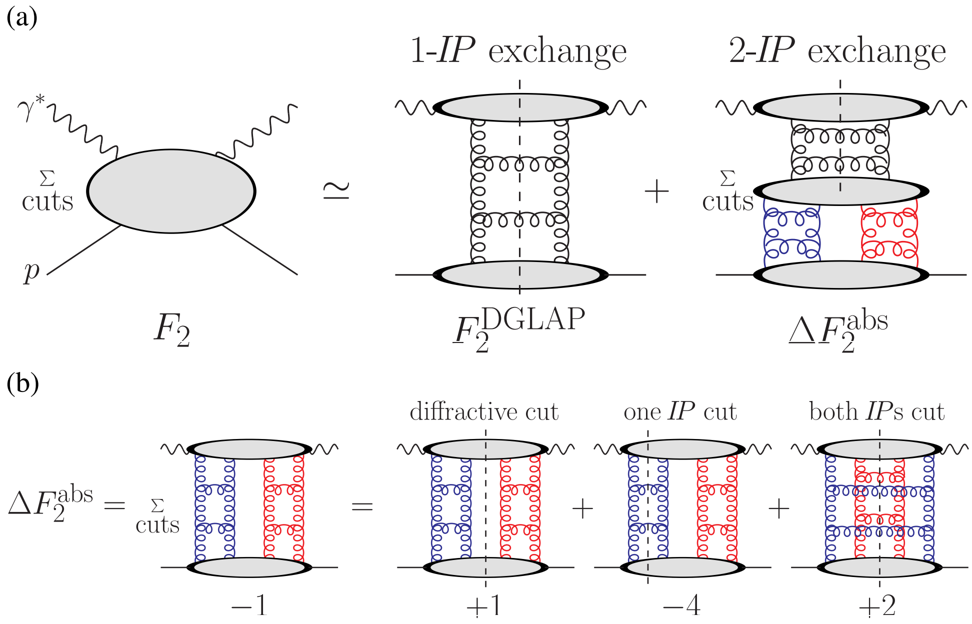


Fig. 2: (a) Absorptive corrections to F_2 due to the $2 \to 1$ Pomeron contribution. (b) Application of the AGK cutting rules. For simplicity, the upper parton ladder, shown in the right-hand diagram of (a), is hidden inside the upper blob in each diagram of (b).

2 Non-linear evolution from diffractive DIS

The inclusive proton structure function, $F_2(x_B, Q^2)$, as measured by experiment, can be approximately written as a sum of the single Pomeron exchange (DGLAP) contribution and absorptive corrections due to a $2 \to 1$ Pomeron merging; see Fig. 2(a). That is,

$$F_2(x_B, Q^2) = F_2^{\rm DGLAP}(x_B, Q^2) + \Delta F_2^{\rm abs}(x_B, Q^2). \tag{2}$$

In computing $\Delta F_2^{\rm abs}$ we need to sum over all possible cuts. The Abramovsky-Gribov-Kancheli (AGK) cutting rules [14] were originally formulated in Reggeon field theory but have been shown to also hold in pQCD [15]. Application of the AGK rules gives the result that relative contributions of $+1$, -4, and $+2$ are obtained according to whether neither Pomeron, one Pomeron, or both Pomerons are cut; see Fig. 2(b). Therefore, the sum over cuts is equal to *minus* the diffractive cut and so the absorptive corrections can be computed from a calculation of the t-integrated diffractive structure function $F_2^{\rm D(3)}(x_{I\!P}, \beta, Q^2)$, where $\beta \equiv x_B/x_{I\!P}$ and $x_{I\!P}$ is the fraction of the proton's momentum transferred through the rapidity gap.

The pQCD description of $F_2^{\rm D(3)}$ is described in [16, 17], and in a separate contribution to these proceedings. Working in the fixed flavour number scheme (FFNS), it can be written as

$$F_2^{\rm D(3)} = \underbrace{F_{2,\rm non-pert.}^{\rm D(3)}}_{\text{soft Pomeron}} + \underbrace{F_{2,\rm pert.}^{\rm D(3)} + F_{2,\rm direct}^{{\rm D(3)},c\bar{c}} + F_{L,\rm tw.4}^{\rm D(3)}}_{\text{QCD Pomeron}}, \tag{3}$$

apart from the secondary Reggeon contribution. The separation between the soft Pomeron and QCD Pomeron is provided by a scale $\mu_0 \sim 1$ GeV. For simplicity, we take μ_0 to be the same as the scale Q_0 at which the input PDFs are taken in the analysis of F_2 data, so $\mu_0 = Q_0 = 1$ GeV, the value used in

[1] Eskola *et al.* [12] have found that taking input gluon and sea-quark distributions at $Q^2 = 1.4$ GeV^2, then evolving upwards with the GLRMQ equations based on LO DGLAP evolution, improves the agreement with F_2 data at small x_B and low Q^2 compared to the standard CTEQ sets, and leads to an enhanced small-x gluon distribution for $Q^2 \lesssim 10$ GeV^2. Note, however, that there is a large NLO correction to the splitting function P_{qg} which changes completely the relationship between the quark and gluon distributions, and so weakens the conclusion of Ref. [12].

the MRST2001 NLO analysis [11]. The contribution to the absorptive corrections arising from the soft Pomeron contribution of (3) is already included in the input PDFs, therefore

$$\Delta F_2^{\rm abs} = -\frac{1}{1-f_{\rm p.diss.}} \int_{x_B}^{1} {\rm d}x_{I\!P} \, \left[F_{2,{\rm pert.}}^{\rm D(3)} + F_{2,{\rm direct}}^{{\rm D(3)},c\bar{c}} + F_{L,{\rm tw.4}}^{\rm D(3)} \right], \tag{4}$$

where $f_{\rm p.diss.}$ is the fraction of diffractive events in which the proton dissociates. In practice, we take $f_{\rm p.diss.} = 0.5$ and take an upper limit of 0.1 instead of 1 for $x_{I\!P}$ in (4).[2]

First consider the contribution to (4) from the $F_{2,{\rm pert.}}^{\rm D(3)}$ term.[3] It corresponds to a $2 \to 1$ Pomeron merging with a cut between the two Pomeron ladders and can be written as

$$F_{2,{\rm pert.}}^{\rm D(3)}(x_{I\!P}, \beta, Q^2) = \sum_{a=q,g} C_{2,a} \otimes a_{\rm pert.}^{\rm D}, \tag{5}$$

where $C_{2,a}$ are the *same* coefficient functions as in inclusive DIS. The diffractive PDFs, $a^{\rm D} = zq^{\rm D}$ or $zg^{\rm D}$, where $z \equiv x/x_{I\!P}$, satisfy an *inhomogeneous* evolution equation [17]:

$$a_{\rm pert.}^{\rm D}(x_{I\!P}, z, Q^2) = \int_{\mu_0^2}^{Q^2} \frac{{\rm d}\mu^2}{\mu^2} \, f_{I\!P}(x_{I\!P}; \mu^2) \; a^{I\!P}(z, Q^2; \mu^2) \tag{6}$$

$$\Longrightarrow \frac{\partial a_{\rm pert.}^{\rm D}}{\partial \ln Q^2} = \frac{\alpha_S}{2\pi} \sum_{a'=q,g} P_{aa'} \otimes a'^{\rm D}_{\rm pert.} \; + \; P_{aI\!P}(z) \, f_{I\!P}(x_{I\!P}; Q^2). \tag{7}$$

Here, $f_{I\!P}(x_{I\!P}; Q^2)$ is the perturbative Pomeron flux factor,

$$f_{I\!P}(x_{I\!P}; \mu^2) = \frac{1}{x_{I\!P} B_D} \left[R_g \frac{\alpha_S(\mu^2)}{\mu} \, x_{I\!P} g(x_{I\!P}, \mu^2) \right]^2. \tag{8}$$

The diffractive slope parameter B_D comes from the t-integration, while the factor R_g accounts for the skewedness of the proton gluon distribution [19]. There are similar contributions from (light) sea quarks, where g in (8) is replaced by $S \equiv 2(\bar{u} + \bar{d} + \bar{s})$, together with an interference term. A sum over all three contributions is implied in (6) and in the second term of (7). The Pomeron PDFs in (6), $a^{I\!P}(z, Q^2; \mu^2)$, are evolved using NLO DGLAP from a starting scale μ^2 up to Q^2, taking the input distributions to be LO Pomeron-to-parton splitting functions, $a^{I\!P}(z, \mu^2; \mu^2) = P_{aI\!P}(z)$ [17].

From (2),

$$a(x, Q^2) = a^{\rm DGLAP}(x, Q^2) + \Delta a^{\rm abs}(x, Q^2), \tag{9}$$

where $a(x, Q^2) = xg(x, Q^2)$ or $xS(x, Q^2)$, and

$$\Delta a^{\rm abs}(x, Q^2) = -\frac{1}{1-f_{\rm p.diss.}} \int_x^1 {\rm d}x_{I\!P} \; a_{\rm pert.}^{\rm D}(x_{I\!P}, x/x_{I\!P}, Q^2). \tag{10}$$

Differentiating (9) with respect to Q^2 gives the evolution equations for the (inclusive) gluon and sea-quark PDFs:

$$\boxed{\frac{\partial a(x, Q^2)}{\partial \ln Q^2} = \frac{\alpha_S}{2\pi} \sum_{a'=q,g} P_{aa'} \otimes a' \; - \; \frac{1}{1-f_{\rm p.diss.}} \int_x^1 {\rm d}x_{I\!P} \; P_{aI\!P}(x/x_{I\!P}) \, f_{I\!P}(x_{I\!P}; Q^2).} \tag{11}$$

Thus (11) is a more precise version of the GLRMQ equations (1), which goes beyond the DLLA and accounts for sea-quark recombination as well as gluon recombination. Consider the recombination of

[2]The value of $f_{\rm p.diss.} = 0.5$ is justified by a ZEUS comparison [18] of proton-tagged diffractive DIS data with data which allowed proton dissociation up to masses of 6 GeV, where $f_{\rm p.diss.} = 0.46 \pm 0.11$ was obtained.

[3]The other two contributions to (4) are described after (13).

gluons into gluons, for example, in the DLLA where $x \ll x_{I\!P}$, then $P_{gI\!P} = 9/16$ [17]. Taking $R_g = 1$ and $f_{\text{p.diss.}} = 0$, then (11) becomes

$$\frac{\partial xg(x,Q^2)}{\partial \ln Q^2} = \frac{\alpha_S}{2\pi} \sum_{a'=q,g} P_{ga'} \otimes a' \; - \; \frac{9}{16} \, \frac{\alpha_S^2(Q^2)}{B_D \, Q^2} \int_x^1 \frac{\mathrm{d}x_{I\!P}}{x_{I\!P}} \left[x_{I\!P} g(x_{I\!P}, Q^2) \right]^2 . \tag{12}$$

Comparing to (1) this is simply the GLRMQ equation with $R^2 = 8B_D$. For numerical results we take $B_D = 6$ (4) GeV^{-2} for light (charm) quarks, which would correspond to $R = \sqrt{8B_D} = 1.4$ (1.1) fm.

The procedure for incorporating absorptive corrections into a (NLO) global parton analysis (in the FFNS) is as follows:

1. Parameterise the x dependence of the input PDFs at a scale $Q_0 \sim 1$ GeV.
2. Evolve the PDFs $xg(x,Q^2)$ and $xS(x,Q^2)$ using the non-linear evolution equation (11). (The non-singlet distributions are evolved using the usual linear DGLAP equations.)
3. Compute

$$F_2(x_B, Q^2) = \sum_{a=q,g} C_{2,a} \otimes a \; - \; \frac{1}{1 - f_{\text{p.diss.}}} \int_{x_B}^1 \mathrm{d}x_{I\!P} \, \left[F_{2,\text{direct}}^{\text{D(3)},c\bar{c}} + F_{L,\text{tw.4}}^{\text{D(3)}} \right], \tag{13}$$

   and compare to data. Here, the two terms inside the square brackets are beyond collinear factorisation, that is, they cannot be written as a convolution of coefficient functions with the PDFs. The first term inside the square brackets corresponds to the process $\gamma^* I\!P \to c\bar{c}$. The second term corresponds to the process $\gamma^* I\!P \to q\bar{q}$, for light quarks with a longitudinally polarised photon. These contributions are calculated as described in Ref. [17].

As usual, these three steps should be repeated with the parameters of the input PDFs adjusted until an optimal fit is obtained. This procedure is our recommended way of accounting for absorptive corrections in a global parton analysis. However, in practice, available NLO DGLAP evolution codes, such as the QCDNUM [20] program, are often regarded as a 'black box', and it is not trivial to modify the usual linear DGLAP evolution to the non-linear evolution of (11). Therefore, we adopt an alternative iterative procedure which avoids the explicit implementation of non-linear evolution, but which is equivalent to the above procedure.

3 Effect of absorptive corrections on inclusive PDFs

We model our analysis of HERA F_2 data [21] on the MRST2001 NLO analysis [11], which was the first in which a negative gluon distribution was required at the input scale of $Q_0 = 1$ GeV. (The more recent MRST sets have not changed substantially at small x.) We apply cuts $x_B \le 0.01$, $Q^2 \ge 2$ GeV^2, and $W^2 \ge 12.5$ GeV^2, leaving 280 data points. The input gluon and sea-quark distributions are taken to be

$$xg(x,Q_0^2) \; = \; A_g \, x^{-\lambda_g} (1-x)^{3.70} (1 + \epsilon_g \sqrt{x} + \gamma_g x) \; - \; A_- \, x^{-\delta_-} (1-x)^{10}, \tag{14}$$

$$xS(x,Q_0^2) \; = \; A_S \, x^{-\lambda_S} (1-x)^{7.10} (1 + \epsilon_S \sqrt{x} + \gamma_S x), \tag{15}$$

where the powers of the $(1-x)$ factors are taken from [11], together with the valence-quark distributions, u_V and d_V, and $\Delta \equiv \bar{d} - \bar{u}$. The A_g parameter is fixed by the momentum sum rule, while the other nine parameters are allowed to go free. Since we do not fit to DIS data with $x_B > 0.01$, we constrain the input gluon and sea-quark distributions, and their derivatives with respect to x, to agree with the MRST2001 NLO parton set [11] at $x = 0.2$. This is done by including the value of these MRST PDFs at $x = 0.2$, and their derivatives, as data points in the fit, with an error of 10% on both the value of the MRST PDFs and their derivatives. Therefore, the PDFs we obtain are not precisely constrained at large x, but this paper is primarily concerned with the small-x behaviour of the PDFs.

The procedure we adopt is as follows:

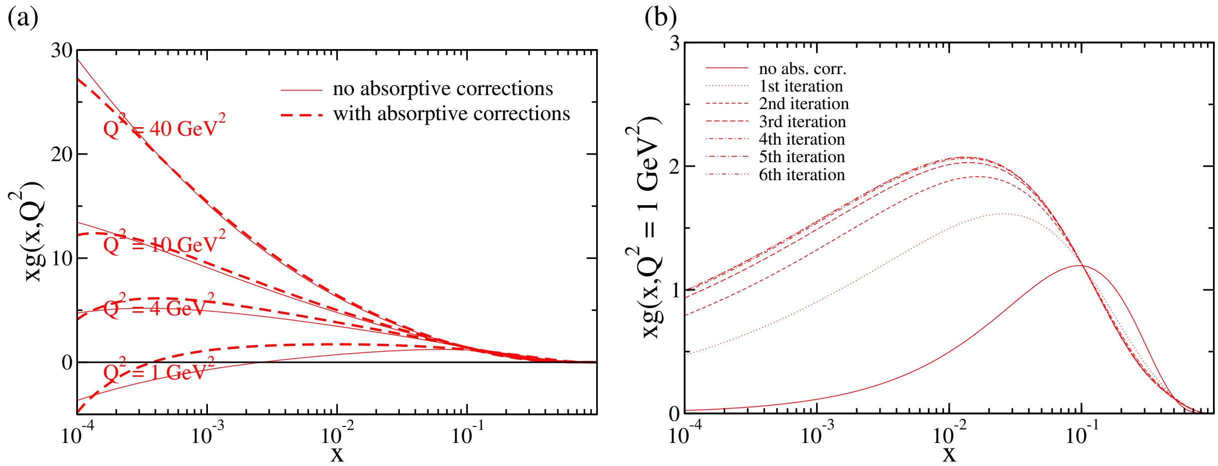


Fig. 3: (a) The gluon distribution obtained from fits to F_2 data, before and after absorptive corrections have been included. (b) The effect of successive iterations on the gluon distribution obtained from fits to F_2, taking a positive definite input gluon at 1 GeV. Each iteration introduces another level of $2 \to 1$ Pomeron mergings.

(i) Start by performing a standard NLO DGLAP fit to F_2 data with no absorptive corrections.

(ii) Tabulate $\Delta F_2^{\rm abs}$, given by (4), and $\Delta a^{\rm abs}$, given by (10), using PDFs $g(x_{I\!P}, \mu^2)$ and $S(x_{I\!P}, \mu^2)$ obtained from the previous fit.

(iii) Perform a standard NLO DGLAP fit to 'corrected' data, $F_2^{\rm DGLAP} = F_2 - \Delta F_2^{\rm abs}$, to obtain PDFs $a^{\rm DGLAP}$. Then correct these PDFs to obtain $a = a^{\rm DGLAP} + \Delta a^{\rm abs}$. These latter PDFs a then satisfy the non-linear evolution equations (11).

(iv) Go to (ii).

Each successive iteration of steps (ii) and (iii) introduces another level of $2 \to 1$ Pomeron mergings, so that eventually all the 'fan' diagrams are included, achieving the same effect as the procedure described at the end of Section 2.

Note that the correction to the PDFs, $a = a^{\rm DGLAP} + \Delta a^{\rm abs}$, in each step (iii), was omitted in our previous analysis [13]. Consequently, the effect of the absorptive corrections on the PDFs at large scales was overestimated. Also in [13], the known LO $P_{aI\!P}(z)$ were multiplied by free parameters ('K-factors'), determined from separate fits to diffractive DIS data, in an attempt to account for higher-order pQCD corrections to the LO Pomeron-to-parton splitting functions. However, since these K-factors took unreasonable values, with some going to zero, here we have chosen to fix them to 1. Therefore, the updated analysis, presented here, does not require a simultaneous fit to the diffractive DIS data.

In Fig. 3(a) we show the gluon distribution at scales $Q^2 = 1$, 4, 10, and 40 GeV^2 obtained from fits before and after absorptive corrections have been included. Both fits are almost equally good with $\chi^2/{\rm d.o.f.}$ values of 0.86 and 0.87 for the fits without and with absorptive corrections respectively. At low Q^2 the absorptive corrections give an increased gluon distribution at small x, apart from at $x \lesssim 10^{-4}$ where there are only a few data points and where additional absorptive effects (Pomeron loops) may become important. The non-linear term of (11) slows down the evolution, so that by 40 GeV^2 the two gluon distributions are roughly equal; see Fig. 3(a).

We repeated the fits without the negative term in the input gluon distribution, that is, without the second term in (14). When absorptive corrections were included, almost the same quality of fit was obtained ($\chi^2/{\rm d.o.f.} = 0.90$), while without absorptive corrections the fit was slightly worse ($\chi^2/{\rm d.o.f.} = 0.95$). We conclude that absorptive corrections lessen the need for a negative gluon distribution at $Q^2 = 1$ GeV^2. The gluon distributions obtained from six successive iterations of steps (ii) and (iii) above are shown in Fig. 3(b). The convergence is fairly rapid, with only the first three iterations having a significant effect, that is, the 'fan' diagrams which include $8 \to 4 \to 2 \to 1$ Pomeron mergings.

Although we have seen that the inclusion of absorptive corrections has reduced the need for a *negative* gluon, it has not solved the problem of the *valence-like* gluon. That is, the gluon distribution at low scales still decreases with decreasing x, whereas from Regge theory it is expected to behave as $xg \sim x^{-\lambda_{\rm soft}}$ with $\lambda_{\rm soft} \simeq 0.08$. We have studied several possibilities of obtaining a satisfactory fit with this behaviour [13]. The only modification which appears consistent with the data (and with the desired $\lambda_g = \lambda_S$ equality) is the inclusion of power-like corrections, specifically, a global shift in all scales by about 1 GeV^2. (Note that a similar shift in the scale is required in the dipole saturation model [22].) However, we do not have a solid theoretical justification for this shift. Therefore, a more detailed, and more theoretically-motivated, investigation of the effect of power corrections in DIS is called for.

References

[1] Gribov, L. V. and Levin, E. M. and Ryskin, M. G., Phys. Rept. **100**, 1 (1983).

[2] Mueller, Alfred H. and Qiu, Jian-wei, Nucl. Phys. **B268**, 427 (1986).

[3] Jalilian-Marian, Jamal and Kovchegov, Yuri V., Prog. Part. Nucl. Phys. (in press) (hep-ph/0505052).

[4] Thorne, Robert S., Phys. Lett. **B474**, 372 (2000);
Altarelli, Guido and Ball, Richard D. and Forte, Stefano, Nucl. Phys. **B674**, 459 (2003);
Ciafaloni, Marcello and Colferai, Dimitri and Salam, Gavin P. and Stasto, Anna M., Phys. Lett. **B587**, 87 (2004).

[5] Khoze, V. A. and Martin, A. D. and Ryskin, M. G. and Stirling, W. J., Phys. Rev. **D70**, 074013 (2004).

[6] Salam, G. P., hep-ph/0501097.

[7] Kwiecinski, J. and Martin, Alan D. and Stirling, W. James and Roberts, R. G., Phys. Rev. **D42**, 3645 (1990).

[8] Donnachie, A. and Landshoff, P. V., Phys. Lett. **B296**, 227 (1992).

[9] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Phys. Lett. **B604**, 61 (2004).

[10] Pumplin, J. and others, JHEP **07**, 012 (2002).

[11] Martin, Alan D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C23**, 73 (2002).

[12] Eskola, K. J. and Honkanen, H. and Kolhinen, V. J. and Qiu, Jian-wei and Salgado, C. A., Nucl. Phys. **B660**, 211 (2003).

[13] Martin, A. D. and Ryskin, M. G. and Watt, G., Phys. Rev. **D70**, 091502 (2004).

[14] Abramovsky, V. A. and Gribov, V. N. and Kancheli, O. V., Yad. Fiz. **18**, 595 (1973).

[15] Bartels, Jochen and Ryskin, M. G., Z. Phys. **C76**, 241 (1997).

[16] Martin, A. D. and Ryskin, M. G. and Watt, G., Eur. Phys. J. **C37**, 285 (2004).

[17] Martin, A. D. and Ryskin, M. G. and Watt, G., Eur. Phys. J. **C**, (in press) (hep-ph/0504132).

[18] Chekanov, S. and others, Eur. Phys. J. **C25**, 169 (2002).

[19] Shuvaev, A. G. and Golec-Biernat, K. J. and Martin, Alan D. and Ryskin, M. G., Phys. Rev. **D60**, 014015 (1999).

[20] Botje, M., QCDNUM version 16.12, ZEUS Note 97-066 (unpublished), available from `http://www.nikhef.nl/~h24/qcdcode/`.

[21] Breitweg, J. and others, Eur. Phys. J. **C7**, 609 (1999);
Chekanov, S. and others, Eur. Phys. J. **C21**, 443 (2001);
Adloff, C. and others, Eur. Phys. J. **C21**, 33 (2001);
Adloff, C. and others, Eur. Phys. J. **C19**, 269 (2001);
Adloff, C. and others, Eur. Phys. J. **C13**, 609 (2000).

[22] Bartels, J. and Golec-Biernat, K. and Kowalski, H., Phys. Rev. **D66**, 014001 (2002);
Kowalski, Henri and Teaney, Derek, Phys. Rev. **D68**, 114005 (2003).

Multiple Scattering at HERA and at LHC - Remarks on the AGK Rules

J. Bartels
II. Institut für Theoretische Physik, Universität Hamburg
Luruper Chaussee 149, D-22761 Hamburg, Germany

Abstract
We summarize the present status of the AGK cutting rules in perturbative QCD. Particular attention is given to the application of the AGK analysis to multiple scattering in DIS at HERA and in pp collisions at the LHC

1 Introduction

Multiple parton interactions play an important role both in electron proton scattering at HERA and in high energy proton proton collisions at the LHC. At HERA, the linear QCD evolution equations provides, for not too small Q^2, a good description of the F_2 data (and of the total $\gamma^* p$ cross section, $\sigma_{tot}^{\gamma^* p}$). This description corresponds to the emission of partons from a single chain (Fig.1a). However, at low Q^2 where the transition to nonperturbative strong interaction physics starts, this simple picture has to supplemented with corrections. First, there exists a class of models [1] which successfully describe this transition region; these models are based upon the idea of parton saturation: they assume the existence of multiple parton chains (Fig.1b) which interact with each other, and they naturally explain the observed scaling behavior, $F_2(Q^2, x) \approx F_2(Q^2/Q_s^2(x))$ with $Q_s^2(x) = Q_0^2(1/x)^\lambda$. Next, in the photoproduction region, $Q^2 \approx 0$, direct evidence for the presence of multiple interactions also comes from the analysis of final states [2]. A further strong hint at the presence of multi-chain configurations comes from the observation of a large fraction of diffractive final states in deep inelastic scattering at HERA. In the final states analysis of the linear QCD evolution equations, it is expected that the produced partons are not likely to come with large rapidity intervals between them. In the momentum-ordered single chain picture (Fig.1a), therefore, diffractive final states should be part of the initial conditions (inside the lower blob in Fig.1a), i.e. they should lie below the scale Q_0^2 which seperates the parton description from the nonperturbative strong interactions. This assignment of diffractive final states, however, cannot be complete. First, data have shown that the Pomeron which generates the rapidity gap in DIS diffraction is harder than in hadron - hadron scattering; furthermore, there are specific diffractive final states with momentum scales larger than Q_0^2, e.g. vector mesons built from heavy quarks and diffractive dijets (illustrated in Fig.2): the presence of such final states naturally requires corrections to the single chain picture (Fig.2b). From a t-channel point of view, both Fig.1b and Fig.2b belong to the same class of corrections, characterized by four gluon states in the t-channel.

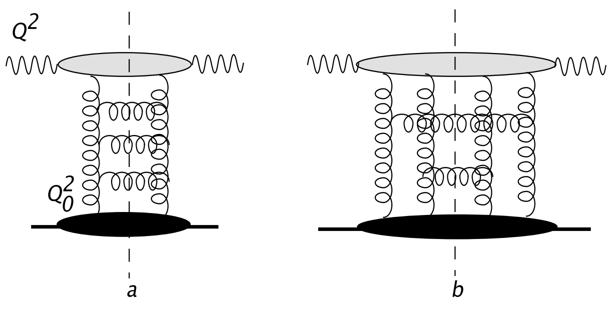


Fig. 1: Contributions to the total cross section $\sigma_{tot}^{\gamma^* p}$: (a) the single chain representing the linear QCD evolution equations; (b) gluon production from two different gluon chains.

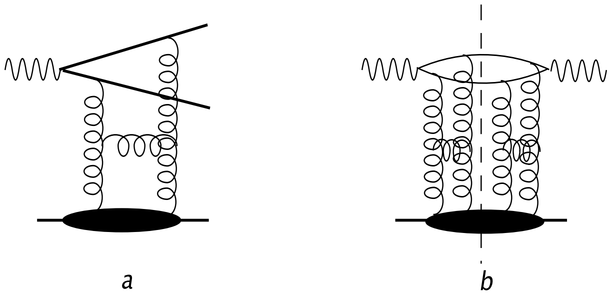


Fig. 2: Hard diffractive final states.(a) dijet production; (b) the diffractive cross section as s-channel discontinuity of a two-ladder diagram.

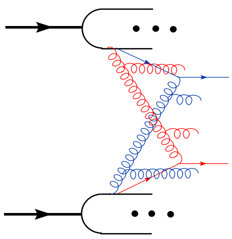


Fig. 3: Jet production in pp collisions from two different parton chains

In proton-proton collisions corrections due to multiple interactions should be important in those kinematic regions where parton densities for small momentum fractions values and for not too large momentum scales are being probed, e.g. jet production near the forward direction. Another place could be the production of multijet final states (Fig.3): multiple jets may come from different parton chains, and these contributions may very well affect the background to new physics beyond the standard model. Moreover, the modelling of multijet configurations will be necessary for understanding the underlying event structure in pp collidions [3].

From the point of view of collinear factorization, multiple interactions with momentum ordered parton chains are higher-twist effects, i.e they are suppressed by powers of the hard momentum scale. At small x, however, this suppression is compensated by powers of the large logarithms, $\ln 1/x$: multiple interactions, therefore, are mainly part of small-x physics. In this kinematic region the Abramovsky-Gribov-Kanchelli (AGK) [4] rules can be applied to the analysis of multi-gluon chains, and it is the aim of this article to present a brief overview about the current status of the AGK rules in pQCD.

As we will discuss below, in the analysis of multiple parton chains the couplings of n gluons to the proton play an essential role. Regge factorization suggests that these coupling should be universal, i.e. the couplings in $\gamma^* p$ collisions at HERA are the same as those in pp scattering at the LHC. Therefore, a thorough analysis of the role of multiple interactions in deep inelastic electron-proton scattering at HERA should be useful for a solid understanding of the structure of events at the LHC.

2 Basics of the AGK cutting rules

The original AGK paper [4], which was written before the advent QCD, addresses the question how, in the optical theorem,

$$\sigma_{tot}^{pp} = \frac{1}{s} Im\, T_{2\to 2} \;=\; \sum_f \int d\Omega_f |T_{i\to f}|^2 \tag{1}$$

the presence of multi-Pomeron exchanges (Fig.4) in the total hadron-hadron cross section leads to ob-

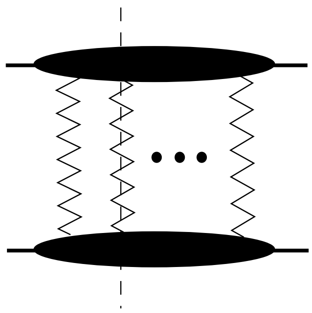


Fig. 4: s-cut through a multi-Pomeron exchange: the zig-zag lines stand for nonperturbative Pomerons.

servable effects in the final states (rhs of eq.(1)). Based upon a few model-independent assumptions on the couplings of multi-Pomeron exchanges to the proton, the authors derived simple 'cutting rules': different contributions to the imaginary part belong to different cuts across the multi-Pomeron diagrams, and each cut has its own, quite distinct, final state characteristics. As a result, the authors found counting rules for final states with different particle multiplicities, and they proved cancellations among rescattering corrections to single-particle and double-particle inclusive cross sections.

In the QCD description of hard (or semihard) final states a close analogy appears between (color singlet) gluon ladders and the nonperturbative Pomeron: multiple parton chains (for example, the two chains in Fig.1b) can be viewed as cuts through two perturbative BFKL Pomerons. In the same way as in the original AGK paper, the question arises how different cuts through a QCD multi-ladder diagram can be related to each other. In the following we briefly describe how AGK cutting rules can be derived in pQCD [5, 6]. In the subsequent section we will present a few new results which come out from pQCD calculations, going beyond the original AGK rules.

One of the few assumptions made in the original AGK paper states that the coupling of the Pomerons to the external particle are (i) symmetric under the exchange of the Pomerons (Bose symmetry), and (ii) that they remain unchanged if some of the Pomerons are beeing cut. These properties also hold in pQCD, but they have to be reformulated: (i') the coupling of (reggeized) gluons to external particles is symmetric under the exchange of reggeized gluons, and (ii') it remains unchanged if we introduce cutting lines between the gluons. In QCD, however, the color degree of freedom also allows for another possibility: inside the n-gluon state (with total color zero), a subsystem of two gluons can form an antisymmetric color octet state: in this case the two gluons form a bound state of a reggeized gluon (bootstrap property). For the case of $\gamma^*\gamma^*$ scattering, explicit calculations [7] have shown that the coupling of n gluons to virtual photons can be written as a sum of several pieces: the fully symmetric ('irreducible') one which satisfies (i') and (ii'), and other pieces which, by using the bootstrap property, can be reduced to symmetric couplings of a smaller number of gluons ('cut reggeons'). This decomposition is illustrated in Fig.5.

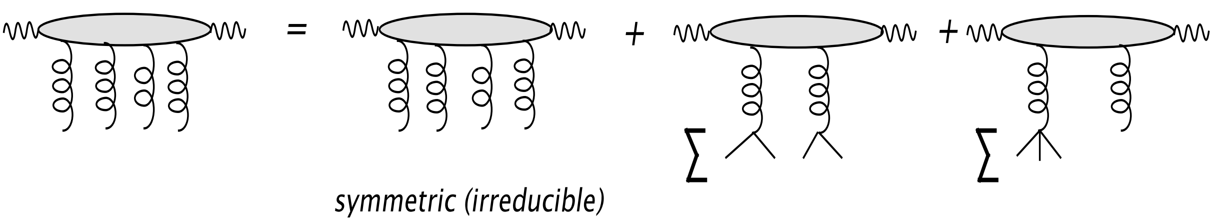


Fig.5 Decomposition of the coupling of four gluons to a virtual photon. In the last two terms on the rhs it is understood that we have to sum over different pairings of gluons at the lower end.

Since the bootstrap property is related to the regeization of the gluon and, therefore, is expected to be valid to all orders perturbation theory, also these properties of the couplings of multi-gluon states to external particles should be of general validity. In this short review we will mainly concentrate on the symmetric couplings.

As an illustrative example, we consider the coupling of four gluons to a proton. The simplest model of a symmetric coupling is a sum of three pieces, each of which contains only the simplest color structure:

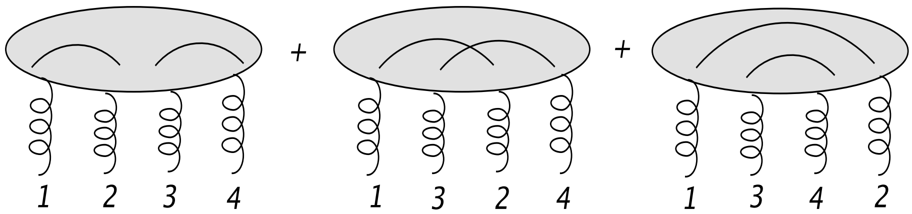


Fig.6 The symmetric coupling of four gluons to an external particle. The lines inside the blob denote the color connection, e.g. the first term has the color structure $\delta_{a_1 a_2}\delta_{a_3 a_4}$.

The best-known cutting rule for the four gluon exchange which follows [5, 6] from this symmetry requirement is the ratio between the three different pairings of lines:

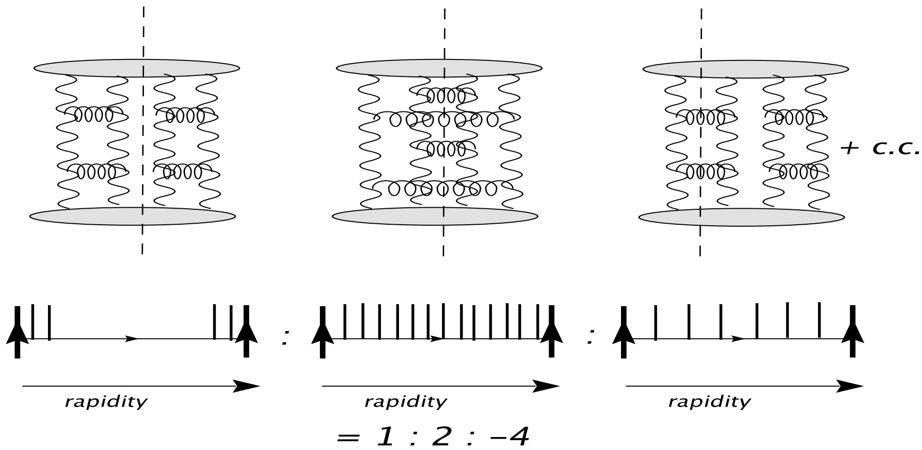


Fig 7: different cutting lines in the four-gluon exchange.

Each term, on the partonic level, corresponds to a certain multiplicity structure of the final state: a rapidity gap ('zero multiplicity'), double multiplicity, and single multiplicity. Simple combinatorics then leads to the ratio

$$1 : 2 : -4. \tag{2}$$

In order to be able to generalize and to sum over an arbitrary number of gluon chains, it is convenient to use an eikonal ansatz:

$$N_{2n}^{A}(\boldsymbol{k}_1, a_1; \ldots; \boldsymbol{k}_{2n}, a_{2n}; \omega) =$$

$$\frac{1}{\sqrt{(N_c^2-1)^n}}\left(\sum_{Pairings} \phi^A(\boldsymbol{k}_1, \boldsymbol{k}_2; \omega_{12})\delta_{a_1 a_2} \cdot \ldots \cdot \phi^A(\boldsymbol{k}_{2n-1}, \boldsymbol{k}_{2n}; \omega_{2n-1,2n})\delta_{a_{2n-1}a_{2n}}\right). \tag{3}$$

Inserting this ansatz into the hadron - hadron scattering amplitude, using the large-N_c approximation, and switching to the impact parameter representation, one obtains, for the contribution of k cut gluon ladders, the well-known formula:

$$ImA_k = 4s\int d^2\boldsymbol{b}e^{i\boldsymbol{q}\boldsymbol{b}}P(s,\boldsymbol{b}) \tag{4}$$

where

$$P(s, \boldsymbol{b}) = \frac{[\Omega(s, \boldsymbol{b})]^k}{k!} e^{-\Omega(s, \boldsymbol{b})}, \tag{5}$$

and Ω stands for the (cut) two-gluon ladder.

Another result [6] which follows from the symmetry properties of the n gluon-particle coupling is the cancellation of rescattering effects in single and double inclusive cross sections. In analogy with the AGK results on the rescattering of soft Pomerons, it can be shown that the sum over multi-chain contributions and rescattering corrections cancels (Fig.8),

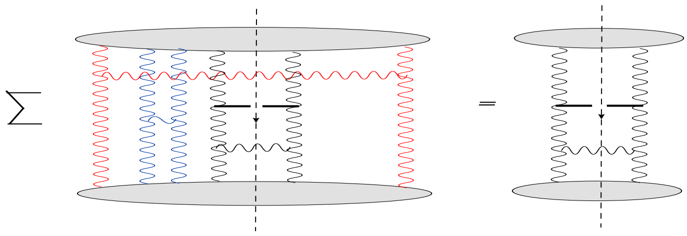


Fig 8: AGK cancellations in the one-jet inclusive cross section.

leaving only the single-chain contribution (in agreement with the factorization obtained in the collinear analysis). This statement, however, holds only for rescattering between the two projectile: it does not affect the multiple exchanges between the tagged jet and the projectile (Fig.9) which require a seperate discussion (see below).

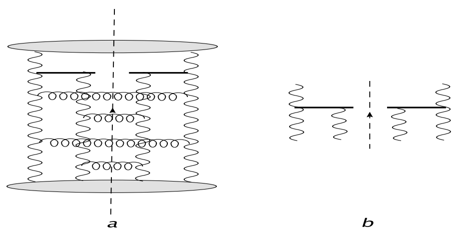


Fig 9: (a) Nonvanishing rescattering corrections in the one-jet inclusive cross section; (b) a new vertex: $g + 2g \rightarrow jet$.

All these results can be generalized to include also the soft Pomeron: all one needs to assume is that the couplings of soft Pomerons and reggeized gluons are symmetric under interchanges, and they are not altered if cutting lines are introduced.

3 New results

Explicit calculations in QCD lead to futher results on multiple interactions. First, in the four gluon exchange there are other configurations than those shown in Fig.7; one example is depicted in Fig.10. Here the pairing of gluon chains switches from $(14)(23)$ in the upper part (= left rapidity interval) to $(12)(34)$ in the lower part (= right rapidity interval).

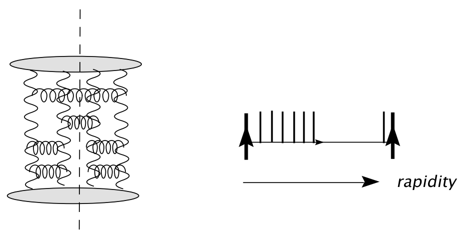


Fig 10: Decomposition into two rapidity intervals: the upper (left) interval has double multiplicity, the lower (right) one corresponds to a rapidity gap.

One can show that the ratio $1:2:-4$ holds for each rapidity interval. In [6] this has been generalized to an arbitrary number of exchanged gluon lines.

Another remark applies to the applicability of the cutting rules to rescattering corrections in the single jet inclusive cross section (Fig.9). Below the jet vertex we, again, have an exchange of four gluon lines, similar to the diagram in the middle of Fig.7. As to the cutting rules, however, there is an important difference between the two situations. In Fig.7, the blob above the four gluons is totally inclusive, i.e. it contains an unrestricted sum over s-channel intermediate states, whereas in Fig.9 the part above the four gluon state is semi-inclusive , i.e. it contains the tagged jet. This 'semi-inclusive' nature destroys the symmetry above the four gluon states, and the cutting rules have to be modified [8, 9]. In particular, eqs.(3) - (4) are not applicable to the rescattering corrections between the jet and projectile. A further investigation of these questions is in progress [10].

Finally a few comments on reggeization and cut reggeons. Clearly there are more complicated configurations than those which we have discussed so far; an example appears in $\gamma^* p$ scattering (deep inelastic electron proton scattering). In contrast to pp scattering, the coupling of multi-gluon chains to the virtual photon can be computed in pQCD, and the LO results, for the case of $n = 4$ gluons, are illustrated in Fig.11.

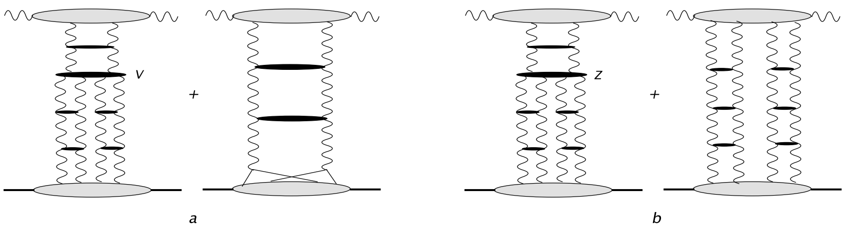


Fig.11: Four-gluon contributions to $\gamma^* p$ proton scattering: two equivalent ways of summing over all contributions. (a) the decomposition of Fig.5 with the pQCD triple Pomeron vertex. (b) an alternative way of summation which explicitly shows the coupling of two Pomerons to the photon vertex and which leads to a new vertex Z.

It turns out that we have two alternative possibilities: in the completely inclusive case (total cross section), it is convenient to chose Fig.11a, i.e. the sum of all contributions can be decomposed into two sets of diagrams. In the first set, at the top of the diagram two gluons couple to the quark-antiquark pair, and the subsequent transition to the four-gluon state goes via the pQCD triple Pomeron vertex. This vertex, as a function of the 4 gluons below, has the symmetry properties described above. As a result, we can apply the cutting rules to the four gluon state, as discussed before. However, there is also the second term in Fig.11a, which consists of a two gluon state only: this is the reggeizing contribution we have mentioned before. As indicated in the figure, the splitting of the reggized gluons at the bottom amounts to a change in the (nonperturbative) coupling. We want to stress that, because of the inclusive nature of this set of diagrams, the triple Pomeron vertex V in Fig.11a, similar to the BFKL kernel, contains both real and virtual contributions. For this reason, the decomposition in Fig.11a is applicable to inclusive cross sections, and it is not convenient for investigating specific final states such as, for example, diffractive final states with a fixed number of quarks and gluons in the final state.

There exists an alternative way of summing all contributions (Fig.11b) which is completely equivalent to Fig.11a but allows to keep track of diffractive $q\bar{q}$, $q\bar{q}g$,... final states: this form is illustrated in Fig.11b. One recognizes the 'elastic intermediate state' which was not visible in Fig.11a, and the new triple Pomeron vertex Z which contains only real gluon production. This vertex Z, as discussed in [11] is no longer symmetric under permutations of the gluons at the lower end; consequently, we cannot apply the AGK cutting rules to the four gluon states below.. These findings for multiple scattering effects in DIS imply, strictly speaking, that cross sections for diffractive $q\bar{q}$ or $q\bar{q}g$ states cannot directly be inserted into the counting rules (2).

Also pp scattering will contain corrections due to multiple interactions which are more complex. There are, for example, graphs which contain the $2 \rightarrow 4$ gluon vertex V, leading to a change of the number of gluon lines (Fig.12).

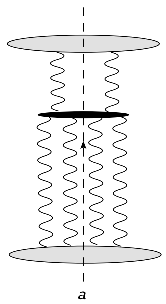


Fig 12: A correction in which the number of lines changes.
The black vertex denotes the $2 \rightarrow 4$ gluon vertex.

Since this $2 \rightarrow 4$ gluon vertex, as a function of the four gluons below the vertex, satisfies the symmetry requirements listed above, we can apply our previous analysis to the cutting lines below the vertex. In addition, however, one can ask how the lines continue above the $2 \rightarrow 4$ gluon vertex: we show two examples, one of them containing a cut (reggeized) gluon. Concentrating on this two-gluon state (i.e. we imagine that we have already summed over all possible cutting lines below the vertex V), the counting rules are quite different: in contrast to the even-signature Pomeron, the gluon is a odd-signature reggeon. Consequently, the cut gluon is suppressed w.r.t. the uncut gluon by one power in α_s, and this suppression leads to the following hierarchy of cutting lines: the cut between the gluons belongs to leading order, the cut through one of the two reggeized gluons is supressed by one power in α_s, the cut through both reggeized gluons is double suppressed (order α_s^2). A closer analysis of this question is under investigation [10].

4 Conclusions

Corrections due to multiple interactions seem to be important in DIS at small x and low Q^2; they are expected to play a significant role also in multijet production in pp scattering. The study of the AGK rules to pQCD provides help in understanding the systematics of multiple gluon chains. Results described in this review represent the beginning of a systematic analysis. We have listed a few questions which require further work.

As an immediate application, we believe that a quantitative analysis of multiple scattering at HERA will provide a useful input to the modelling of final states at the LHC.

A question of practical importance which we have not addressed at all is the hadronization of partonic final states. All statements on ratios of 'particle densities in the final states' made in this paper refer to the parton (gluon) level. However, the hadronization of events which, for example, belong to a double-cut ladder configuration may be quite different from the one obtained by applying just the normal single-chain hadronization to each chain seperately. The answer to this question [1] goes beyond the AGK analysis discussed in this paper.

References

[1] K.Golec-Biernat and M.Wusthoff, *Phys.Rev.* **D59**: 014017,1999; e-Print Archive: hep-ph/9807513; *Phys.Rev.***D60**: 114023,1999; e-Print Archive: hep-ph/9903358.
J.Bartels, K.Golec-Biernat, H.Kowalski, *Phys.Rev.***D66**: 014001,2002; e-Print Archive: hep-ph/0203258.

[2] C.Gwenlan, *Acta Phys.Polon.***B35**: 377-386,2004.

[1] I thank G.Gustafson for a very useful discussion on this point.

[3] T.Sjostrand and P.Z.Skands, *Eur.Phys.J.***C39**: 129-154, 2005; e-Print Archive: hep-ph/0408302 and references therein.

[4] V.A.Abramovsky, V.N.Gribov, O.V.Kanchelli, *Yad.Fiz.* **18**, 595 (1973) [Sov.J.Nucl.Phys.**18** 308 (1974)].

[5] J.Bartels, M.G.Ryskin, *Z.Phys.***C76**: 241-255,1997; e-Print Archive: hep-ph/9612226

[6] J.Bartels, M.Salvadore, G.P.Vacca, *Eur.Phys.J.* **C42**: 53-71,2005; e-Print Archive: hep-ph/0503049.

[7] J.Bartels, M.Wüsthoff, *Z.Phys.***C66**: 157-180,1995.

[8] Yu.V.Kovchegov, K.Tuchin, *Phys.Rev.***D65**: 074026,2002; e-Print Archive: hep-ph/0111362.

[9] M.Braun, *Eur.Phys.J.***C42**: 169-181,2005; e-Print Archive: hep-ph/0502184.

[10] J.Bartels, M.Salvadore, G.P.Vacca, in preparation.

[11] J.Bartels, M.Braun, G.P.Vacca, *Eur.Phys.J.* **C40**: 419-433,2005; e-Print Archive: hep-ph/0412218.

Multiple Interactions in DIS

Henri Kowalski
Deutsches Elektronen Synchrotron DESY, 22603 Hamburg

Abstract
The abundance of diffractive reactions observed at HERA indicates the presence of multiple interactions in DIS. These interactions are analysed, first in a qualitative way, in terms of QCD Feynman diagrams. Then a quantitative evaluation of diffractive and multiple interaction is performed with the help of the AGK cutting rules applied within an Impact Parameter Dipole Saturation Model. The cross-sections for multiple and diffractive interactions are found to be of the same order of magnitude and to exhibit a similar Q^2 dependence.

1 Introduction

One of the most important observations of HERA experiments is the rapid rise of the structure function F_2 with decreasing x indicating the presence of abundant gluon radiation processes [1]. The observation of a substantial diffractive component in DIS processes, which is also quickly rising with decreasing x, is equally important. The diffractive contribution at HERA is of a leading-twist type, i.e. the fraction of diffractive events remains constant or decreases only logarithmically with increasing Q^2. The presence of a substantial diffractive component suggests that, in addition to the usual partonic single ladder contribution, also multi-ladder processes should be present.

In this talk I will first discuss the general role of multi-ladder contributions in DIS scattering, called for historical reasons multi-Pomeron processes. The concept of a Pomeron is very useful in the discussion of high energy scattering processes since it relates, by the AGK cutting rules [2], seemingly different reactions like inclusive, diffractive and multiple scattering. I will present a numerical estimate of the magnitude of diffractive and of multi-Pomeron contributions, using AGK cutting rules within a dipole model which has been shown to provide a good description of HERA DIS data [3].

2 General Analysis

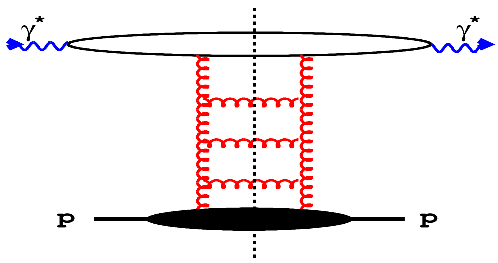


Fig. 1: *The single gluon-ladder contribution to the total $\gamma^* p$ cross section. The blob at the lower end of the diagrams contains the physics below the scale Q_0^2 which seperates hard from soft physics, whereas the blob at the upper end contains hard physics that can be described by pQCD. The dashed line denotes the cut.*

Let us first recall that the main properties of HERA interactions can be related to the properties of the elastic amplitude, $A_{\gamma^* p \to \gamma^* p}$, which, by the optical theorem, is directly related to the total $\gamma^* p$ cross-section:

$$\sigma_{\gamma^* p} = \frac{1}{W^2} Im\, A_{\gamma^* p \to \gamma^* p}(W^2, t = 0). \tag{1}$$

Here W denotes the γ^*p CMS energy and t the 4-momentum transfer of the elastically scattered proton. At not too small Q^2, the total cross section is dominated by the single ladder exchange shown in Fig. 1; the ladder structure also illustrates the linear DGLAP evolution equations that are used to describe the F_2 data. In the region of small x, gluonic ladders are expected to dominate over quark ladders. The cut lines in Fig. 1 mark the final states produced in a DIS event: a cut parton (gluon) hadronizes and leads to jets or particles seen in the detector. It is generally expected that partons produced from a single chain are unlikely to generate large rapidity gaps between them, since large gaps are exponentially suppressed as a function of the gap size. Therefore, in the single ladder contribution of Fig. 1, diffractive final states only reside inside the blob at the lower end, i.e. lie below the initial scale Q_0^2.

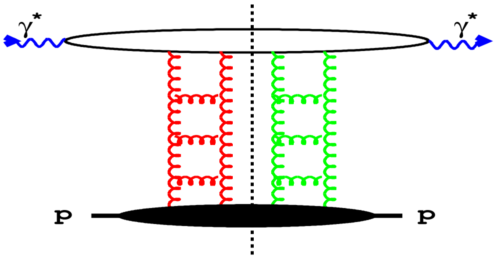


Fig. 2: *The double-gluon ladder contribution to the inclusive diffractive γ^*p cross section.*

The properties of diffractive reactions at HERA, however, give clear indications that significant contributions from multi-ladder exchanges should be present: not all diffractive final states are soft, in particular the diffractive production of jets and charm was observed [4, 5]. In addition, the inclusive diffractive cross-section is rising as quickly as the total cross-section with increasing W [6] and the exclusive diffractive production of J/Ψ and Υ vector meson exhibits a rise with energy which is about twice as fast [7]. In short, the Pomeron exchanged in inclusive diffractive DIS is harder than the hadronic soft Pomeron and therefore, one should expect that the majority of the observed diffractive final states cannot be absorbed into the blob of soft physics of Fig. 1. Instead, double ladder exchange, Fig. 2, provides a potential source for these harder diffractive states: the cut blob at the upper end may contain $q\bar{q}$ and $q\bar{q}g$ states which hadronize into harder jets or particles. Further evidence for the presence of multi-ladder contribution comes from saturation models which have been shown to successfully describe HERA F_2 data in the transition region at low Q^2 and small x: these models are explicitly built on the idea of summing over multiple exchanges of single ladders (or gluon densities).

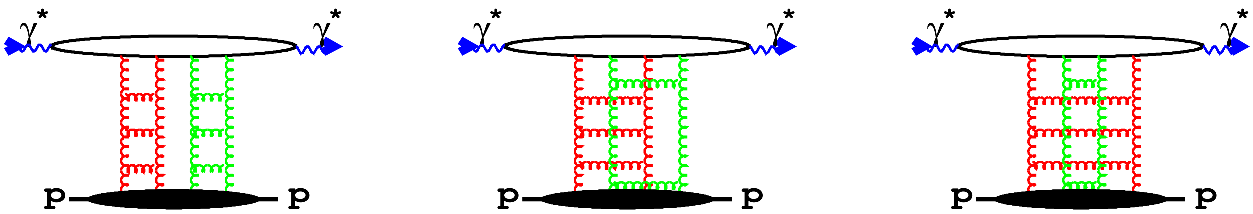


Fig. 3: *The double-gluon ladder contribution to the elastic γ^*p amplitude*

Let us analyse the content of a double ladder exchange contribution (for a more detailed analysis see ref. [8]). It is easiest to begin with the elastic γ^*p scattering amplitude, Fig. 3: from a t-channel point of view, the two gluon ladders form a four gluon intermediate state which has to be symmetric under permutations of the gluon lines (Bose symmetry). Therefore, on the amplitude level one cannot distinguish between different diagrams of Fig. 3. Invoking now the optical theorem, (1), different contributions to the total cross section correspond to different cuts through the two-ladder diagrams: they are shown in Fig. 4, ordered w.r.t. the density of cut gluons. In Fig. 4a, the cut runs between the two ladders: on the both sides of the cut there is a color singlet ladder, and we expect a rapidity gap between the upper blob (containing, for example, a diffractive $q\bar{q}$ final state) and the proton remnants inside the lower blob. Similarly, the diagram of Fig. 4b describes a single cut ladder with a final state similar to the one ladder contribution in Fig. 1; this contribution simply represents a correction to the one ladder contribution. Finally, the diagram of Fig. 4c belongs to final states with double density of cut partons. As outlined in [9], the correct counting of statistic factors and combinatorics leads to the result that the contributions shown in Fig. 4 a - c are identical, up to the overall counting factors $1 : -4 : 2$.

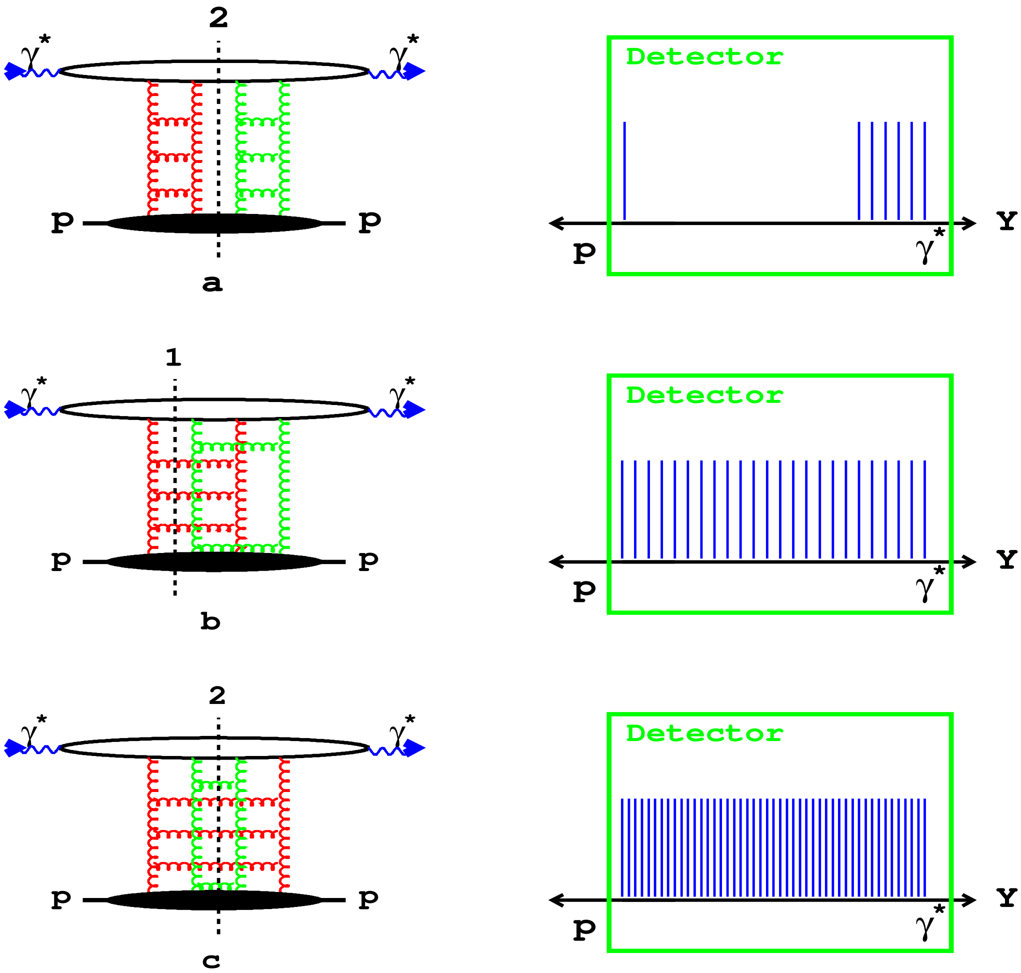


Fig. 4: *Three examples of 2-ladder contributions (lhs), with the corresponding, schematical, detector signatures (rhs).* Top row*: the diagram (a) with the cut positions (2) describes diffractive scattering.* Middle row*: the diagram (b) with the cut position (1) describes inclusive final states with single densityof cut partons.* Bottom row*: the diagram (c) with the cut position (2) describes inclusive final states with increased multiplicity.*

Experimentally it is easy to differentiate between diffractive and *single* or *multiple* inclusive final states since diffractive states exhibit large rapidity gaps. The *multiple* inclusive final states should also be distinct from the *single* inclusive ones since, at least naively, we would expect that in the *multiple* case the particle multiplicity should be considerably higher. At low x, however, the relation between the number of virtual states excited in the interaction (as measured by F_2) and the final particle multiplicity cannot be straight-forward since the growth of F_2 with decreasing x is faster than the multiplicity increase. This may indicate that the hadronization mechanism may be different from the string picture commonly used in the hadronization procedure of single chain parton showers. The influence of multiple scattering on the particle multiplicity of the final states should also be damped by the energy conservation. The cut through several Pomerons leads clearly to more gluons produced in the final state, but the available energy to produce particles in the hadronization phase remains the same. A detailed Monte Carlo program is therefore necessary to evaluate this effect.

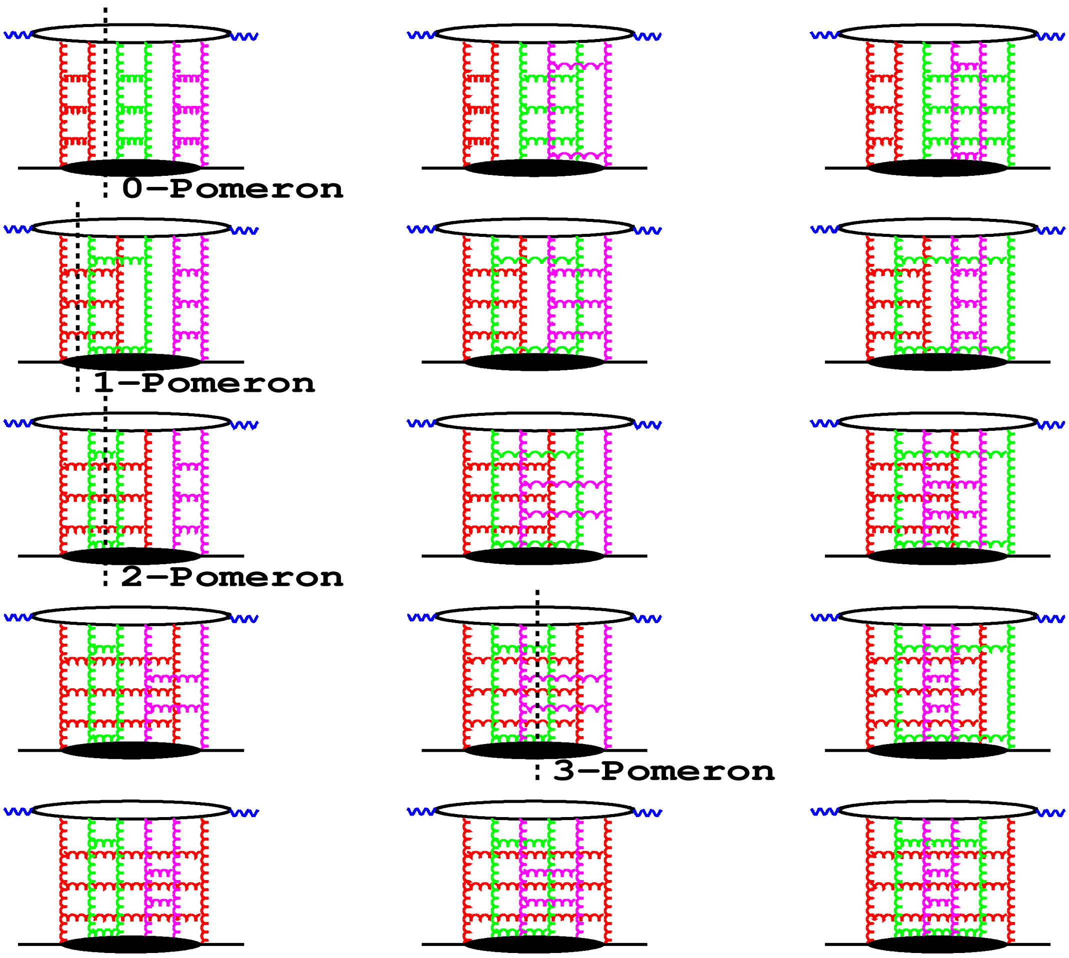


Fig. 5: *3-Pomeron contributions to the elastic $\gamma^* p$ amplitude. All 15 possible diagrams are shown with some examples of Pomeron cuts.*

The number of diagrams contributing to the reaction amplitude increases very quickly with the number of Pomerons. For the 3-Pomeron amplitude the gluons can be paired in 15 possible ways, shown in Fig. 5 with the examples of 0-Pomeron, 1-Pomeron, 2-Pomeron and 3-Pomeron cuts. For m-Pomerons the number of possible gluon pairs and also diagrams is:

$$(2m-1)(2m-3)(2m-5).... = (2m-1)!/(2^{m-1}(m-1)!).$$

Assuming that all the diagrams for a given multi-Pomeron exchange amplitude contribute in the same way, the above analysis suggests that the probability for different cuts to contribute should be given by the combinatorial factors. This is the content of the AGK rules which were obtained from the analysis of field theoretical diagrams well before QCD was established [2] and which relate the cross-section, σ_k, for observing a final state with k-cut Pomerons with the amplitudes for exchange of m Pomerons, $F^{(m)}$:

$$\sigma_k = \sum_{m=k}^{\infty} (-1)^{m-k}\, 2^m \, \frac{m!}{k!(m-k)!} F^{(m)}. \tag{2}$$

The same result is also obtained from a detailed analysis of the Feynman diagram contributions in QCD with the oversimplified assumption that only the symmetric part of the two-gluon couplings contributes [9].

3 Multiple Interactions in the Dipole Model

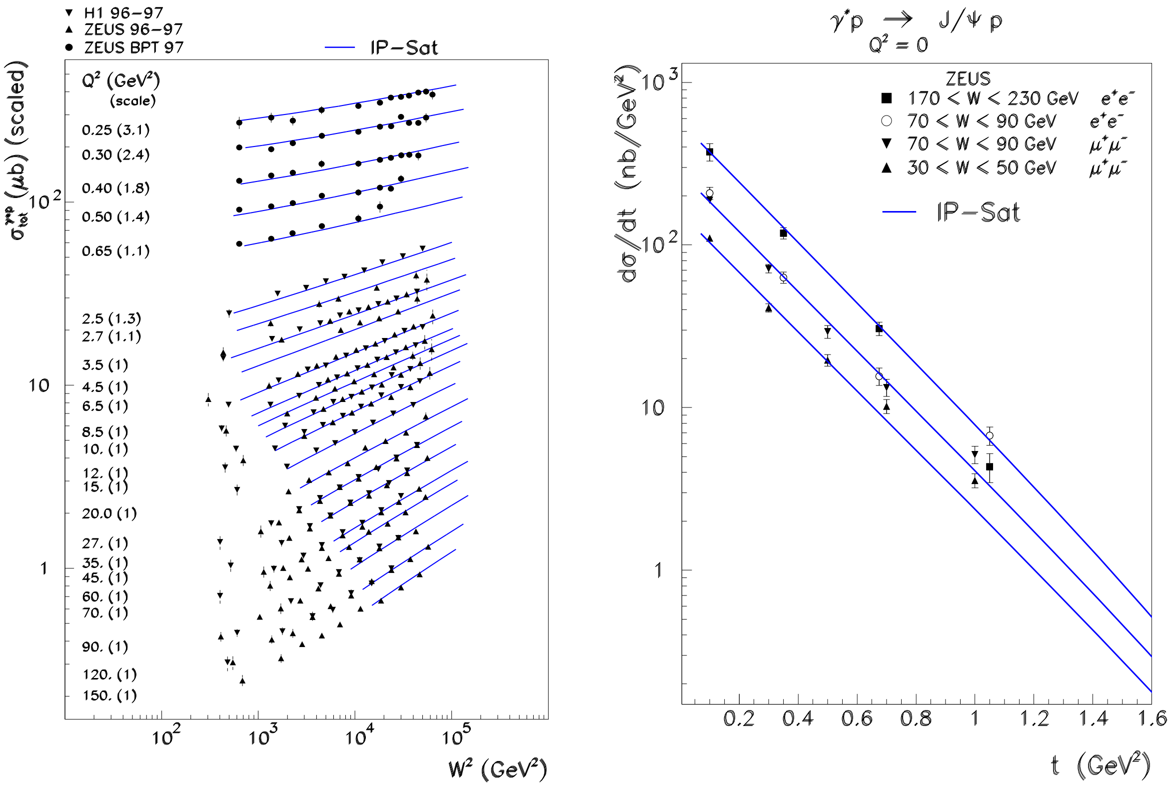


Fig. 6: *LHS: The $\gamma^* p$ cross-section as a function of W^2. RHS: The differential cross section for exclusive diffractive J/Ψ production as a function of the four-momentum transfer t. The solid line shows a fit by the IP saturation model.*

The properties of the multi-Pomeron amplitude and of the cut Pomeron cross-sections can be quantitatively studied in a dipole model. Let us first recall the main properties of the dipole picture, see ref. [10, 11] and [3]. In the model the $\gamma^* p$ interaction proceeds in three stages: first the incoming vitual photon fluctuates into a quark-antiquark pair, then the $q\bar{q}$ pair elastically scatters on the proton, and finally the $q\bar{q}$ pair recombines to form a virtual photon. The total cross-section for $\gamma^* p$ scattering, or equivalently F_2, is obtained by averaging the dipole cross-sections with the photon wave functions, $\psi(r, z)$, and integrating over the impact parameter, b:

$$F_2 = \frac{Q^2}{4\pi^2 \alpha_{em}} \int d^2 r \int \frac{dz}{4\pi} \psi^* \psi \int d^2 b \frac{d\sigma_{qq}}{d^2 b}. \tag{3}$$

Here $\psi^*\psi$ denotes the probability for a virtual photon to fluctuate into a $q\bar{q}$ pair, summed over all flavors and helicity states. The dipole cross-section is assumed to be a function of the opacity Ω:

$$\frac{d\sigma_{qq}}{d^2b} = 2\left(1 - \exp(-\frac{\Omega}{2})\right). \tag{4}$$

At small-x the opacity Ω can be directly related to the gluon density, $xg(x, \mu^2)$, and the transverse profile of the proton, $T(b)$:

$$\Omega = \frac{\pi^2}{N_C} r^2 \alpha_s(\mu^2)\, xg(x, \mu^2)\, T(b) \tag{5}$$

The parameters of the gluon density are determined from the fit to the total inclusive DIS cross-section, as shown in Fig. 6 [3]. The transverse profile was determined from the exclusive diffractive J/Ψ cross-sections shown in the same figure. The opacity function Ω determined in this way has predictive properties; it allows to describe other measured reactions, e.g. charm structure function or elastic diffractive J/Ψ production shown in Fig.7.

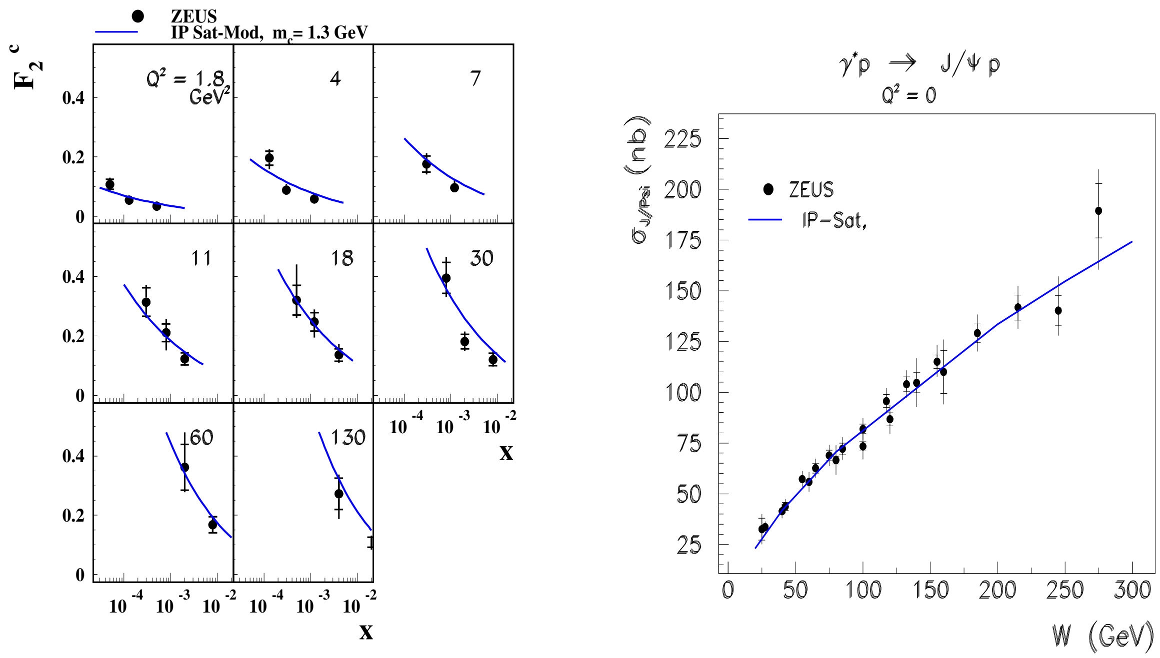


Fig. 7: *LHS: Charm structure function, F_2^c. RHS: Total elastic J/Ψ cross-section. The solid line shows the reswult of the IP saturation model.*

For a small value of Ω the dipole cross-section, eg. 4, is equal to Ω and therefore proportional to the gluon density. This allows to identify the opacity with the single Pomeron exchange amplitude of Fig. 1. The multi-Pomeron amplitude is determined from the expansion:

$$\frac{d\sigma_{qq}}{d^2b} = 2\left(1 - \exp(-\frac{\Omega}{2})\right) = 2\sum_{m=1}^{\infty}(-1)^{m-1}\left(\frac{\Omega}{2}\right)^m \frac{1}{m!} \tag{6}$$

as

$$F^{(m)} = \left(\frac{\Omega}{2}\right)^m \frac{1}{m!}, \tag{7}$$

since the dipole cross-section can be expressed as a sum of multi-Pomeron amplitudes [12] in the following way:

$$\frac{d\sigma_{qq}}{d^2b} = 2\sum_{m=1}^{\infty}(-1)^{m-1}\,F^{(m)} \tag{8}$$

The cross-section for k cut Pomerons is then obtained from the AGK rules, eq. 2, and from the multi-Pomeron amplitude, eq. 7, as:

$$\frac{d\sigma_k}{d^2b} = \sum_{m=k}^{\infty}(-1)^{m-k}\,2^m\,\frac{m!}{k!(m-k)!}\left(\frac{\Omega}{2}\right)^m\frac{1}{m!} = \frac{\Omega^k}{k!}\sum_{m=k}^{\infty}(-1)^{m-k}\frac{\Omega^{m-k}}{(m-k)!} \tag{9}$$

which leads to a simple expression:

$$\frac{d\sigma_k}{d^2b} = \frac{\Omega^k}{k!}\exp(-\Omega) \tag{10}$$

The diffractive cross-section is given by the difference between the total and the sum over all cut cross-sections:

$$\frac{d\sigma_{diff}}{d^2b} = \frac{d\sigma_{tot}}{d^2b} - \sum_{k=1}^{\infty}\frac{d\sigma_k}{d^2b} = 2\left(1-\exp\left(-\frac{\Omega}{2}\right)\right) - (1-\exp(-\Omega)) = \left(1-\exp\left(-\frac{\Omega}{2}\right)\right)^2 \tag{11}$$

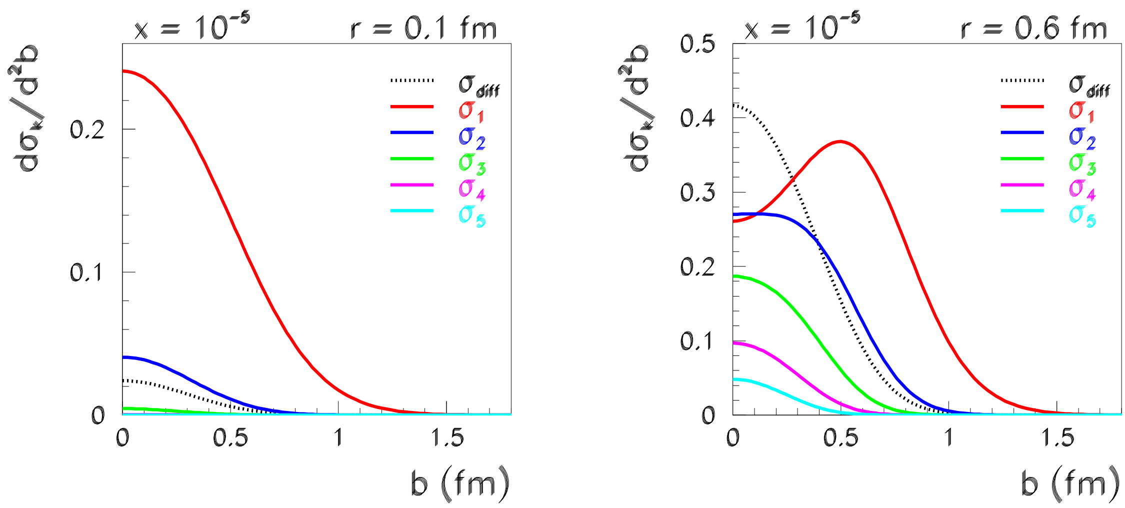


Fig. 8: *Examples of b dependence of various cut dipole and diffractive cross-sections.*

The cut cross-sections determined in the dipole model analysis of HERA data have several interesting properties shown in Fig. 8: for small dipoles ($r = 0.1$ fm) the opacity Ω is also small, so the single cut cross-section, σ_1, dominates. This leads to particle production emerging only from the one-cut pomeron, which should correspond, in the context of e.g. the LUND model, to a fragmentation of only one string. For larger dipoles ($r = 0.6$ fm) the dipole cross-section starts to be damped in the middle of the proton (at $b \approx 0$) by saturation effects. Therefore, the single cut cross-section is suppressed in the middle while the multiple cut cross-sections, σ_2, σ_3, etc, become substantial and increasingly concentrated in the proton center. These, fairly straight-forward properties of dipoles indicate that in the central scattering events the multiple scattering probability will be enhanced, which may lead at the LHC to substantial effects in a surrounding event multiplicity.

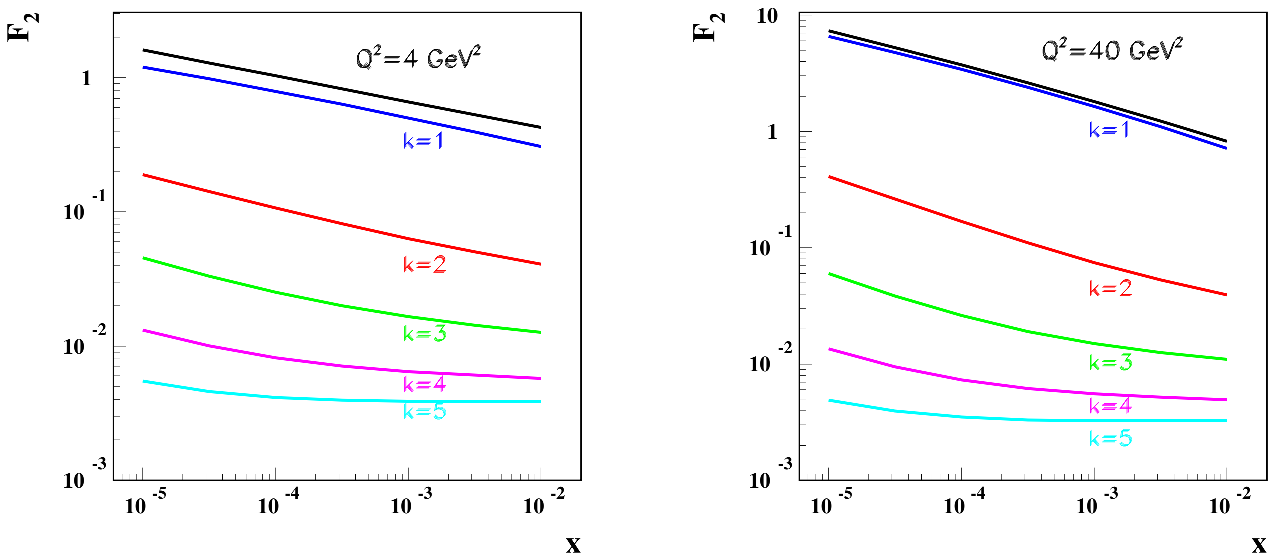


Fig. 9: *F_2 and the contributions of k-cut Pomeron processes, F_2^k.*

The contribution to F_2 from the k-cut Pomeron exchanges are computed in the analogous way to F_2:

$$F_2^k = \frac{Q^2}{4\pi^2\alpha_{em}} \int d^2r \int \frac{dz}{4\pi}\psi^*\psi \int d^2b \frac{d\sigma_k}{d^2b}. \tag{12}$$

These contributions are shown, together with F_2, as a function of x for two representative Q^2 values in Fig. 9. One finds that multiple interaction contributions, i.e. $k \geq 2$, in the perturbative region, at $Q^2 = 4$ GeV^2, are substantial. In the typical HERA range of $x \approx 10^{-3} - 10^{-4}$, the $k = 2$ contribution is around 10% of F_2 and the contributions of higher cuts are also non-negligible. For example, the contribution of the 5-cut Pomeron exchanges is still around 0.5%, which means that at HERA, many thousand events may come from this type of process. Figure 10 shows the fraction of the multpile interaction processes, $F_2^{MI} = F_2^{k=2} + F_2^{k=3} + F_2^{k=4} + F_2^{k=5}$ in F_2, at the same Q^2 values. At $Q^2 = 4$ GeV^2 the fraction of multiple scattering events is around 14% and at $Q^2 = 40$ GeV^2 around 6%, in the HERA x region, which indicates that the decrease of multiple scattering with increasing Q^2 is only logarithmic. The fraction of diffractive processes, shown for comparison, is of the same order, and drops also logarithmically with Q^2. The logarithmic drop of the diffractive contribution expected in the dipole model is confirmed by the data [6].

The dipole model provides a straight-forward extrapolation to the region of low Q^2, which is partly perturbative and partly non-perturbative. Figure 11 shows the contribution to F_2 of k-cut Pomeron processes and the fractions of multiple interactions and diffractive processes at $Q^2 = 0.4$ GeV^2.

Note also that, as a byproduct of this investigation, the ratio of diffractive and inclusive cross-sections, F_2^D/F_2 is found to be almost independent of x, in agreement with the data and also other dipole model predictions [6, 13, 14]. The absolute amount of diffractive effects is underestimated, since the evaluation of diffraction through AGK rules is oversimplified. It is well known [14], that a proper evaluation of diffraction should also take into account the $q\bar{q}g$ contribution which is missing in the simple AGK schema.

In conclusion, we find that the impact parameter dependent dipole saturation model [3] reproduces well the main properties of the data and leads to the prediction that multiple interaction effects at

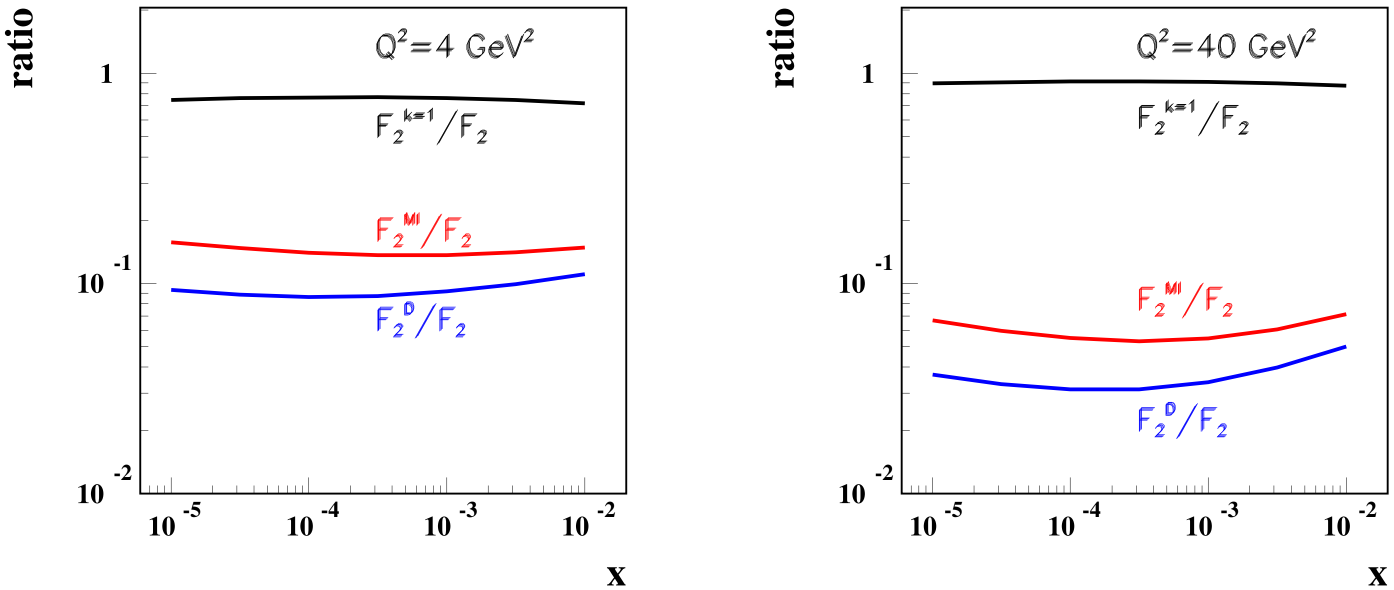


Fig. 10: *Fractions of single (k=1), multiple interaction (MI) and diffraction (D) in DIS*

HERA should be of the order of diffractive effects, which are known to be substantial. The multiple interaction effects should decrease slowly (logarithmically) with increasing Q^2, similarly to the diffractive contribution.

4 Acknowledgement

I would like to thank Jochen Bartels for many illuminating discussions about physics of the multi-ladder QCD diagrams.

References

[1] G. Wolf, International Europhysics Conference on HEP (2001).

[2] Abramovski, Gribov, Kancheli, Sov. J. Nucl. Phys. **18**, 308 (1974).

[3] H. Kowalski and D. Teaney, Phys. Rev. D **68**, 114005 (2003).

[4] ZEUS Collaboration, S. Chekanov et al., Nucl. Phys. B **672**, 3 (2003).

[5] ZEUS Collaboration, S. Chekanov et al., Phys. Letters B **545**, 244 (2002).

[6] ZEUS Collaboration, S. Chekanov et al., Nucl. Phys. B **713**, 3 (2005).

[7] A. Levy, Nucl. Phys. B (Proc. Suppl) **146**, 92 (2005).

[8] J. Bartels and M. Ryskin, Z. Phys. **C 76**, 241 (1997).

[9] J. Bartels, M. Salvadore, and G.P. Vacca, Eur. Phys. J. C **42**, 53 (2005).

[10] N.N. Nikolaev and B.G. Zakharov, Z. Phys. C **49**, 607 (1991).

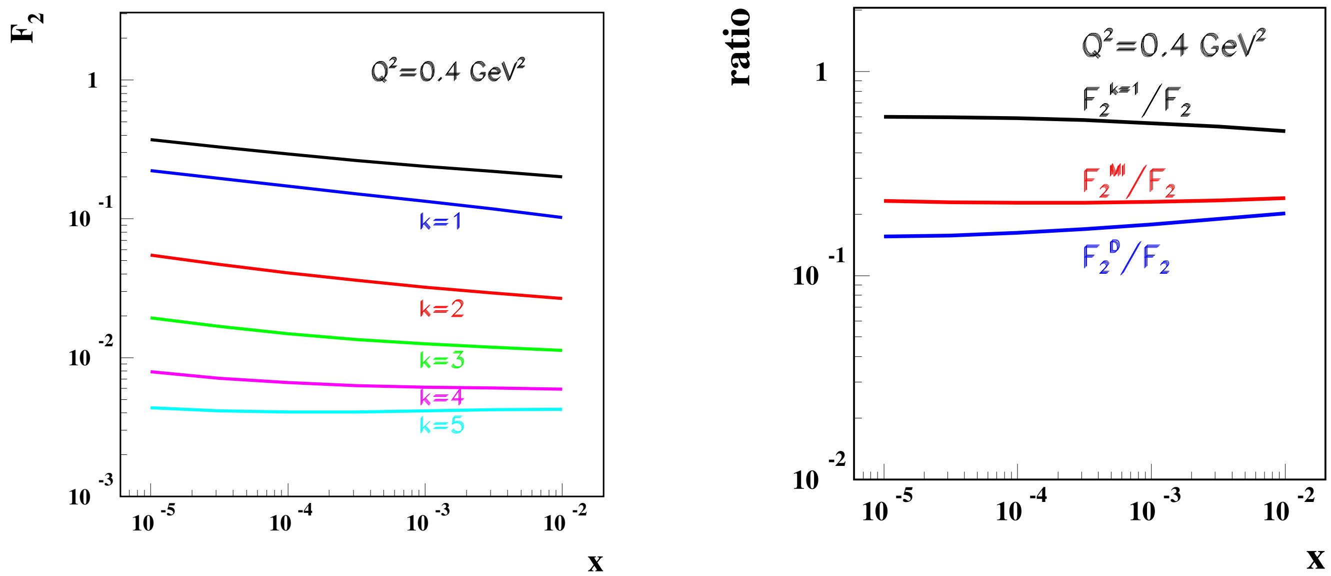


Fig. 11: *Left:* F_2 *and the contributions of* k*-cut Pomeron processes. Right: Fractions of single (k=1), multiple interaction (MI) and diffraction (D) in DIS at* $Q^2 = 0.4$ *GeV*^{2}

[11] A.H. Mueller, Nucl. Phys. B **415**, 373 (1994).

[12] A.H. Mueller and G.P. Salam, Nucl. Phys. B **475**, 293 (1996).

[13] K. Golec-Biernat and M. Wuesthoff, Phys. Rev. D **60**, 014023 (1999).

[14] J. Bartels, K. Golec-Biernat and H. Kowalski, Phys. Rev. D **66**, 014001 (2002).

From HERA to LHC through the Color Glass Condensate

Raju Venugopalan
Physics Department
Brookhaven National Laboratory
Upton, NY 11973, USA.

Abstract

A classical effective field theory, the Color Glass Condensate (CGC), provides a unified treatment of high parton density effects in both DIS and hadron-hadron collisions at very high energies. The validity and limitations of $k_\perp$ factorization can be studied in this effective theory. Multi-parton correlations in the effective theory are described by universal dipole and multipole operators. The evolution of these operators with energy provides a sensitive test of multi-parton dynamics in QCD at high energies.

1 Introduction

In the Bjorken limit of QCD, $Q^2 \to \infty$, $s \to \infty$, $x_{\rm Bj} \approx Q^2/s = {\rm fixed}$, we have a powerful framework to compute a large number of processes to high accuracy. Underlying this machinery is the Operator Product Expansion (OPE), where cross-sections are identified as a convolution of short distance "coefficient functions" which are process dependent and long distance parton distribution functions which are universal. The evolution of the parton distribution functions with x and Q^2 is described by splitting functions, which determine the probability of "parent" partons to split into a pair of "daughter" partons. Both coefficient functions and splitting functions for DIS inclusive cross-sections are now available to Next-Next-Leading-Order (NNLO) accuracy [1].

While this is a tremendous achievement, the contribution of high Q^2 processes to the total cross-section is very small. The bulk of the cross-section can perhaps be better understood in the Regge asymptotic limit: $x_{\rm Bj} \to 0$, $s \to \infty$, $Q^2 = {\rm fixed}$. The BFKL renormalization group equation [3] describes the leading $\alpha_S \ln(1/x)$ behavior of gluon distributions in this limit. The solutions of the BFKL equation predict that gluon distributions grow very rapidly with decreasing x. In the Regge asymptotics, since the transverse size of the partons is fixed, this growth of distributions will lead to the overlapping of partons in the transverse plane of the hadron. In this regime, contributions that were power suppressed in the BFKL scheme become important. These are recombination and screening effects which slow down the growth of gluon distributions leading ultimately to a saturation of these distributions [4, 5]. Such effects must appear at small x because the occupation number [1] of partons in QCD be at most of order $1/\alpha_S$.

Thus qualitatively, the competition between Bremsstrahlung and recombination/screening effects becomes of the same order when

$$\frac{1}{2\,(N_c^2-1)}\,\frac{x\,G(x,Q^2)}{\pi R^2 Q^2} \approx \frac{1}{\alpha_S(Q^2)}\,, \tag{1}$$

where R is the radius of the target. This relation is solved self-consistently when $Q \equiv Q_s(x)$. The scale $Q_s(x)$ is termed the saturation scale and it grows as one goes to smaller values of x. When $Q^2 \leq Q_s^2$, higher twist effects are important; at sufficiently small x, $Q_s^2 \gg \Lambda^2_{\rm QCD}$, which makes feasible a weak

[1] This corresponds to the number of partons per unit transverse area, per unit transverse momentum, per unit rapidity, in light cone gauge. This condition has its gauge invariant counterpart in the requirement that the field strength squared not exceed $1/\alpha_S$.

coupling analysis of these effects. At HERA, reasonable fits of small x inclusive and diffractive data for $x \leq 10^{-2}$ are obtained in saturation models with $Q_s^2(x) \approx Q_0^2\,(x_0/x)^\lambda$, with $Q_0^2 = 1$ GeV^2 and $x_0 = 3 \cdot 10^{-4}$. Detailed estimates suggest that the saturation scale for gluons is $Q_s(x) \approx 1.4$ GeV at $x \approx 10^{-4}$ [7]. The applicability of weak coupling techniques at these scales is dubious. Nevertheless, they cannot be ruled out since the effective scale at which the coupling runs can be larger than the estimate. Leading twist evolution of "shadowed" distributions at the saturation scale can extend out to significantly large values of x. A hint of this possibility is suggested by the fact that geometrical scaling- the dependence of cross-sections on the dimensionless ratio Q^2/Q_s^2 alone-extends out to $Q^2 \approx 450$ GeV^2 at HERA [8].

The possibility that weak coupling may apply at high energies is good news. Some of the remarkable regularities in high energy scattering data may be understood in a systematic way. The OPE, for instance, is no longer a good organizing principle since its usefulness is predicated on the twist expansion. In the next section, we will discuss an effective field theory approach which may provide a more efficient organizing principle at high parton densities.

2 The Color Glass Condensate

The physics of high parton densities can be formulated as a classical effective theory [6] because there is a Born-Oppenheimer separation between large x and small x modes [9] which are respectively the slow and fast modes in the effective theory. Large x partons are static sources of color charge for the dynamical wee (small x) parton fields. The generating functional of wee partons has the form

$$\mathcal{Z}[j] = \int [d\rho]\, W_{\Lambda^+}[\rho] \left\{ \frac{\int^{\Lambda^+} [dA]\delta(A^+) e^{iS[A,\rho] - j\cdot A}}{\int^{\Lambda^+} [dA]\delta(A^+) e^{iS[A,\rho]}} \right\} \tag{2}$$

where the wee parton action has the form

$$S[A,\rho] = \frac{-1}{4} \int d^4x\, F_{\mu\nu}^2 + \frac{i}{N_c} \int d^2x_\perp dx^- \delta(x^-)\, \mathrm{Tr}\left(\rho(x_\perp) U_{-\infty,\infty}[A^-]\right) . \tag{3}$$

In Eq. 2, ρ is a two dimensional classical color charge density and $W[\rho]$ is a weight functional of sources (which sits at momenta $k^+ > \Lambda^+$: note, $x = k^+/P^+_{\rm hadron}$). The sources are coupled to the dynamical wee gluon fields (which in turn sit at $k^+ < \Lambda^+$) via the gauge invariant term which is the second term on the RHS of Eq. 3. Here $U_{-\infty,\infty}$ denotes a path ordered exponential of the gauge field A^- in the x^+ direction. The first term in Eq. 3 is the QCD field strength tensor squared – thus the wee gluons are treated in full generality in this effective theory, which is formulated in the light cone gauge $A^+ = 0$. The source j is an external source – derivatives taken with respect to this source (with the source then put to zero) generate correlation functions in the usual fashion.

The argument for why the sources are classical is subtle and follows from a coarse graining of the effective action. The weight functional for a large nucleus is a Gaussian in the source density [6, 11], with a small correction for SU(N_c) coming from the $N_c - 2$ higher Casimir operators [10]. The variance of the Gaussian, the color charge squared per unit area μ_A^2, proportional to $A^{1/3}$, is a large scale – and is the only scale in the effective action [2]. Thus for $\mu_A^2 \gg \Lambda_{\rm QCD}^2$, $\alpha_S(\mu_A^2) \ll 1$, and one can compute the properties of the theory in Eq. 2 in weak coupling.

The saddle point of the action in Eq. 3 gives the classical distribution of gluons in the nucleus. The Yang-Mills equations can be solved analytically to obtain the classical field of the nucleus as a function of ρ: $A_{\rm cl.}(\rho)$ [6, 11, 12]. One can determine, for Gaussian sources, the occupation number $\phi = dN/\pi R^2/dk_\perp^2 dy$ (the number of partons per unit transverse momentum, per unit rapidity y, where $y = \ln(1/x)$) of wee partons in the classical field of the nucleus. One finds for $k_\perp \gg Q_s^2$, the Weizsäcker-Williams spectrum $\phi \sim Q_s^2/k_\perp^2$; for $k_\perp \leq Q_s$, one obtains a complete resummation to all orders in $k_\perp$,

[2] μ_A^2 is simply related in the classical theory to the saturation scale Q_s^2 via the relation $Q_s^2 = \alpha_S N_c \mu_A^2 \ln(Q_s^2/\Lambda_{\rm QCD}^2)$

which gives $\phi \sim \frac{1}{\alpha_S}\ln(Q_s/k_\perp)$. (The behavior at low $k_\perp$ can, more accurately, be represented as $\frac{1}{\alpha_S}\Gamma(0,z)$ where Γ is the incomplete Gamma function and $z = k_\perp^2/Q_s^2$ [13]).

A high energy hadron is a Color Glass Condensate for the following reasons [2]. The "color" is obvious since the parton degrees of freedom are colored. It is a glass because the sources, static on time scales much larger than time scales characteristic of the system, induce a stochastic (space-time dependent) coupling between the partons under quantum evolution – this is analogous to a spin glass. Finally, the matter is a condensate because the wee partons have large occupation numbers (of order $1/\alpha_S$) and have momenta peaked about Q_s. These properties are enhanced by quantum evolution in x. The classical field retains its structure–while the saturation scale grows: $Q_s(x') > Q_s(x)$ for $x' < x$.

Small fluctuations about the effective action in Eq. 3 give large corrections [14] of order $\alpha_S\ln(1/x)$. The Gaussian weight functional is thus fragile under quantum evolution of the sources. A Wilsonian renormalization group (RG) approach systematically treats these corrections [15]. In particular, the change of the weight functional $W[\rho]$ with x is described by the JIMWLK- non-linear RG equations [15]. These equations form an infinite hierarchy of ordinary differential equations for the gluon correlators $\langle A_1 A_2 \cdots A_n\rangle_Y$, where $Y = \ln(1/x)$ is the rapidity. The JIMWLK equation for an arbitrary operator $\langle O\rangle$ is

$$\frac{\partial\langle\mathcal{O}[\alpha]\rangle_Y}{\partial Y} = \left\langle \frac{1}{2}\int_{x_\perp,y_\perp} \frac{\delta}{\delta\alpha_Y^a(x_\perp)}\,\chi^{ab}_{x_\perp,y_\perp}[\alpha]\,\frac{\delta}{\delta\alpha_Y^b(y_\perp)}\mathcal{O}[\alpha]\right\rangle_Y , \tag{4}$$

where $\alpha = (\nabla_\perp^2)^{-1}\rho$. Here χ is a non-local object expressed in terms of path ordered (in rapidity) Wilson lines of α [2]. This equation is analogous to a (generalized) functional Fokker-Planck equation, where Y is the "time" and χ is a generalized diffusion coefficient. It illustrates the stochastic properties of operators in the space of gauge fields at high energies. For the gluon density, which is proportional to a two-point function $\langle\alpha^a(x_\perp)\alpha^b(y_\perp)\rangle$, one recovers the BFKL equation in the limit of low parton densities.

3 Dipoles in the CGC

In the limit of large N_c and large A ($\alpha_S^2 A^{1/3} \gg 1$), the JIMWLK hierarchy closes for the two point correlator of Wilson lines because the expectation value of the product of traces of Wilson lines factorizes into the product of the expectation values of the traces:

$$\langle \mathrm{Tr}(V_x V_z^\dagger)\mathrm{Tr}(V_z V_y^\dagger)\rangle \longrightarrow \langle \mathrm{Tr}(V_x V_z^\dagger)\rangle\,\langle \mathrm{Tr}(V_z V_y^\dagger)\rangle\,, \tag{5}$$

where $V_x = \mathcal{P}\exp\left(\int dz^-\alpha^a(z^-,x_\perp)T^a\right)$. Here $\mathcal{P}$ denotes path ordering in x^- and T^a is an adjoint SU(3) generator. In Mueller's dipole picture, the cross-section for a dipole scattering off a target can be expressed in terms of these 2-point dipole operators as [16, 17]

$$\sigma_{q\bar{q}N}(x,r_\perp) = 2\int d^2b\,\mathcal{N}_Y(x,r_\perp,b)\,, \tag{6}$$

where $\mathcal{N}_Y$, the imaginary part of the forward scattering amplitude, is defined to be $\mathcal{N}_Y = 1-\frac{1}{N_c}\langle\mathrm{Tr}(V_x V_y^\dagger)\rangle_Y$. Note that the size of the dipole, $\vec{r}_\perp = \vec{x}_\perp - \vec{y}_\perp$, and the impact parameter, $\vec{b} = (\vec{x}_\perp + \vec{y}_\perp)/2$. The JIMWLK equation for the two point Wilson correlator is identical in the large A, large N_c mean field limit to an equation derived independently by Balitsky and Kovchegov – the Balitsky-Kovchegov equation [18], which has the operator form

$$\frac{\partial\mathcal{N}_Y}{\partial Y} = \frac{\alpha_S N_c}{\pi}\,\mathcal{K}_{\mathrm{BFKL}}\otimes\left\{\mathcal{N}_Y - \mathcal{N}_Y^2\right\}\,. \tag{7}$$

Here $\mathcal{K}_{\mathrm{BFKL}}$ is the well known BFKL kernel. When $\mathcal{N} \ll 1$, the quadratic term is negligible and one has BFKL growth of the number of dipoles; when $\mathcal{N}$ is close to unity, the growth saturates. The approach to unity can be computed analytically [19]. The B-K equation is the simplest equation including both the Bremsstrahlung responsible for the rapid growth of amplitudes at small x as well as the repulsive many body effects that lead to a saturation of this growth.

A saturation condition which fixes the amplitude at which this change in behavior is significant, say $\mathcal{N} = 1/2$, determines the saturation scale. One obtains $Q_s^2 = Q_0^2 \exp(\lambda Y)$, where $\lambda = c\alpha_S$ with $c \approx 4.8$. The saturation condition affects the overall normalization of this scale but does not affect the power λ. In fixed coupling, the power λ is large and there are large pre-asymptotic corrections to this relation- which die off only slowly as a function of Y. BFKL running coupling effects change the behavior of the saturation scale completely–one goes smoothly at large Y to $Q_s^2 = Q_0^2 \exp(\sqrt{2b_0 c(Y+Y_0)})$ where b_0 is the coefficient of the one-loop QCD β-function. The state of the art computation of Q_s is the work of Triantafyllopoulos, who obtained Q_s by solving NLO-resummed BFKL in the presence of an absorptive boundary (which corresponds to the CGC) [20]. The pre-asymptotic effects are much smaller in this case and the coefficient $\lambda \approx 0.25$ is very close to the value extracted from saturation model fits to the HERA data [21]. Fits of CGC inspired models to the HERA data have been discussed elsewhere [22] and will not be discussed here.

4 Hadronic scattering and $k_\perp$ factorization in the CGC

Collinear factorization is the pQCD mechanism to compute hard scattering. At collider energies, a new window opens up where $\Lambda_{\rm QCD}^2 \ll M^2 \ll s$, where M is the invariant mass of the final state. In principle, cross-sections in this window can be computed in the collinear factorization language–however, one needs to sum up large logarithmic corrections in s/M^2. An alternative formalism is that of $k_\perp$-factorization [23, 24], where one has a convolution of $k_\perp$ dependent "un-integrated" gluon distributions from the two hadrons with the hard scattering matrix. In this case, the in-coming partons from the wavefunctions have non-zero $k_\perp$. Levin et al. [25] suggested that at high energies the typical $k_\perp$ is the saturation scale Q_s. The rapidity dependence of the unintegrated distributions is given by the BFKL or BK equations. However, unlike the structure functions, it has not been proven that these unintegrated distributions are universal functions.

At small x, both the collinear factorization and $k_\perp$ factorization limits can be understood in a systematic way in the framework of the Color Glass Condensate. The expectation value of an operator $\mathcal{O}$ can be computed as

$$\langle \mathcal{O} \rangle_Y = \int [d\rho_1]\,[d\rho_2]\, W_{x_1}[\rho_1]\, W_{x_2}[\rho_2]\, \mathcal{O}(\rho_1,\rho_2)\,, \tag{8}$$

where $Y = \ln(1/x_F)$ and $x_F = x_1 - x_2$. Quantum information, to leading logarithms in x, is contained in the source functionals $W_{x_1(x_2)}[\rho_1(\rho_2)]$ of the two hadrons. The operator $\mathcal{O}$ corresponding to the final state is expressed in terms of gauge fields $A^\mu[\rho_1,\rho_2](x)$. Inclusive gluon production in the CGC is computed by solving the Yang-Mills equations $[D_\mu, F^{\mu\nu}]^a = J^{\nu,a}$ for $A^\mu[\rho_1,\rho_2]$, where the current is given by $J^\nu = \rho_1\,\delta(x^-)\delta^{\nu+} + \rho_2\,\delta(x^+)\delta^{\nu-}$ with initial conditions determined by the Yang-Mills fields of the two hadrons before the collision. These are obtained self-consistently by matching the solutions of the Yang-Mills equations on the light cone [26]. Since we have argued in Section 2 that we can *compute* the Yang-Mills fields in the nuclei before the collision, the classical problem is in principle completely solvable. Quantum corrections not enhanced by powers of $\alpha_S \ln(1/x)$ can be computed systematically. Those terms enhanced by powers of $\alpha_S \ln(1/x)$ are absorbed into the weight functionals $W[\rho_{1,2}]$.

Hadronic scattering in the CGC can therefore be studied through a systematic power counting in the density of sources in powers of $\rho_{1,2}/k_{\perp;1,2}^2$. This power counting is more relevant at high energies than whether the incoming projectile is a hadron or a nucleus. In addition, one can study the applicability of collinear and $k_\perp$ factorization at small x in this approach.

The power counting is applicable as well to a proton at small x. The relevant quantity here is Q_s, which, as one may recall, is enhanced both for large A and small x. As long as $k_\perp \gg Q_s \gg \Lambda_{\rm QCD}$, one can consider the proton or nucleus as being dilute. To lowest order in $\rho_{p1}/k_\perp^2$ and $\rho_{p2}/k_\perp^2$, one can compute inclusive gluon production analytically [26]. At large transverse momenta, $Q_s \ll k_\perp$, the scattering can be expressed in a $k_\perp$-factorized form. The inclusive cross-section is expressed as the product of two unintegrated ($k_\perp$ dependent) distributions times the matrix element for the scattering. The comparison of this result to the collinear pQCD $gg \to gg$ process and the $k_\perp$ factorized $gg \to g$ was performed in Ref. [27]. At this order, the result is equivalent to the pQCD result first derived by Gunion and Bertsch [28]. This result for gluon production is substantially modified, as we shall discuss shortly, by high parton density effects either because the target is a large nucleus or because small values of x are being probed in the hadron (as in forward pp scattering).

$k_\perp$ factorization is a good assumption at large momenta for quark pair-production. This was worked out in the CGC approach by François Gelis and myself [29]. The result for inclusive quark pair production can be expressed in $k_\perp$ factorized form as

$$\frac{d\sigma_1}{dy_p dy_q d^2p_\perp d^2q_\perp} \propto \int \frac{d^2k_{1\perp}}{(2\pi)^2}\frac{d^2k_{2\perp}}{(2\pi)^2}\delta(k_{1\perp}+k_{2\perp}-p_\perp-q_\perp)$$
$$\times \phi_1(k_{1\perp})\phi_2(k_{2\perp})\frac{\mathrm{Tr}\left(\left|m_{ab}^{-+}(k_1,k_2;q,p)\right|^2\right)}{k_{1\perp}^2 k_{2\perp}^2}\,, \tag{9}$$

where ϕ_1 and ϕ_2 are the unintegrated gluon distributions in the projectile and target respectively (with the gluon distribution defined as $xG(x,Q^2) = \int_0^{Q^2} d(k_\perp^2)\,\phi(x,k_\perp)$). The matrix element $\mathrm{Tr}\left(\left|m_{ab}^{-+}(k_1,k_2;q,p)\right|^2\right)$ is identical to the result derived in the $k_\perp$–factorization approach [23, 24]. In the limit $|\vec{k_{1\perp}}|, |\vec{k_{2\perp}}| \to 0$, $\mathrm{Tr}\left(\left|m_{ab}^{-+}(k_1,k_2;q,p)\right|^2\right)\Big/(k_{1\perp}^2 k_{2\perp}^2)$ is well defined–after integration over the azimuthal angles in Eq. 9, one obtains the usual matrix element $|\mathcal{M}|^2_{gg\to q\bar{q}}$, recovering the lowest order pQCD collinear factorization result.

4.1 Gluon and quark production in forward pp and pA collisions

Many analytical results are available when one of the hadrons is dilute and the other is dense. This may correspond to either pA collisions or forward pp collisions. One solves the Yang–Mills equations $[D_\mu, F^{\mu\nu}] = J^\nu$ with the light cone sources $J^{\nu,a} = \delta^{\nu+}\,\delta(x^-)\,\rho_p^a(x_\perp) + \delta^{\nu-}\,\delta(x^+)\,\rho_A^a(x_\perp)$, to determine the gluon field produced-to lowest order in the source density of one projectile ($\rho_p/k_\perp^2 \ll 1$)and to all orders ($\rho_A/k_\perp^2 \sim 1$) in the source density of the other. The inclusive gluon production cross-section, in this framework, was first computed by Kovchegov and Mueller [30] and shown to be $k_\perp$ factorizable in Ref. [34, 31]. The "unintegrated" gluon distribution in the dense system however is here replaced by the gluon "dipole" distribution $\mathcal{N}_Y$ we discussed previously. It is no longer a leading twist object but includes all twists enhanced by high parton density effects. The well known "Cronin" effect observed in Deuteron-Gold collisions at RHIC is obtained in this formalism and can be simply understood in terms of the multiple scattering of a parton from the projectile with those in the target. The energy evolution of the dipole distribution is given by the BK equation, leading to a suppression of the Cronin effect at high densities due to the shadowing of nuclear distributions. This prediction appears to be confirmed by the RHIC data. The "dipole" operators extracted from DIS can therefore be used to predict inclusive hadron production in pp and pA collisions. One can similarly compute Drell-Yan and photon production in forward pp and pA collisions [33, 35].

Unlike gluon production, neither quark pair-production nor single quark production is strictly $k_\perp$ factorizable. The pair production cross-section can however still be written in $k_\perp$ factorized form as a product of the unintegrated gluon distribution in the proton times a sum of terms with three unintegrated distributions, $\phi_{g,g}$, $\phi_{q\bar{q},g}$ and $\phi_{q\bar{q},q\bar{q}}$. These are respectively proportional to 2-point (dipole), 3-point and 4-point correlators of the Wilson lines we discussed previously. Again, these operators include all twist contributions. For instance, the distribution $\phi_{q\bar{q},g}$ is the product of fundamental Wilson lines coupled to a $q\bar{q}$ pair in the amplitude and adjoint Wilson lines coupled to a gluon in the complex conjugate amplitude. For large transverse momenta or large-mass pairs, the 3-point and 4-point distributions collapse to the unintegrated gluon distribution, and we recover the previously discussed $k_\perp$-factorized result for pair production in the dilute/pp-limit. Single quark distributions are straightforwardly obtained and depend only on the 2-point quark and gluon correlators and the 3-point correlators. For Gaussian sources, as in the McLerran-Venugopalan-model, these 2-,3- and 4-point functions can be computed exactly as discussed in Ref. [32].

The situation gets complicated when one enters a regime where both projectiles are dense–as defined in our power counting. $k_\perp$ factorization breaks down decisively and analytical approaches are likely not possible. Nevertheless, numerical techniques have been developed, which allow the computation of final states, at least to leading logs in x [38].

The results for gluon and quark production in forward pp and pA or dA collisions (for a review, see Ref. [37]), coupled with the previous results for inclusive and diffractive [33, 35, 34, 36] distributions in DIS, suggest an important new paradigm. *At small x in DIS and hadron colliders, previously interesting observables such as quark and gluon densities are no longer the only observables to capture the relevant physics. Instead, they should be complemented by dipole and multipole correlators of Wilson lines that seem ubiquitous in all high energy processes and are similarly gauge invariant and process independent. The renormalization group running of these operators may be a powerful and sensitive harbinger of new physics.*

References

[1] A. Vogt, S. Moch and J. A. M. Vermaseren, Nucl. Phys. B **691**, 129 (2004); S. A. Larin, P. Nogueira, T. van Ritbergen and J. A. M. Vermaseren, Nucl. Phys. B **492**, 338 (1997).

[2] E. Iancu and R. Venugopalan, hep-ph/0303204, in QGP3, Ed. R. Hwa and X.-N. Wang, World Scientific (2004).

[3] E. A. Kuraev, L. N. Lipatov and V. S. Fadin, Sov. Phys. JETP **45**, 199 (1977) [Zh. Eksp. Teor. Fiz. **72**, 377 (1977)]; I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978) [Yad. Fiz. **28**, 1597 (1978)].

[4] L.V. Gribov, E.M. Levin, M.G. Ryskin, Phys. Rept. **100**, 1 (1983).

[5] A.H. Mueller, J-W. Qiu, Nucl. Phys. **B 268**, 427 (1986).

[6] L. D. McLerran and R. Venugopalan, Phys. Rev. D **49**, 2233 (1994); *ibid.*, 3352, (1994); *ibid.*, **50**, 2225 (1994).

[7] H. Kowalski and D. Teaney, Phys. Rev. D **68**, 114005 (2003).

[8] A. M. Stasto, K. Golec-Biernat and J. Kwiecinski, Phys. Rev. Lett. **86**, 596 (2001).

[9] L. Susskind, Phys. Rev. **165**, 1535, (1968).

[10] S. Jeon and R. Venugopalan, Phys. Rev. D **70**, 105012 (2004).

[11] Yu.V. Kovchegov, Phys. Rev. **D 54**, 5463 (1996); *ibid.*, **55**, 5445 (1997).

[12] J. Jalilian-Marian, A. Kovner, L.D. McLerran, H. Weigert, Phys. Rev. **D 55**, 5414 (1997).

[13] D. Triantafyllopoulos, hep-ph/0502114.

[14] A. Ayala, J. Jalilian-Marian, L.D. McLerran, R. Venugopalan, Phys. Rev. **D 52**, 2935 (1995); *ibid.*, **53**, 458 (1996).

[15] J. Jalilian-Marian, A. Kovner, A. Leonidov, H. Weigert, Phys. Rev. **D 59**, 014014 (1999); E. Iancu, A. Leonidov, L.D. McLerran, Nucl. Phys. **A 692**, 583 (2001).

[16] A. H. Mueller, Nucl. Phys. B **415**, 373 (1994).

[17] N. N. Nikolaev and B. G. Zakharov, Z. Phys. C**49**, (1991) 607.

[18] I. Balitsky, Nucl. Phys. **B 463**, 99 (1996); Yu.V. Kovchegov, Phys. Rev. **D 61**, 074018 (2000).

[19] E.M. Levin, K. Tuchin, Nucl. Phys. **B 573**, 833 (2000).

[20] D. N. Triantafyllopoulos, Nucl. Phys. B **648**, 293 (2003).

[21] K. Golec-Biernat, M. Wüsthoff, Phys. Rev. **D 59**, 014017 (1999).

[22] R. Venugopalan, arXiv:hep-ph/0412396.

[23] S. Catani, M. Ciafaloni, F. Hautmann, Nucl. Phys. B **366**, 135 (1991).

[24] J. C. Collins and R. K. Ellis, Nucl. Phys. B **360**, 3 (1991).

[25] E. M. Levin, M. G. Ryskin, Y. M. Shabelski and A. G. Shuvaev, Sov. J. Nucl. Phys. **53**, 657 (1991).

[26] A. Kovner, L. D. McLerran and H. Weigert, Phys. Rev. D **52**, 3809 (1995); Y. V. Kovchegov and D. H. Rischke, Phys. Rev. C **56**, 1084 (1997).

[27] M. Gyulassy and L. D. McLerran, Phys. Rev. C **56**, 2219 (1997).

[28] J. F. Gunion and G. Bertsch, Phys. Rev. D **25**, 746 (1982).

[29] F. Gelis and R. Venugopalan, Phys. Rev. D **69**, 014019 (2004).

[30] Y. V. Kovchegov and A. H. Mueller, Nucl. Phys. B **529**, 451 (1998).

[31] J. P. Blaizot, F. Gelis and R. Venugopalan, Nucl. Phys. A **743**, 13 (2004).

[32] J. P. Blaizot, F. Gelis and R. Venugopalan, Nucl. Phys. A **743**, 57 (2004); H. Fujii, F. Gelis and R. Venugopalan, hep-ph/0504047.

[33] F. Gelis and J. Jalilian-Marian, Phys. Rev. D **66**, 094014 (2002); Phys. Rev. D **66**, 014021 (2002); A. Dumitru and J. Jalilian-Marian, Phys. Rev. Lett. **89**, 022301 (2002).

[34] Y. V. Kovchegov and K. Tuchin, Phys. Rev. D **65**, 074026 (2002); D. Kharzeev, Yu. Kovchegov, K. Tuchin, Phys. Rev. D **68**, 094013 (2003).

[35] B. Z. Kopeliovich, J. Raufeisen, A. V. Tarasov and M. B. Johnson, Phys. Rev. C **67**, 014903 (2003).

[36] A. Kovner and U. A. Wiedemann, Phys. Rev. D **64**, 114002 (2001).

[37] J. Jalilian-Marian and Y. V. Kovchegov, arXiv:hep-ph/0505052.

[38] A. Krasnitz and R. Venugopalan, Nucl. Phys. B **557**, 237 (1999); Phys. Rev. Lett. **84**, 4309 (2000); A. Krasnitz, Y. Nara and R. Venugopalan, Phys. Rev. Lett. **87**, 192302 (2001); Nucl. Phys. A **717**, 268 (2003); Nucl. Phys. A **727**, 427 (2003); T. Lappi, Phys. Rev. C **67**, 054903 (2003).

Vector Boson Fusion at CMS

N. Amapane[1,a], *M. Arneodo*[2], *R. Bellan*[1], *S.Bolognesi*[1], *G. Cerminara*[1], *C. Mariotti*[3]
1) Torino University and INFN Torino, 2) University of Eastern Piedmont, Novara and INFN Torino, 3) INFN Torino, a) now at CERN

Abstract
The processes of boson-boson scattering and of Higgs production in boson-boson fusion hold the key to electroweak symmetry breaking. A preliminary study has been performed using a fast simulation of the CMS detector. The results are encouraging and suggest that, after few years of data taking at LHC, the region above 1 TeV can be explored, which is interesting if the Higgs is not found.

1 Vector Boson Fusion at CMS

1.1 Introduction

The Standard Model predicts that, without the Higgs boson, the scattering amplitude of the *longitudinally polarized vector boson* (V_L) fusion process violates unitarity at about 1-1.5 TeV. The longitudinal polarization of the *V* arises from the *V* getting massive, i.e. when the symmetry breaks spontaneously. The cross section as a function of the V_LV_L invariant mass will show a resonance at M(V_LV_L)=M(H) if the Higgs is there; otherwise, the cross section will deviate from the Standard Model prediction at high values of M(VV). Therefore, *VV* scattering can probe the Electroweak Symmetry Breaking with or without the assumption the Higgs mechanism.

1.2 The Signal Selection

Two channels have been studied using *Pythia* [1] and the *CMS Fast Simulation* [2]:

- $pp \to \mu\mu jjjj$ [3] through the processes:
  - $pp \to V_LV_Ljj \to Z_LZ_Ljj \to \mu\mu jjjj$,
  - $pp \to Z_LW_Ljj \to Z_LW_Ljj \to \mu\mu jjjj$.
- $pp \to \mu\nu jjjj$ [4] through the process:
  - $pp \to V_LV_Ljj \to W_LW_Ljj \to \mu\nu jjjj$.

The study has been done for high Higgs masses: $m_H = 500$ GeV and $m_H = 1000$ GeV, and for the no-Higgs scenario. The latter has been simulated in Pythia by setting $m_H = 10000$ GeV (the Higgs exchange diagram is suppressed by a m_H^2 term in the denominator of the Higgs propagator). The cross sections of the signal processes are shown in Table 1.

Table 1: Signal cross section (in fb) for different Higgs masses.

Processes	$m_H = 500$ GeV	$m_H = 1000$ GeV	$m_H = 10000$ GeV
$pp \to Z_LZ_Ljj \to \mu\mu jjjj$	9.1	3.0	1.7
$pp \to Z_LW_Ljj \to \mu\mu jjjj$	0.7	1.0	1.5
$pp \to W_LW_Ljj \to \mu\nu jjjj$	64.4	26.9	19.7

The signal has a well defined topology (see Figure 1):

- one μ^+ and one μ^- (or one μ and one ν) in the final state, with high p_T and low η coming from the *Z* (*W*) boson;

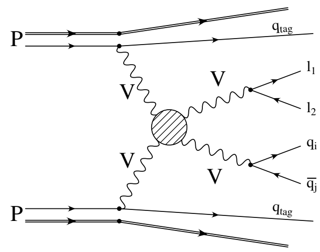


Fig. 1: The signal topology. l_1 and l_2 can be $\mu^{\pm}$ or μ and ν.

- two jets with high p_T and low η, coming from the vector boson decay;
- two energetic jets with high p_T, in the forward-backward regions (large η and $\Delta\eta$).

The aim of the work is to reconstruct the invariant mass of the *VV*-fusion system in both the channels and estimate its resolution. We also attempted a first estimate of the signal to background ratio assuming that the main background processes are:

- $t\bar{t}$ *background*: a six fermion final state, like the signal, but the jets are mainly in the central region; therefore, by requiring two jets at high η and with a large $\Delta\eta$ between them this kind of background can be rejected.
- *VV associated production*: a four fermion final state; it needs however to be kept under control in the case in which one boson decays leptonically since there are several jets from gluon radiation in the final state. The most effective variables to distinguish this background from the signal are the transverse momenta of the jets and of the leptons.
- *V plus one and two hard jets*: it is simple to reject this background because it has a topology not very similar to that of the signal and the additional jets have a very low p_T (since they are generated by the parton shower). However it is fundamental to keep it under control since it has a very large cross section.

The cross section of the background processes are shown in Table 2.

Table 2: Background cross section (in fb).

Background	Cross Section [fb]	**Background**	Cross Section [fb]
$t\bar{t}$, 1 μ	$622 \cdot 10^3$	$t\bar{t}$, 1 μ^- and 1 μ^+	$620 \cdot 10^3$
$ZZ \to \mu^-\mu^+jj$	653	$ZW \to \mu^-\mu^+jj$	663
$WW \to \mu\nu^+jj$	$11 \cdot 10^3$	$W + jj \to \mu\nu jj$	$77 \cdot 10^3$
$Z + jet \to \mu^-\mu^+j$	$13 \cdot 10^6$	$W + j \to \mu\nu j$	$184 \cdot 10^6$

1.3 The Results

A set of cuts has been applied to enhance the signal with respect to the background. A good resolution (estimated using MC info) on the most important observables has been achieved. In particular:

- $Z \to \mu\mu$ invariant mass: $R_z \sim 1.5\%$;
- $V \to jj$ invariant mass, $\mu\mu jjjj$ channel: $R_v \sim 13\%$;
- $V \to jj$ invariant mass, $\mu\nu jjjj$ channel: $R_v \sim 10\%$.

The difference between the two latter resolutions reflects the fact that for the $jjjj\mu\mu$ channel the pile-up has been considered whereas in the $jjjj\mu\nu$ it was not. The resolution on the energy scale of the process ($M_{inv}(VV)$) is:

- 4% for the $pp \rightarrow \mu\mu jjjj$ channel;
- 8% for the $pp \rightarrow \mu\nu jjjj$ channel.

The difference is due to the worse resolution on the neutrino p_T and p_z reconstruction. The resulting background efficiency is lower than one percent while the signal efficiency reaches 30% for the $jjjj\mu\mu$ channel and 50% for the $jjjj\mu\nu$ channel. A high significance ($S/\sqrt{B}$) has been achieved for an integrated luminosity of $100 fb^{-1}$: for the $\mu\mu jjjj$ samples it is about 8 in the interval $M_{inv}^{VV} \in [0,1]$ TeV for the Higgs mass set to 500 GeV and about 10 for $M_{inv}^{VV} > 1$ TeV for the no-Higgs scenario. Similar values have been obtained for the $\mu\nu jjjj$ channel: a significance of about 5, in the interval $M_{inv}^{VV} \in [0,1]$ TeV, for the Higgs mass set to 500 GeV and about 2.4 in the interval $M_{inv}^{VV} > 1$ TeV for the no-Higgs scenario. In the Figures 2 (no-Higgs scenario) and 3 ($m_H = 500$ GeV) the number of reconstructed events and the selection efficiency as a function of the *VV* invariant mass are shown.

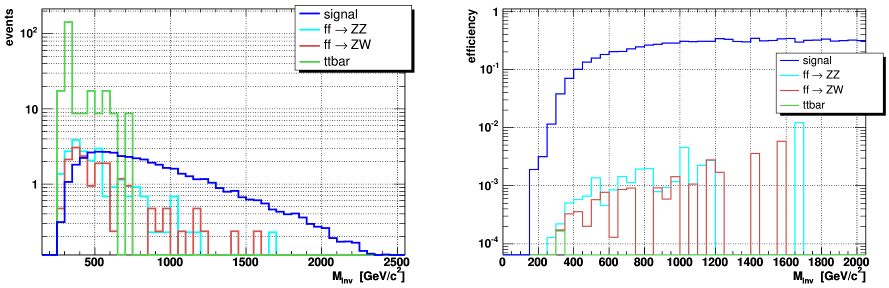


Fig. 2: (Left) Number of reconstructed events as a function of the *VV* invariant mass and (Right) the selection efficiency as a function of the invariant mass of the VV-fusion process; both for the $\mu\mu jjjj$ final state in the no-Higgs scenario and an integrated luminosity of 100 fb^{-1}.

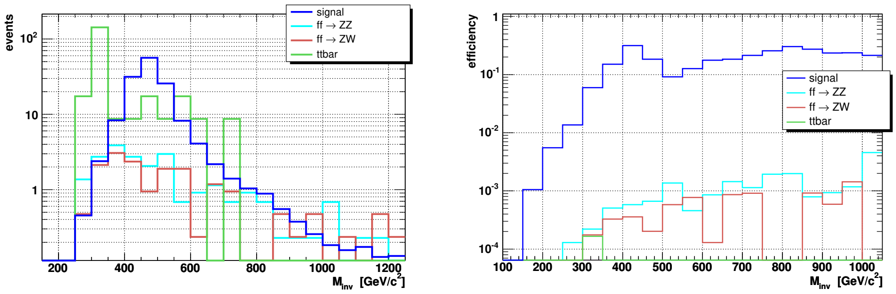


Fig. 3: (Left) Number of reconstructed events as a function of the *VV* invariant mass and (Right) the selection efficiency as a function of the invariant mass of the VV-fusion process; both for the $\mu\mu jjjj$ final state for $m_H = 500$ GeV and an integrated luminosity of 100 fb^{-1}.

1.4 Future Plans

Further studies are in progress, since for those presented here the Pythia generator was used, which only simulates a subset of the relavant diagrams, and cannot simulate the full set of background processes (notably not the scattering of transversely polarised vector bosons). To better describe the signal (and the background as well) a *Matrix Element Monte Carlo* must be used. *Phase* [5] is the best candidate, since it simulates all processes that lead to a six fermion final state, at order α_{QED}^6. Up to now only the channel $pp \rightarrow \mu\nu jjjj$ has been computed; therefore for the the $\mu\mu jjjj$ final state the *MadGraph* [6] event generator was used. This can simulate the $2l4j$ final state through the production (in Narrow Width Approximation) of intermediate vector bosons and their subsequent semileptonic decay.

Moreover it is crucial to redo the analysis, processing the events through the *Full Simulation* [7] of the CMS detector in order to properly take into account the detector resolution.

1.5 Summary

In conclusion, Electroweak Symmetry Breaking can be probed through the fusion of longitudinally polarized vector bosons with the CMS detector at LHC. The signal reconstruction and the background rejection algorithms have been successfully tested with the Fast Simulation. In the near future the study will be repeated with the Full Simulation of the detector and with dedicated generators.

References

[1] T. Sjöstrand et al., Comp. Phys. Comm. **135**, 238 (2001).

[2] S. Abdullin, A. Khanov and N. Stepanov, CMS TN/94-180.

[3] R. Bellan, *Observables in VV-fusion at the CMS Experiment*. Thesis, Università degli Studi di Torino, 2004.

[4] G. Cerminara, *A Study of the WW-fusion Process at CMS as a Probe of Symmetry Breaking*. Thesis, Università degli Studi di Torino, 2003.

[5] E. Accomando, A. Ballestrero and E. Maina, hep-ph/0504009.

[6] F. Maltoni and T. Stelzer, hep-ph/0208156.

[7] CMS Software and Computing Group, http://cmsdoc.cern.ch/oscar/;
CMS Software and Computing Group, http://cmsdoc.cern.ch/orca/.

5 Working group 5: Monte Carlo Tools

List of participants in the working group:

Victor Lendermann (Heidelberg), Sasha Nikitenko (Imperial College), Elzbieta Richter-Was (Jagiellonian University), Patrick Robbe (LAL Paris), Mike Seymour (Manchester/CERN)
Paolo Bartalini (CERN), Olaf Behnke (Heidelberg), Daniel Beneckestein (Heidelberg), Dimitri Bourilkov (Florida), Graziano Bruni (INFN Bologna), Andrew Buckley (Durham), Armen Bunyatyan (MPIK Heidelberg / Yerevan), Craig Buttar (Glasgow), Jonathan Butterworth (UCL), Susanna Butterworth (UCL), Sascha Caron (NIKHEF), Sergei Chekanov (Argonne), John Collins (Penn State), Massimo Corradi (INFN Bologna), Brian Cox (Manchester), Zofia Czyczula (IFJ-PAN Krakow/Copenhagen), Andrea Dainese (Padova), Giovanna Davatz (Zurich), Rick Field (Florida), Achim Geiser (DESY), Stefan Gieseke (Karlsruhe), Tanju Gleisberg (Dresden), Piotr Golonka (CERN), Günter Grindhammer (Munich), Robert Craig Group (Florida), Oliver Gutsche (DESY/Hamburg), Stefan Höche (Dresden), Giuseppe Iacobucci (INFN Bologna), Stanislaw Jadach (IFJ-PAN Krakow), Andreas Werner Jung (Heidelberg), Hannes Jung (DESY), Borut Kersevan (Ljubljana), Frank Krauss (Dresden), Stefan Lausberg (Heidelberg), Nils Lavesson (Lund), Kristin Lohwasser (Dortmund), Leif Lönnblad (Lund), Stephen Maxfield (Liverpool), Michelangelo Mangano (CERN), James Monk (Manchester), Arthur Moraes (Glasgow), Christian Pascaud (LAL Orsay), Klaus Rabbertz (Karslruhe), Peter Richardson (Durham), Lorenzo Rinaldi (INFN Bologna), Eduardo Rodrigues (NIKHEF), Marta Ruspa (University of Eastern Piedmont, Novara), Andreas Schälicke (Dresden), Thomas Schörner-Sadenius (Hamburg), Steffen Schumann (Dresden), Torbjorn Sjöstrand (Lund/CERN), Maciej Skrzypek (IFJ-PAN Krakow), Andrei Sobol (Nebraska), James Stirling (Durham), Paul Szczypka (Bristol), Paul Thompson (Birmingham), Niels Tuning (NIKHEF), Jacek Turnau (IFJ-PAN Krakow), Bennie Ward (Baylor, Texas), Zbigniew Was (IFJ-PAN Krakow), Ben Waugh (UCL), Mike Whalley (Durham), Jan Winter (Dresden)

Introduction to Monte Carlo Tools

V. Lendermann [1], *A. Nikitenko* [2], *E. Richter-Was* [3], *P. Robbe* [4], *M. H. Seymour* [5]
[1] Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany;
[2] Imperial College, London, UK; on leave from ITEP, Moscow, Russia;
[3] Institute of Physics, JU/IFJ-PAN, 30-059 Krakow, ul. Reymonta 4, Poland;
[4] Laboratoire de l'Accélérateur Linéaire, Université Paris-Sud, 91898 Orsay, France;
[5] School of Physics & Astronomy, University of Manchester, UK, and Theoretical Physics Group, CERN.

Abstract
The activities of working group 5 "Monte Carlo Tools" of the HERA–LHC Workshop are summarised. The group concerned itself with the developments and tunings of Monte Carlo models in light of the HERA–LHC connection, interfaces and libraries for parton density functions, Monte Carlo running, validation and tuning frameworks as well as some data analysis tools.

1 Introduction

The goals of working group 5 were:

- to review existing and developing Monte Carlo (MC) models used for studies at HERA and the LHC;
- to examine and possibly improve Monte Carlo models for the LHC physics using HERA data;
- to prioritise possible measurements at HERA which would allow tuning of these MC models;
- to pursue the development of frameworks for running, validating and tuning of Monte Carlo and analysis programs;
- to improve and further develop common interfaces and libraries used with Monte Carlo event generators;
- to review data analysis tools developed by the HERA collaborations which can be useful for studies at the LHC.

Both theorists and experimentalists from the HERA and LHC communities came together, in order to share their experience, identify crucial issues and discuss the future developments of the programs, libraries and frameworks.

The physics topics discussed in the group have overlapped with those of all the other working groups in this workshop. Therefore many presentations were given in common sessions with other groups, most notably with working group 2 "Multijet Final States and Energy Flows". These presentations covered the models of multiparton interactions, new developments in parton shower models, matrix element / parton shower (ME+PS) matching and simulations of multijet final states. The contributions to the present proceedings reviewing these studies are published in the chapter of working group 2. Further contributions are summarised below.

2 Libraries for Parton Density Functions

In the past, the PDFLIB library [1] was the standard package containing parameterisations of the proton, photon and pion parton density functions (PDFs). The LHC studies have necessitated the development of a new library which should include not only the central values of PDFs but also the error sets. The PDFLIB interface appeared not well suited to meet the new requirements. Therefore a new library, LHAPDF (Les Houches Accord PDF library [2]) was created following the Les Houches meeting in 2001. During this workshop the library was extensively developed by *M. Whalley* and *D. Bourilkov* [3]. Several recent PDF sets were included, both those by the leading theory groups and by the H1 and ZEUS collaborations. The fits are important for better estimations of PDF uncertainties. A particularly interesting cross check would be, e.g., the comparison of the fits provided by the TeVatron collaborations with the independent predictions obtained by the DGLAP evolution of the HERA PDFs to the TeVatron region.

The photon and pion PDFs were also included into the library, thus allowing its usage for all HEP analyses, in particular, for HERA studies. The library is thus developed to the level at which it can fully replace PDFLIB. Several tests of the applicability and performance of the library were made within the H1 collaboration (*V. Lendermann*).

Another topic discussed is the creation of a collection of diffractive PDF parameterisations. The project was presented by *F. P. Schilling* in a meeting of working group 4 "Diffraction".

3 Monte Carlo Event Generators

The status and plans for major leading order (LO) and next-to-leading order (NLO) QCD programs, as well as generators with k_T factorisation was discussed.

3.1 Leading Order MC programs

Currently, the major fortran MC event generators, PYTHIA [4] and HERWIG [5], are undergoing the transition to object-oriented software technologies. The C++ versions of both generators, PYTHIA7 [6] and HERWIG++ [7], are built in the common framework THEPEG [8] which is based on the CLHEP class library [9].

PYTHIA7/THEPEG includes some basic $2 \to 2$ matrix elements (ME), several built-in PDF parameterisations, remnant handling, initial and final state parton showers, Lund string fragmentation and particle decays. There have been plans to rework the fragmentation model, in order to include junction strings, and to implement multiple interactions (*L. Lönnblad*). However, recently *T. Sjöstrand* started a completely new C++ implementation, PYTHIA8 [10].

In parallel to the work on the C++ PYTHIA versions, the development of the fortran PYTHIA6 continues. It remains the main platform for new physical concepts. The version 6.3 [11, 12] includes a completely new framework for simulation of parton showers and multiple interactions by *T. Sjöstrand* and *P. Skands*. Currently this version works for pp interactions only.

The development of HERWIG continues mainly in C++ (*S. Gieseke*, *A. Ribon*, *P. Richardson*, *M. H. Seymour*, *P. Stephens* and *B.R. Webber*). The current fortran version 6.5 is foreseen as the final fortran version of HERWIG (apart from possible bug fixes). It is interfaced to the JIMMY generator for multiple interactions (*J. M. Butterworth*, *J. R. Forshaw*, *M. H. Seymour*, and *R. Walker*).

HERWIG++ includes a new parton shower algorithm and an improved model of cluster fragmentation. The e^+e^- event generation is implemented in HERWIG++ 1.0. The next version including hadronic interactions is in progress. The plans for the near future are to fully implement the matrix element – parton shower matching according to the Catani–Krauss–Kuhn–Webber (CKKW) scheme [13], as well as multiple interactions. A new framework for accessing particle data and simulations of particle decays is currently being constructed by *P. Richardson*. The treatment of hadronic decays will include spin correlations.

Further physics models can be incorporated into the same THEPEG framework. In particular, it is planned to make a C++ version of ARIADNE [14] based on THEPEG (*L. Lönnblad*). ARIADNE implements the Dipole Cascade Model (DCM) [15, 16] as an alternative to the DGLAP based shower models used in PYTHIA or HERWIG.

Despite a great success of ARIADNE in modeling hadronic final state observables, as measured at LEP and HERA, additional work is necessary to make ARIADNE fully suitable for modeling interactions at the LHC. This was, in particular, shown in a study presented by *Z. Czyczula* on the impact of parton shower models on generation of bbH, $H \to \tau\tau$ at the LHC [17]. The planned features include a remodeling of initial state $g \to q\bar{q}$ splittings as well as the introduction of the $q \to g^{\star}q$ process.

Another study was presented by *N. Lavesson* of ME+PS matching in ARIADNE on the example of W+jet production at the TeVatron [18]. It is also planned to include the matching to fixed order tree-level matrix elements à la CKKW [13, 18, 19] for the most common subprocesses at the LHC. When these plans are realised (hopefully during 2006), it should be safe to use ARIADNE for LHC predictions.

An alternative C++ implementation is performed in the SHERPA program (*T. Gleisberg*, *S. Höche*, *F. Krauss*, *A. Schälicke*, *S. Schumann*, *J. Winter*) [20] which is capable of simulating lepton–lepton, lepton–photon, photon–photon and fully hadronic collisions, such as proton–proton reactions. In its current version SHERPA includes the ME generator AMEGIC++ providing the matrix elements for hard processes and decays in the SM, MSSM and the ADD model, the parton shower module APACIC++ containing virtuality ordered initial and final state parton showers, ME+PS matching using the CKKW algorithm, the AMISIC++ module for a simple hard underlying event model taken from PYTHIA and an interface to the PYTHIA string fragmentation and hadron decays. Studies were presented on ME+PS matching considering W/Z+jet production at the TeVatron and at the LHC (*S. Schumann*) [18], and on the underlying event simulations (*S. Höche*) [12].

None of the above C++ programs is available for ep interactions yet, and so no applications and tests of these programs at HERA have been possible.

Further talks were given on the RAPGAP event generator [21] by *H. Jung* and on the ACERMC event generator [22] by *B. Kersevan* and *E. Richter-Was*. RAPGAP is one of the major generators used at HERA. It includes leading order QCD matrix elements, LEPTO [23] and ARIADNE parton cascade models, as well as simulations of hard diffraction. Both ep and pp versions are available. Recently, the Les Houches Accord interface for fragmentation models was included which allows the choice between the PYTHIA and HERWIG fragmentation models. This feature may allow better estimations of measurement uncertainties accounting for the transition from parton to hadron level of final states. It is planned to include double diffractive scattering for pp collisions to allow simulation of diffractive Higgs production.

The ACERMC event generator simulates the Standard Model backgrounds to the Higgs production in pp collisions. It includes LO QCD matrix elements produced by MADGRAPH/HELAS [24], as well as both PYTHIA and HERWIG parton shower and fragmentation models via the Les Houches Accord interface. During this workshop, the ARIADNE parton shower model and the LHAPDF library were implemented. The program is also interfaced to TAUOLA [25] for precise treatment of τ decays and to PHOTOS [26, 25] for simulations of QED radiative decays. The study on the impact of parton shower models on generation of bbH, $H \to \tau\tau$ at the LHC [17], mentioned above, was performed using ACERMC.

The program can be linked with the ACERDET package [27] which provides a fast and simplified simulation of the expected ATLAS detector effects (energy smearing, acceptance corrections) as well as the usual analysis steps (jet reconstruction algorithms, isolation criteria, etc.). This allows a quick estimation of the feasibility of measurements in an LHC experiment, not necessarily by the members of the experimental collaboration.

3.2 NLO MC Programs

NLO QCD calculations are required to make theoretical predictions at the level of precision currently reached in particle scattering experiments. However, writing a hadron level MC program implementing an NLO model is a very complicated task, which is currently solved only for a few pp reactions [28].

An important step forward would be an *ep* version of MC@NLO. It would be a major benefit for HERA studies of heavy quark and multijet production and would also allow an extensive validation of the NLO QCD calculations with HERA data. The development of the program started recently [29].

3.3 Monte Carlo Programs with k_T Factorisation

The CASCADE event generator presented by *H. Jung* [30] provides an implementation of the CCFM model for parton cascades [31]. The program was very successful in describing hadronic final states at low x at HERA. First applications of CASCADE for the studies at the LHC were presented by *G. Davatz* [17]. The plans include an implementation of quark lines into CCFM cascades (currently, only gluon lines are implemented), as well as a new model for multiparton interactions based on the AGK cutting rules [32].

A reformulation of the CCFM model into the link dipole chain (LDC) model [33] provides a simplified formalism, which is incorporated into the LDCMC program by *H. Kharraziha* and *L. Lönnblad* [34, 14]. An LDCMC version for deep inelastic *ep* scattering is available within the framework of the ARIADNE event generator. A *pp* version is planned.

In conjunction with these models, special sessions of working group 2 were dedicated to possibilities of determining the unintegrated parton distributions in the proton [35].

4 Comparisons of MC Models with Data

Models for particular subprocesses and their tuning using HERA data were reviewed in the corresponding working groups. As mentioned above, many discussions were carried out in the common sessions of working group 5 with the other groups.

One topic considered in WG5 is a comparison of leading proton data with several MC models (*G. Bruni*, *G. Iacobucci*, *L. Rinaldi*, *M. Ruspa*) [36]. The *ep* data from ZEUS and *pp* data from ISR and fixed target experiments were confronted to the HERWIG (together with POMWIG [37] and SANG to simulate diffraction), LEPTO, ARIADNE and PYTHIA simulations. This exercise has revealed that the simulation of the leading-proton momenta, both longitudinal and transverse to the beams, does not reproduce the properties of the data.

This study can be especially important for understanding of diffractive processes and backgrounds for them at the LHC.

5 MC Running, Tuning and Validation Frameworks

During this workshop much progress was made in developing common frameworks that provide a convenient handling of MC and analytical programs and allow quick comparisons of MC simulations and analytical calculations with the results of HERA and other HEP experiments. The developments of HZTOOL/JETWEB, RUNMC and NLOLIB packages were presented and actively discussed.

The HZTOOL [38, 39] library provides a comprehensive collection of fortran routines to produce various distributions using Monte Carlo event generators. The routines allow easy reproduction of the experimental distributions by modeling programs and give access to published data from EMC, SPS, LEP, HERA and TeVatron experiments. A number of studies for the LHC and the future linear collider are also included. The library can be linked with all major fortran MC event generators, and with a number of NLO QCD programs from the NLOLIB package (see below). The development of the library started within the workshop "Future Physics at HERA" and steadily continued in the last years.

In the current workshop, the emphasis was put on the HERA results relevant for the LHC [38]. Several measurements by H1 and ZEUS were implemented which allow tuning of multiparton interaction models in MC event generators (work by *D. Beneckenstein, A. Bunyatyan, J. M. Butterworth, H. Jung, S. Lausberg, K. Lohwasser, V. Lendermann, B. M. Waugh*). Common tunings of multiple interaction models based on the TeVatron and HERA results may constitute in future an interesting outcome of the current efforts.

Recent H1 and ZEUS results on heavy quark production in ep collisions were also added (*A.W. Jung, A. Geiser, O. Gutsche, P. Thompson*). In addition, calculations of benchmark cross sections for heavy flavour production were included [40].

Based on HZTOOL, JETWEB [38, 41] is a facility for tuning and validating Monte Carlo models through a World Wide Web interface. A relational database of reaction data and predictions from different models is accessed through a Java servlet, enabling a user to find out how well a given model agrees with a range of data and to request predictions using parameter sets that are not yet in the database.

The transition of experimental analysis frameworks and Monte Carlo generators to object-oriented software technologies necessitates a proper development of the MC running, tuning and validation frameworks. For this reason, HZTOOL/JETWEB is currently subject to a major redesign within the CEDAR ("Combined e-Science Data Analysis Resource for high energy physics") project [42]. CEDAR should comprise:

- an extensive archive of data from particle scattering experiments, based on the Durham HEP database [43];
- validation and tuning of Monte Carlo programs, parton distribution functions and other high energy physics calculation programs, building on JETWEB;
- access to well defined versions of these programs and code management support for developers;
- a standardised set of data formats for specifying HEP measurements as used in HepData and Monte Carlo event generator configurations as used in JETWEB;
- Grid compatibility for distribution of JETWEB Monte Carlo submissions and to enable secure addition of experimental data to the HepData catalogue by experimental collaborations.

A particularly important step in building CEDAR will be designing a C++ equivalent for HZTOOL, as well as providing an interface to the new C++ MC event generators.

A complementary approach using the object-oriented software design is realised in the RUNMC framework [44] by *S. Chekanov*. While JETWEB is a Web server system, RUNMC is a desktop application written in C++ and java. It provides a unified approach to generate MC events and to analyse different MC models. All major fortran MC event generators can be run via RUNMC. The output of fortran MC programs is converted to C++ classes for further analysis and for graphical representation (histograms). The graphical user interface of RUNMC allows an initialisation of MC models and histograms in a unified manner, and provides monitoring of the event generation. The program provides an interface to HZTOOL. It also allows loading of "project files" which can contain external calculations, MC tunings, histogram definitions, etc. In particular, these files can include C++ data analysis code, similar to the HZTOOL fortran analysis routines.

A further project, discussed in working group 5, is a common framework for the NLO QCD programs, NLOLIB [45], which was initiated within the workshop "Monte Carlo Generators for HERA Physics". Since hadron level Monte Carlo programs implementing QCD NLO calculations are not (yet) available for many processes, parton level NLO calculations are extensively used. NLOLIB is aimed to become a container for virtually all NLO QCD programs. It provides:

- a setup for compiling and linking the programs on diverse UNIX platforms;
- a unified access to the NLO event records;
- a unified steering for parameters and settings;
- a unified access to PDF libraries;
- an interface to HZTOOL, thus allowing easy comparisons with experimental results;
- examples of the analysis code which can be linked with the library.

During the workshop the structure of the framework was further developed by *K. Rabbertz*. In addition to already implemented programs for ep (DISENT [46], DISASTER++ [47], MEPJET [48]) and e^+e^-

(RACOONWW [49]) physics, an effort was made to integrate further ep programs: NLOJET++ [50] (*K. Rabbertz*) and JETVIP [51, 52] (*T. Schörner-Sadenius*). The integration of the NLO programs for pp physics is surely possible, but requires additional effort.

6 Data Analysis Tools

Apart from MC related topics, general analysis tools, aimed for searches for specific final states, were presented.

One such tool is SBUMPS [53], currently being developed by *S. Chekanov*, which performs automatic searches for resonance peaks in invariant mass distributions of two or more tracks. The program can be useful for searches of new states as well as for the reconstruction of known resonances.

Recently, interest in hadron spectroscopy at HERA increased sharply with the observations of narrow peaks in inclusive invariant mass distributions which can be interpreted as pentaquarks [54]. These studies have inspired the development of the automated peak searching tool, which can be used in data analyses at any particle scattering experiment.

A general strategy for searches for new physics was presented by *S. Caron*. The approach was developed and used by the H1 collaboration for searches of new phenomena at HERA [55]. It involves a statistical algorithm to search for deviations from the Standard Model in the distributions of the scalar sum of transverse momenta or invariant mass of final state particles and to quantify their significance.

7 Conclusions

A number of interesting developments of MC models, programs, libraries and frameworks were presented in working group 5. The general status and prospects for major established and currently developed MC generators were reviewed. In common sessions with working group 2, the models of multiparton interactions, new developments in parton shower models, matrix element / parton shower matching and simulations of multijet final states were extensively discussed. Direct communication of theoreticians and experimentalists from the HERA and LHC communities further pursued several developments and studies. In particular the recent advances in the development of the LHAPDF library, HZTOOL, RUNMC and NLOLIB frameworks were inspired by discussions within working group 5. This will hopefully help further studies on validation and tuning of the MC models for multiparton interactions, parton showers and heavy flavour production.

References

[1] H. Plothow-Besch, Comput. Phys. Commun. **75**, 396 (1993).

[2] W. Giele *et al.*, Workshop on Physics at TeV Colliders, Les Houches, France, May 2001.

[3] M. Whalley, D. Bourilkov and R. C. Group, these proceedings, working group 5;
LHAPDF Web page, available on `http://hepforge.cedar.ac.uk/lhapdf/`.

[4] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun. **135**, 238 (2001);
PYTHIA Web page, available on
`http://www.thep.lu.se/~torbjorn/Pythia.html`.

[5] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101**, 010 (2001);
HERWIG Web page, available on
`http://hepwww.rl.ac.uk/theory/seymour/herwig/`.

[6] PYTHIA7 Web page, available on `http://www.thep.lu.se/Pythia7/`.

[7] S. Gieseke, these proceedings, working group 5;
HERWIG++ Web page, available on
`http://www.hep.phy.cam.ac.uk/theory/Herwig++/`.

[8] L. Lönnblad, these proceedings, working group 5;
ThePEG Web page, available on `http://www.thep.lu.se/ThePEG/`.

[9] CLHEP Web page, available on
`http://wwwasd.web.cern.ch/wwwasd/lhc++/clhep/`.

[10] T. Sjöstrand, these proceedings, working group 5.

[11] T. Sjöstrand, L. Lönnblad, S. Mrenna and P. Skands, *Pythia 6.3: Physics and manual*. Preprint hep–ph/0308153, 2003.

[12] C. Buttar *et al.*, these proceedings, working group 2.

[13] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, JHEP **11**, 063 (2001);
F. Krauss, JHEP **08**, 015 (2002).

[14] L. Lönnblad, these proceedings, working group 5;
L. Lönnblad, Comput. Phys. Commun. **71**, 15 (1992).

[15] G. Gustafson, Phys. Lett. B **175**, 453 (1986).

[16] G. Gustafson and U. Pettersson, Nucl. Phys. B **306**, 746 (1988).

[17] Z. Czyczula, G. Davatz, A. Nikitenko, E. Richter-Was, E. Rodrigues and N. Tuning, these proceedings, working group 2.

[18] S. Höche *et al.*, these proceedings, working group 2.

[19] L. Lönnblad, JHEP **05**, 046 (2002);
N. Lavesson and L. Lönnblad, JHEP **07**, 054 (2005).

[20] T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann and J. Winter, JHEP **402**, 056 (2004);
T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann and J. Winter, these proceedings, working group 5;
SHERPA Web page, available on
`http://www.physik.tu-dresden.de/~krauss/hep/`.

[21] H. Jung, these proceedings, working group 5;
H. Jung, Comput. Phys. Commun. **86**, 147 (1995);
RapGap Web page, available on `http://www.desy.de/~jung/rapgap/`.

[22] B. Kersevan and E. Richter-Was, these proceedings, working group 5;
AcerMC Web page, available on `http://borut.home.cern.ch/borut/`.

[23] G. Ingelman, A. Edin and J. Rathsman, Comput. Phys. Commun. **101**, 108 (1997);
Lepto Web page, available on `http://www3.tsl.uu.se/thep/lepto/`.

[24] F. Maltoni and T. Stelzer, JHEP **02**, 027 (2003);
MadGraph Web page, available on `http://madgraph.hep.uiuc.edu/`.

[25] Tauola and Photos Web page, available on
`http://wasm.home.cern.ch/wasm/goodies.html`.

[26] P. Golonka and Z. Was, these proceedings, working group 2.

[27] AcerDET Web page, available on
`http://erichter.home.cern.ch/erichter/AcerDET.html`.

[28] S. Frixione and B. R. Webber, JHEP **0206**, 029 (2002);
S. Frixione, P. Nason and B. R. Webber, JHEP **0308**, 007 (2003);
MC@NLO Web page, available on
`http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO/`.

[29] S. Frixione, H. Jung and T. Toll, private communication.

[30] H. Jung and G. P. Salam, Eur. Phys. J. C **19**, 351 (2001);
H. Jung, these proceedings, working group 5;
CASCADE Web page, available on `http://www.desy.de/~jung/cascade/`.

[31] M. Ciafaloni, Nucl. Phys. B **296**, 49 (1988);
S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. B **234**, 339 (1990);
S. Catani, F. Fiorani and G. Marchesini, Nucl. Phys. B **336**, 18 (1990);
G. Marchesini, Nucl. Phys. B **445**, 49 (1995).

[32] J. Bartels, M. Salvadore and G. P. Vacca, Eur. Phys. J. C **42**, 53 (2005). And references therein.

[33] Bo Andersson, G. Gustafson and J. Samuelsson, Nucl. Phys. B **463**, 217 (1996).

[34] H. Kharraziha and L. Lönnblad, Comput. Phys. Commun. **123**, 153 (1999).

[35] J. Collins, M. Diehl, H. Jung, L. Lönnblad, M. Lublinksy and T. Teubner, these proceedings, working group 2.

[36] G. Bruni, G. Iacobucci, L. Rinaldi and M. Ruspa, these proceedings, working group 5.

[37] B. E. Cox and J. R. Forshaw, Comput. Phys. Commun. **144**, 104 (2002);
POMWIG Web page, available on `http://www.pomwig.com/`.

[38] J. Butterworth, H. Jung, V. Lendermann and B. Waugh, these proceedings, working group 5.

[39] HZTool Web page, available on `http://www.cedar.ac.uk/heptools/hztool/`.

[40] O. Behnke, M. Cacciari, M. Corradi, A. Dainese, H. Jung, E. Laenen, I. Schienbein and H. Spiesberger, these proceedings, working group 3.

[41] JetWeb Web page, available on `http://www.cedar.ac.uk/jetweb/`.

[42] A. Buckley, J. M. Butterworth, S. Butterworth, L. Lönnblad, W. J. Stirling, M. Whalley and B. M. Waugh, these proceedings, working group 5;
CEDAR Web page, available on `http://www.cedar.ac.uk/`.

[43] Durham HEP database Web page, available on `http://durpdg.dur.ac.uk/hepdata/`.

[44] S. Chekanov, these proceedings, working group 5;
RunMC Web page, available on `http://hepforge.cedar.ac.uk/runmc/`.

[45] K. Rabbertz and T. Schörner-Sadenius, these proceedings, working group 5;
NLOLIB Web page, available on `http://www.desy.de/~nlolib/`.

[46] S. Catani and M. H. Seymour, Nucl. Phys. B **485**, 291 (1997);
M. H. Seymour, DISENT Web page, available on
`http://hepwww.rl.ac.uk/theory/seymour/nlo/`.

[47] D. Graudenz. Preprint hep–ph/9710244, 1997;
DISASTER Web page, available on
`http://graudenz.home.cern.ch/graudenz/disaster.html`.

[48] E. Mirkes and D. Zeppenfeld, Phys. Lett. B **380**, 205 (1996);
E. Mirkes and D. Zeppenfeld, Phys. Rev. Lett. **78**, 428 (1997).

[49] A. Denner, S. Dittmaier, M. Roth and D. Wackeroth, Phys. Lett. B **475**, 127 (2000);
RacoonWW Web page, available on
`http://ltpth.web.psi.ch/racoonww/racoonww.html`.

[50] Z. Nagy, Phys. Rev. Lett. **88**, 122003 (2002);
NLOJET++ Web page, available on `http://www.cpt.dur.ac.uk/~nagyz/nlo++/`.

[51] B. Pötter, Comput. Phys. Commun. **133**, 105 (2000).

[52] B. Pötter, Comput. Phys. Commun. **119**, 45 (1999);
B. Pötter, Eur. Phys. J. Direct C **5**, 1 (1999);
G. Kramer and B. Pötter, Eur. Phys. J. C **5**, 665 (1998);
M. Klasen, G. Kramer and B. Pötter, Eur. Phys. J. C **1**, 261 (1998);
B. Pötter, Nucl. Phys. B **559**, 323 (1999);
B. Pötter, Nucl. Phys. B **540**, 382 (1999);
JetVip Web page, available on `http://www.desy.de/~poetter/jetvip.html`.

[53] S. Chekanov, these proceedings, working group 5;
SBumps Web page, available on `http://www.desy.de/~chekanov/sbumps/`.

[54] A. Airapetian *et al.* [HERMES Collaboration], Phys. Lett. B **585**, 213 (2004);
S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **591**, 7 (2004);
A. Aktas *et al.* [H1 Collaboration], Phys. Lett. B **588**, 17 (2004);
S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **38**, 29 (2004);
S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **610**, 212 (2005).

[55] A. Aktas *et al.* [H1 Collaboration], Phys. Lett. B **602**, 14 (2004).

The Les Houches Accord PDFs (LHAPDF) and LHAGLUE

M R Whalley[†], *D Bourilkov*[‡] *and R C Group*[‡]
[†]University of Durham, Durham, DH1 3LE, UK
[‡]University of Florida, Gainesville, FL 32611, USA

Abstract

We describe the development of the LHAPDF library from its initial implementation following the Les Houches meeting in 2001 to its present state as a functional replacement for PDFLIB. Brief details are given of how to install and use the library together with the PDF sets available. We also describe LHAGLUE, an add-on PDFLIB look-a-like interface to LHAPDF, which facilitates using LHAPDF with existing Monte Carlo generators such as PYTHIA and HERWIG.

1 LHAPDF – Introduction

Parton Density Functions (PDFs), which describe the partonic content of hadrons, need to be well understood and of sufficiently high precision if theoretical predictions are to match the experimental accuracies expected from future LHC data. These PDFs, which are produced by several different groups (e.g. MRST, CTEQ, Alekhin and more recently ZEUS and H1), are derived from fitting deep inelastic and related hard scattering data using parameterisations at low Q_0^2 ($\approx$ 1–7 (GeV/c)^2) and evolving these to higher Q^2. These PDFs are typically presented as grids in x-Q^2 with suitable interpolation codes provided by the PDF authors. The CERN PDFLIB library [1] has to date provided a widely used standard FORTRAN interface to these PDFs with the interpolation grids built into the PDFLIB code itself. However, it is realised that PDFLIB would be increasingly unable to meet the needs of the new generation of PDFs which often involve large numbers of sets ($\approx$20–40) describing the uncertainties on the individual partons from variations in the fitted parameters. As a consequence of this, at the Les Houches meeting in 2001 [2], the beginnings of a new interface were conceived — the so-call "Les Houches Accord PDF"— LHAPDF. This has further been developed over the course of the HERA-LHC workshop incorporating many new features to enable it to replace PDFLIB as the standard tool to use. The development is briefly described in this writeup together with LHAGLUE, an interface to LHAPDF, which provides PDF access using almost identical calling routines as PDFLIB.

2 LHAPDF – Development during the Workshop

In its initial incarnation (Version 1), LHAPDF had two important features which distinguished it from the methods used by PDFLIB in handling PDFs.

Firstly the PDFs are defined by the analytical formulae used in the original fitting procedures, with external files of parameters, which describe the momentum x distributions of the partons at the relevant Q_0^2. Evolution codes within LHAPDF then produce the PDF at any desired Q^2 at the users request. At present LHAPDF provides access to two evolution codes, EVLCTEQ for the CTEQ distributions and QCDNUM 16.12 [3] for the other PDF sets. This represents a radical difference from the existing methods used by the PDF authors to present their distributions where large grid files and interpolation routines are the norm. In PDFLIB these interpolation codes and grids are essentially compiled into a single FORTRAN library. The advantage of the LHAPDF method is that the compiled code is separate from the parameter files, which are typically small. Thus to add new PDF sets does not necessarily need the code to be recompiled and the library rebuilt.

Secondly, the concept is introduced of a "set" being a related collection of PDFs (e.g. an error set) all of which are accessible to the program after initialisation of that set. This allows LHAPDF to handle the multi-set "error" PDFs produced in recent years which give predicted uncertainties to the PDF values. All the PDFs in a set are initialised together and are therefore available to the user.

V1 was written by Walter Giele of Fermilab who in 2002 released a working version which could be downloaded from a web-site together with the parameter files for a limited number of PDF sets. There was also a manual and example files. One of the present authors MRW became involved and took over maintenance and development of LHAPDF in March 2003. The limitations of the idealised situation in V1 with respect to making LHAPDF a replacement tool for PDFLIB soon became apparent.

The primary problem was that V1 contained only a limited number of PDF sets and, since the method was reliant on the x parameterisations at Q_0^2 being available, it would be virtually impossible to include many of the older sets which are still needed for comparisons. A second and serious problem is the compute time taken in the initialisation phase of the individual members of a PDF set (i.e. calling the routine InitPDF described later). This can take in the region of 2 seconds per call on a 1GHz machine and is therefore unacceptable in the situation of a program which makes repeated use of the different members. [1]

A solution introduced in LHAPDF Version 2, which helps to solve the above problems, was to include the option to make the original grid files and interpolation codes available in LHAPDF **in addition** to the V1 method of parameter files and "on-the-fly" evolution. For some PDF sets both methods would be available and for others only the latter. The operation of the program was made identical for both methods with the content of the input file (with extension ".LHpdf" for the former and ".LHgrid" for the latter) dictating which is used. Not only does this allow all the older PDF sets to be included but also there is no time penalty in changing between members of the same set since all are loaded in the initialisation phase. LHAPDF V2 was released in March 2004 including many of the older PDF sets as well as some new ones.

LHAPDF Version 3 was released in September 2004 and, as well incorporating more older and some new PDF sets (e.g. ZEUS and H1), it also included the code for LHAGLUE, a newly developing add-on interface to LHAPDF which provides PDFLIB look-alike access. In addition to having subroutine calls identical to those in PDFLIB it also incorporates a PDF numbering scheme to simplify usage. It should be noted however that, because of the greatly increased number of new PDF sets, it was not possible to follow the original numbering scheme of PDFLIB and a new one was devised. This is described in more detail in Section 5.

The major feature of Version 4, which was released in March 2005, was the incorporation of the photon and pion PDFs. All the photon and pion PDFs that were implemented in PDFLIB were put into LHAPDF using identical code and using the ".LHgrid" method. The LHAGLUE numbering scheme in these cases more closely resembles that of PDFLIB than it does for the protons.

In addition in V4 there were new proton PDFs (MRST2004 and an updated Alekhin's a02m), a new simpler file structure with all the source files being in a single "src" directory, some code changes to incorporate access to $\Lambda_{4/5}^{QCD}$ and a more rigorous implementation of the α_s evolution as being exactly that used by the PDF author.

All the LHAPDF and LHAGLUE data and code, in addition to being made available on the new web site (http://hepforge.cedar.ac.uk/lhapdf/), is also included in the GENSER subproject of the LHC Computing Grid.

3 LHAPDF – Development after the Workshop

Since the last HERA-LHC meeting there has been one minor release of LHAPDF (Version 4.1 in August 2005). In this version the installation method has changed to be more standard with the "configure; make;

[1] A third problem reported at the workshop concerning small differences (up to $\approx 0.5\%$) between the PDFs produced by LHAPDF for MRST and the authors' code directly is now believed to be due to slight mismatches of grid boundaries at the heavy quark thresholds and will be corrected in future MRST grids.

make install" sequence familiar to many and also a small amount of code has been altered to be more compliant with proprietary FORTRAN 95 compilers. As mentioned in the previous section the web site for public access to LHAPDF from which the source code can be obtained has changed. Since this is the current and most recent version we assume V4.1 in the following referring to earlier versions where necessary.

4 Using LHAPDF

Once the code and PDF data sets have been downloaded from the relevant web site and installed following the instructions given therein, using LHAPDF is simply a matter of linking the compiled FORTRAN library **libLHAPDF.a** to the users program. Table 1 lists the LHAPDF routines available to the user, which are of three types:

- Initialisation (selecting the required PDF set and its member)
- Evolution (producing the momentum density functions (*f*) for the partons at selected *x* and *Q*)
- Information (displaying for example α_s, descriptions, etc.)

Command	Description
call InitPDFset(*name* **)**	Initialises the PDF set to use.
call InitPDF(*member* **)**	Selects the member from the above PDF set.
call evolvePDF(*x,Q,f* **)**	Returns the momentum density function, $f(x, Q)$, for protons or pions.
call evolvePDFp(*x,Q,P2,ip2,f* **)**	Returns the momentum density function for photons. [2].
call numberPDF(*num* **)**	Returns the number (*num*) of PDF members in the set.
call GetDesc( )	Prints a description of the PDF set.
alphasPDF(*Q* **)**	Function giving the value of α_s at *Q* GeV.
call GetLam4(*mem,qcdl4* **)**	Returns the value of Λ_4^{QCD} for the specific member.
call GetLam5(*mem,qcdl5* **)**	Returns the value of Λ_5^{QCD} for the specific member.
call GetOrderPDF(*order* **)**	Returns the order of the PDF evolution.
call GetOrderAs(*order* **)**	Returns the order of the evolution of α_s .
call GetRenFac(*muf* **)**	Returns the renormalisation factor.
call GetQmass(*nf,mass* **)**	Returns the mass of the parton of flavour *nf*.
call GetThreshold(*nf,Q* **)**	Returns the threshold value for parton of flavour *nf*.
call GetNf(*nfmax* **)**	Returns the number of flavours.

Table 1: LHAPDF commands

The evolution commands utilise a double precision array *f(-6:6)* where the arguments range from -6 to +6 for the different (anti)partons as shown in Table 2 below.

parton	$\bar{t}$	$\bar{b}$	$\bar{c}$	$\bar{d}$	$\bar{u}$	$\bar{d}$	g	d	u	s	c	b	t
n	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6

Table 2: The flavour enumeration scheme used for *f(n)* in LHAPDF

Specifying the location of the PDF sets in the code should be especially mentioned at this point. The argument (*name*) in **InitPDFset** should specify the complete path (or at least to a symbolic link to this path). From version 4.1 onwards, however, a new routine **InitPDFsetByName** can be used in which only the name of the PDF set need by specified. This works in conjunction with the script "lhapdf-config"

[2]In **evolvePDFp** $P2$ is the vitruality of the photon in GeV ^{2}, which should by 0 for an on-shell photon, and *ip2* is the parameter to evaluate the off-shell anomalous component. See the PDFLIB manual [1] for details.

which is generated at the configure stage of the installation which provides the correct path to the PDF sets. The location of this script must therefore be in the users execution path. Tables 3 and 4 list the complete range of PDF set available. The equivalent numbers to use in LHAGLUE, as described in the next section, are also listed in these tables.

Ref	Prefix	Suffix (# of sets)	type	LHAGLUE numbers
		Proton PDFs		
[4]	alekhin_	100 (100), 1000 (1000)	p	**40100-200**, **41000-1999**
[5]	a02m_	lo (17), nlo (17), nnlo (17)))	g	40350-67, 40450-67, 40550-67
[6]	botje_	100 (100),1000 (1000)	p	**50100-200**, **51000-1999**
[7]	cteq	61 (41)	p,g	**10100-40**, 10150-90
[8]	cteq	6 (41)	p,g	**10000-40**, 10050-90
	cteq	6m, 6l, 6ll	p	**10000, 10041, 10042**
[9]	cteq	5m, 5m1, 5d, 5l	g	19050, 19051, 19060, 19070
[10]	cteq	4m, 4d, 4l	g	19150, 19160, 19170
[11]	fermi2002_	100 (100), 1000 (1000)	p	**30100-200**, **31000-2000**
[12]	GRV98	lo, nlo(2)	g	80060, 80050-1
[13]	H12000	msE (21), disE (21), loE (21)	g	70050-70, 70150-70, 70250-70
[14]	MRST2004	nlo	p,g	**20400**, 20450
	MRST2004	nnlo	g	20470
[15]	MRST2003	cnlo	p,g	**20300**, 20350
	MRST2003	cnnlo	g	20370
[16]	MRST2002	nlo (2)	p,g	**20200**, 20250
	MRST2002	nnlo	g	20270
	MRST2001	E (31)	p,g	**20100-130**, 20150-180
[17]	MRST2001	nlo(4)	p,g	**20000-4**, 20500-4
	MRST2001	lo, nnlo	g	20060, 20070
[18]	MRST98	(3)	p	**29000-3**
	MRST98	lo (5), nlo (5) dis (5), ht	g	29040-5, 29050-5,29060-5,29070-5
[19]	ZEUS2002_	TR (23), FF (23), ZM (23)	p	**60000-22**, **60100-22**, **60200-22**
[20]	ZEUS2005_	ZJ (23)	p	**60300-22**

Notes:

LHAPDF → PrefixSuffix.LHpdf (type p). Where both p and g are present (p,g) then
filename → PrefixSuffix.LHgrid (type g). the user has the choice of either.

LHAGLUE numbers in **bold** are the type p (.LHpdf) sets.

Table 3: The Proton PDF sets available in LHAPDF.

5 LHAGLUE

The LHAGLUE interface [21] to LHAPDF is designed along the lines of the existing interface from PYTHIA to PDFLIB. [3]. For both HERWIG and PYTHIA the existing 'hooks' for PDFLIB have been utilised for the LHAGLUE interface. This makes it possible to link it exactly like PDFLIB with no further changes to PYTHIA's or HERWIG's source code needing to be implemented.

The interface contains three subroutines (similar to PDFLIB) and can be used seamlessly by Monte Carlo generators interfaced to PDFLIB or in standalone mode. These are described in Table 5. In addition any of the LHAPDF routines, **except** the initialisation routines **InitPDFset** and **InitPDF**, de-

[3]DB would like to thank T. Sjöstrand and S. Mrenna for discussions on this topic.

Prefix	Suffix	LHAGLUE numbers	Prefix	Suffix	LHAGLUE numbers
Pion PDFs			**Photon PDFs**		
OWPI	(2)	211-12	DOG	0, 1	311, 312
SMRSPI	(3)	231-3	DGG	(4)	321-4
GRVPI	0, 1	251, 252	LACG	(4)	331-4
ABFKWPI	(3)	261-3	GSG	0 (2), 1	341-2, 343
All filenames are PrefixSuffix.LHgrid			GSG96	0, 1	344, 345
The nomenclature used here is essentially the same as in PDFLIB and the relevant publication references can be found in the PDFLIB manual [1].			GRVG	0 (2), 1 (2)	351-2, 353-4
			ACFGPG	(3)	361-3
			WHITG	(6)	381-6
			SASG	(8)	391-8

Table 4: The Pion and Photon PDF sets available in LHAPDF.

scribed in Table 1, can also be used, for example to return the value of the strong coupling constant α_s (**alphasPDF(**Q**)**), or to print the file description (**call GetDesc()**).

There are also several CONTROL switches specified through the 20 element character array LHAPARM and COMMON blocks which determine how the interface operates.

- Location of the LHAPDF library of PDFs (pathname):
  From version LHAPDF v4.1 onwards, and the LHAGLUE routines distributed with it, the location of the PDFsets data files is set automatically using the "lhapdf-config" script as described in the previous section, provided that the prescribed installation instructions have been used.
  For previous versions (4.0 and earlier) the common block **COMMON/LHAPDFC/LHAPATH** is used where **LHAPATH** is a **character*132** variable containing the full path to the PDF sets. The default path is subdir 'PDFsets' of the current directory.
- Statistics on under/over-flow requests for PDFs outside their validity ranges in x and Q^2.
  a) **LHAPARM(16) .EQ. 'NOSTAT'** → No statistics (faster)
  b) **LHAPARM(16) .NE. 'NOSTAT'** → Default: collect statistics
  c) **call PDFSTA** at the end to print out statistics.
- Option to use the values for α_s as computed by LHAPDF in the Monte Carlo generator as well in order to ensure uniform α_s values throughout a run
  a) **LHAPARM(17) .EQ. 'LHAPDF'** → Use α_s from LHAPDF
  b) **LHAPARM(17) .NE. 'LHAPDF'**→ Default (same as LHAPDF V1/V3)
- Extrapolation of PDFs outside the LHAPDF validity range given by $x_{min/max}$ and $Q^2_{min/max}$.
  a) Default → PDFs "frozen" at the boundaries.
  b) **LHAPARM(18) .EQ. 'EXTRAPOLATE'** → Extrapolate PDFs at own risk
- Printout of initialisation information in PDFSET (by default)
  a) **LHAPARM(19) .EQ. 'SILENT'** → No printout (silent mode).
  b) **LHAPARM(19) .EQ. 'LOWKEY'** → Print 5 times (almost silent).
- Double Precision values of $\Lambda^{QCD}_{4/5}$ applicable to the selected PDF are available (as read-only) in the COMMON block: **COMMON/W50512/QCDL4,QCDL5** → as in PDFLIB.

The LHAGLUE interface can be invoked in one of 3 ways, Standalone, PYTHIA or HERWIG, depending on the value of *parm(1)* when calling **PDFSET(***parm,value)*.

- **Standalone** mode
  PARM(1)= 'DEFAULT'
  VALUE(1) = "*PDF number*"
- **PYTHIA mode**
  PARM(1) = 'NPTYPE' ← set automatically in PYTHIA

Command	Description
call PDFSET(*parm,value)*	For initialisation (called once) where PARM and VALUE are LOCAL arrays in the calling program specified as **CHARACTER*20 PARM(20)** **DOUBLE PRECISION VALUE(20)**
call STRUCTM(*X,Q,UPV,DNV,USEA,DSEA,STR,CHM,BOT,TOP,GLU)*	
	For the proton (and pion) PDFs: where X and Q are the input kinematic variables and the rest are the output PDF of the valence and sea quarks and the gluon.
call STRUCTP(*X,Q2,P2,IP2,UPV,DNV,USEA,DSEA,STR,CHM,BOT,TOP,GLU)*	
	For the photon PDFs: as above with the additional input variables P2 and IP2 [2].

Table 5: LHAGLUE commands

In this case the user must supply MSTP(51) and MSTP(52) in the PYTHIA common block
COMMON/PYPARS/MSTP(200),PARP(200),MSTI(200).PARI(200)
MSTP(52) = 2 ← to use an external PDF library
MSTP(51)= "*PDF number*"

- **HERWIG mode**
  PARM(1) = 'HWLHAPDF'← set by the user.
  In this case one sets for the beam and target particles separately
  AUTPDF(1) = 'HWLHAPDF'
  AUTPDF(2) = 'HWLHAPDF'
  MODPDF(1) = "*PDF number*"
  MODPDF(2) = "*PDF number*"
  Note that HERWIG specifies the"*PDF number*" for each of the colliding particles separately and care should be taken that the same PDF members are used when appropriate.

The user then simply links their own standalone code, or the HERWIG/PYTHIA main program and the HERWIG/PYTHIA code [4], with the LHAPDF library **libLHAPDF.a** making sure the 'PDFsets' directory is specified as described above.

The LHAGLUE interface has been tested extensively at TEVATRON and LHC energies for the proton PDFs and with HERA examples for the photon PDFs. Results with new and legacy PDF sets, using LHAPDF, PDFLIB or internal implementations in the Monte Carlo generators, and comparing cross sections produced with PYTHIA and HERWIG, give us confidence in the consistency of the LHAGLUE interface and the underlying LHAPDF library [22].

6 Summary and Future Development

Both LHAPDF and the interface LHAGLUE have been developed over the period of the Workshop to a point where they can now be used as a serious replacement for PDFLIB. Indeed, except for the PDF authors' own code, they are the only place to obtain the latest PDFs. There is however still considerable development in progress and the latest PDF sets will be incorporated as and when they become available. One major development area is to include the possibility of having more than one PDF set initialised concurrently. This may be necessary in interactions between different beam and target particles types

[4]It is important when starting with a fresh PYTHIA or HERWIG download the user must first rename the 'dummy' subroutines **STRUCTM**, **STRUCTP** and **PDFSET** in the PYTHIA/ HERWIG source codes exactly as if one were to link to PDFLIB.

and also including photon and pion PDFs. This will be the aim of the next LHAPDF release.

Acknowledgements

MRW wishes to thank the UK PPARC for support from grant PP/B500590/1. DB wishes to thank the USA National Science Foundation for support under grants NSF ITR-0086044 and NSF PHY-0122557.

References

[1] H. Plothow-Besch, Comput. Phys. Commun. **75**, 396 (1993).

[2] W. Giele *et al.*, in *The QCD/SM working group: Summary report*. 2002. Also in preprint hep-ph/0204316.

[3] M. Botje, *QCDNUM version 16.12* (unpublished). Available on http://www.nikhef.nl/ h24/qcdcode/index.html.

[4] S. I. Alekhin, Phys. Rev. **D63**, 094022 (2001).

[5] S. I. Alekhin, Phys. Rev. **D68**, 014002 (2003).

[6] M. Botje, Eur. Phys. J. **C14**, 285 (2000).

[7] D. Stump *et al.*, JHEP **10**, 046 (2003).

[8] J. Pumplin *et al.*, JHEP **07**, 012 (2002).

[9] H. L. Lai *et al.*, Eur. Phys. J. **C12**, 375 (2000).

[10] H. L. Lai *et al.*, Phys. Rev. **D55**, 1280 (1997).

[11] W. T. Giele, S. A. Keller and D. A. Kosower, *Parton distribution function uncertainties*. Preprint hep-ph/0104052, 2001.

[12] M. Glück, E. Reya and A. Vogt, Eur. Phys. J. **C5**, 461 (1998).

[13] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. **C30**, 1 (2003).

[14] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C39**, 155 (2005).

[15] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, *Mrst partons and uncertainties*. Preprint hep-ph/0307262, 2003.

[16] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C28**, 455 (2003).

[17] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C23**, 73 (2002).

[18] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C4**, 463 (1998).

[19] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Rev. **D67**, 012007 (2003).

[20] S. Chekanov *et al.* [ZEUS Collaboration], *An NLO QCD analysis of inclusive cross-section and jet- production data from the ZEUS experiment - DESY-05-050.* (unpublished). Preprint hep-ph/0503274, 2005.

[21] D. Bourilkov, *Study of parton density function uncertainties with LHAPDF and PYTHIA at LHC* (unpublished). Preprint hep-ph/0305126, 2003.

[22] DB, CG, MRW. Note in preparation.

THEPEG
Toolkit for High Energy Physics Event Generation

Leif Lönnblad
Department of Theoretical Physics, Lund University, Sweden

Abstract
I present the status of the THEPEG project for creating a common platform for implementing C++ event generators. I also describe briefly the status of the new versions of PYTHIA, HERWIG and ARIADNE which are implemented using this framework.

1 Introduction

Monte Carlo Event Generators (EGs) have developed into essential tools in High Energy Physics. Without them it is questionable if it at all would be possible to embark on large scale experiments such as the LHC. Although the current EGs work satisfactorily, the next generation of experiments will substantially increase the demands both on the physics models implemented in the EGs and on the underlying software technology.

The current EGs are typically written in Fortran and their basic structure was designed almost two decades ago. Meanwhile there has been a change in programming paradigm, towards object oriented methodology in general and C++ in particular. This applies to almost all areas of high-energy physics, but in particular for the LHC program, where all detector simulation and analysis is based on C++. When designing the next generation of EGs it is therefore natural to use C++. Below is a brief description of the THEPEG [1] project for designing a general framework in C++ for implementing EG models, and also the PYTHIA7 and ARIADNE programs which uses THEPEG to implement their respective physics models. Also HERWIG++ is implemented in the THEPEG framework, but this program is described elsewhere in these proceedings [2]

2 Basic structure

THEPEG is a general platform written in C++ for implementing models for event generation. It is made up from the basic model-independent parts of PYTHIA7 [3, 4], the original project of rewriting the Lund family of EGs in C++. When the corresponding rewrite of the HERWIG program [5] started it was decided to use the same basic infrastructure as PYTHIA7 and therefore the THEPEG was factorized out of PYTHIA7 and is now the base of both PYTHIA7 and HERWIG++ [6]. Also the coming C++ version of ARIADNE [7] is using THEPEG.

THEPEG uses CLHEP [8] and adds on a number of general utilities such as smart pointers, extended type information, persistent I/O, dynamic loading and some extra utilities for kinematics, phase space generation etc.

The actual event generation is then performed by calling different *handler* classes for hard partonic sub-processes, parton densities, QCD cascades, hadronization etc. To implement a new model to be used by THEPEG, the procedure is then to write a new C++ class inheriting from a corresponding handler class and implement a number of pre-defined virtual functions. Eg. a class for implementing a new hadronization model would inherit from the abstract `HandronizationHandler` class, and a new parton density parameterization would inherit from the `PDFBase` class.

To generate events with THEPEG one first runs a setup program where an `EventGenerator` object is set up to use different models for different steps of the generation procedure. All objects to be chosen from are stored in a *repository*, within which it is also possible to modify switches and parameters of the implemented models in a standardized fashion, using so called *interface* objects. Typically the user would choose from a number of pre-defined `EventGenerator` objects and only make minor changes for the specific simulation to be made. When an `EventGenerator` is properly set up it is saved persistently to a file which can then be read into a special run program to perform the generation, in which case special `AnalysisHandler` objects may be specified to analyze the resulting events. Alternatively it can be read into eg. a detector simulation program or a user supplied analysis program, where it can be used to generate events.

3 Status

3.1 THEPEG

THEPEG version 1.0α is available [1] and is working. As explained above, it contains the basic infrastructure for implementing and running event generation models. It also contains some simple physics models, such as some $2 \to 2$ matrix elements, a few parton density parameterizations and a near-complete set of particle decays. However, these are mainly in place for testing purposes, and to generate realistic events, the PYTHIA7 and/or HERWIG++ programs are needed.

Currently the program only works under Linux using the `gcc` compiler. This is mainly due to the use of dynamic linking of shared object files, which is inherently platform-dependent. Recently, the build procedure has been redesigned using the `libtool` facility [9], which should allow for easy porting to other platforms in the future.

Although THEPEG includes a general structure for implementing basic fixed-order matrix element generation to produce the initial hard subprocesses in the event generation, a general procedure for reading such parton level events from external programs using the Les Houches accord [10] has been developed and will be included in the next release[1].

The documentation of THEPEG is currently quite poor. Recently the actual code documentation was converted to Doxygen format [11], which will hopefully facilitate the documentation process. The lack of documentation means that there is currently a fairly high threshold for a beginner to start using and/or developing physics modules for THEPEG. The situation is further complicated since the user interface is currently quite primitive. THEPEG has a well worked through low-level interface to be able to set parameter and switches, etc. in classes introduced to the structure from the outside. However, the current external user interface is a simple command-line facility which is not very user-friendly. A Java interface is being worked on, but is not expected to be released until next year.

3.2 PYTHIA 7 (and PYTHIA8)

PYTHIA7 version 1.0α is available [4] and is working. It contains a reimplementation of the parton shower and string fragmentation models currently available in the 6.1 version of PYTHIA [12]. In an unfortunate turn of events, the principal PYTHIA author, Torbjörn Sjöstrand, has decided to leave the THEPEG collaboration and is currently developing a new C++ version of PYTHIA (called PYTHIA8 [13]) on his own. This means that the development of PYTHIA7 is stopped, but hopefully it will be possible to interface the different modules in PYTHIA8 so that they can be used within the general framework of THEPEG.

3.3 ARIADNE

The reimplementation of the ARIADNE [7, 14] program using the framework of THEPEG has just started and is, hence, not publically available yet. Although this is mainly a pure rewrite of the fortran version

[1]A snapshot of the current development version is available from [1]

of ARIADNE, it will contain some improvements, such as the CKKW matching [15, 16] also planned for HERWIG++. In addition, an improved version of the LDCMC [17] is planned.

4 Conclusions

THEPEG was intended to be *the* standard platform for event generation for the LHC era of collider physics. Unfortunately, this does not seem to become a reality. Besides the recent split between PYTHIA and THEPEG, there will also be other separate programs such as SHERPA [18, 19]. This is, of course, not an optimal situation, especially not for the LHC experiments, which surely would have preferred a uniform interface to different event generator models.

References

[1] L. Lönnblad *et al.*, *THEPEG program*. `http://www.thep.lu.se/ThePEG`.

[2] S. Gieseke, *HERWIG++*. These proceedings.

[3] M. Bertini, L. Lönnblad, and T. Sjöstrand, Comput. Phys. Commun. **134**, 365 (2001). `hep-ph/0006152`.

[4] L. Lönnblad *et al.*, *PYTHIA7 program*. `http://www.thep.lu.se/Pythia7`.

[5] G. Corcella *et al.*, JHEP **01**, 010 (2001). `hep-ph/0011363`.

[6] S. Gieseke, A. Ribon, M. H. Seymour, P. Stephens, and B. Webber, JHEP **02**, 005 (2004). `hep-ph/0311208`.

[7] L. Lönnblad, Comput. Phys. Commun. **71**, 15 (1992).

[8] L. Lönnblad, Comput. Phys. Commun. **84**, 307 (1994).

[9] G. Matzigkeit *et al.*, *`libtool` program*. `http://www.gnu.org/software/libtool`.

[10] E. Boos *et al.* (2001). `hep-ph/0109068`.

[11] D. van Heesch, *The doxygen documentation system*. `http://www.doxygen.org`.

[12] T. Sjöstrand, and others, Comput. Phys. Commun. **135**, 238 (2001). `arXiv:hep-ph/0010017`.

[13] T. Sjöstrand, *PYTHIA*. These proceedings.

[14] L. Lönnblad, *ARIADNE at HERA and at the LHC*. These proceedings.

[15] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, JHEP **11**, 063 (2001). `hep-ph/0109231`.

[16] L. Lönnblad, JHEP **05**, 046 (2002). `hep-ph/0112284`.

[17] H. Kharraziha and L. Lönnblad, JHEP **03**, 006 (1998). `hep-ph/9709424`.

[18] T. Gleisberg *et al.*, JHEP **02**, 056 (2004). `hep-ph/0311263`.

[19] T. Gleisberg *et al.*, *The event generator SHERPA*. These proceedings.

PYTHIA

T. Sjöstrand
CERN, Geneva, Switzerland, and
Department of Theoretical Physics, Lund University, Sweden

The PYTHIA program is a standard tool for the generation of high-energy collisions, containing a realistic description of the full story, from a hard interaction involving a few partons to an observable hadronic final state of hundreds of particles. The current PYTHIA 6.3 version is described in detail in the manual [1], with the most recent update notes to be found on the PYTHIA webpage
`http://www.thep.lu.se/∼torbjorn/Pythia.html`,
together with the code itself, sample main programs and some further material. The latest published version is [2] and a recent brief review is found in [3]. The 6.3 version includes new transverse-momentum-ordered showers and a new multiple-interactions and beam-remnant scenario [4], described elsewhere in these proceedings.

From the onset, all PYTHIA code has been written in Fortran 77. For the LHC era, the experimental community has made the decision to move heavy computing completely to C++. Hence the main future development line is PYTHIA 8, which is a re-implementation in C++. Many obsolete options will be removed and various aspects modernized in the process.

With the rise of automatic matrix-element code generation and phase-space sampling, input of process-level events via the Les Houches Accord (LHA) [5] reduces the need to have extensive process libraries inside PYTHIA itself. Thus emphasis is on providing a good description of subsequent steps of the story, involving elements such as initial- and final-state parton showers, multiple parton–parton interactions, string fragmentation, and decays. All the latter components now exist as C++ code, even if in a preliminary form, with finer details to be added, and still to be better integrated and tuned. At the current stage, however, there is not even the beginning of a PYTHIA 8 process library; instead a temporary interface is provided to PYTHIA 6, so that all hard processes available there can be generated and sent on to PYTHIA 8, transparent to the user.

PYTHIA 8 is intended to be a standalone program, i.e. does not require any external libraries. However, in addition to the LHA interface, hooks also exist for external parton distribution functions, particle decays and random numbers, and more may be added.

This project was started in September 2004, and so is still at an early stage. A first public version, PYTHIA 8.040, can be found on the PYTHIA webpage (look under the "Future" link). This should be viewed as a development snapshot, to allow early feedback from the LHC experimental community, and cannot be used for any serious physics studies. It is intended/hoped that a first realistic version, PYTHIA 8.100, could be ready by early 2007, but even this version will be clearly limited in its capabilities, and strongly focused on LHC applications. It is therefore to be expected that PYTHIA 6 and PYTHIA 8 will co-exist for several years.

References

[1] T. Sjöstrand, L. Lönnblad, S. Mrenna, and P. Skands (2003). `hep-ph/0308153`.

[2] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001). `hep-ph/0010017`.

[3] M. A. Dobbs *et al.* (2004). `hep-ph/0403045`.

[4] T. Sjöstrand and P. Z. Skands, JHEP **03**, 053 (2004). `hep-ph/0402078`;
T. Sjöstrand and P. Z. Skands, Eur. Phys. J. **C39**, 129 (2005). `hep-ph/0408302`.

[5] E. Boos *et al.* (2001). `hep-ph/0109068`.

HERWIG

Michael H. Seymour
School of Physics & Astronomy, University of Manchester, and Theoretical Physics Group, CERN

Abstract
I review the status of the current fortran version of HERWIG. Progress towards its replacement, Herwig++, is reviewed elsewhere in these proceedings.

1 Introduction

HERWIG [1] is a Monte Carlo event generator for simulation of hadronic final states in lepton–lepton, lepton–hadron and hadron–hadron collisions. It incorporates important colour coherence effects in the final state [2] and initial state [3] parton showers, as well as in heavy quark processes [4] and the hard process generation [5]. It uses the cluster [6] hadronization model and a cluster-based simulation of the underlying event [7]. While earlier versions [8] concentrated on QCD and a few other SM processes, recent versions contain a vast library of MSSM [9] and other BSM processes. A review of current Monte Carlo event generators including HERWIG can be found in [10].

We are currently in a period of intense activity, finalizing the HERWIG program and writing a completely new event generator, HERWIG++. In this very short contribution, I can do little more than mention the areas of progress and provide references to sources of more details.

2 HERWIG version 6.5

HERWIG version 6.5 was released [11] in October 2002. Its main new features were an interface to the Les Houches Accord event format [12], the hooks needed by the MC@NLO package [13] and various bug fixes and minor improvements. It was advertised as the final fortran version of HERWIG before work switched to HERWIG++.

Despite this, the period since then has seen intense development with several new subversion releases and new features, most notably version 6.505, which featured an improved interface to the Jimmy generator for multiparton interactions, which I will discuss in more detail shortly. The most recent version is 6.507, which can be obtained from the HERWIG web site [14].

Development of fortran HERWIG is now slowing, and the only new feature still foreseen is the implementation of matrix element corrections to the production of Higgs bosons, both SM and MSSM, preliminary versions of which have been discussed in [15]. Beyond this, the HERWIG collaboration has made a commitment to all running (and ceased) experiments to support their use of HERWIG throughout their lifetimes. Due to lack of manpower, making the same promise to the LHC experiments would divert too much effort away from support of HERWIG++, and we will only support their use of HERWIG until we believe that HERWIG++ is a stable alternative for production running.

3 Jimmy

Early versions of the Jimmy model [16] generated jet events in photoproduction using a multiparton interaction picture. The recent update [17] enables it to work efficiently as a generator of underlying events in high E_T jet events and other hard processes in hadron–hadron collisions for the first time. For a given pdf set, the main adjustable parameters are `PTJIM`, the minimum transverse momentum of partonic scattering, and `JMRAD(73)`, related to the effective proton radius. Varying these one is able to get a good description of the CDF data [18] and other data held in the JetWeb database [19] that are sensitive to underlying event effects in hard process events. However, a poor description of minimum bias data in which there is no hard scale is still obtained. This is probably due to the fact that `PTJIM` is a hard cutoff and there is no soft component below it; preliminary attempts to rectify this are encouraging [20]. It is interesting to note that with tunings that give equally good descriptions of current data, Jimmy predicts twice as much underlying event activity as PYTHIA at the LHC.

References

[1] G. Corcella *et al.*, JHEP **0101** (2001) 010 [arXiv:hep-ph/0011363].

[2] G. Marchesini and B.R. Webber, Nucl. Phys. B **238** (1984) 1.

[3] G. Marchesini and B.R. Webber, Nucl. Phys. B **310** (1988) 461.

[4] G. Marchesini and B.R. Webber, Nucl. Phys. B **330** (1990) 261.

[5] R.K. Ellis, G. Marchesini and B.R. Webber, Nucl. Phys. B **286** (1987) 643 [Erratum-ibid. B **294** (1987) 1180].

[6] B.R. Webber, Nucl. Phys. B **238** (1984) 492.

[7] G. Marchesini and B.R. Webber, Phys. Rev. D **38** (1988) 3419.

[8] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, Comput. Phys. Commun. **67** (1992) 465.

[9] S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, JHEP **0204** (2002) 028 [arXiv:hep-ph/0204123].

[10] M.A. Dobbs *et al.*, "Les Houches guidebook to Monte Carlo generators for hadron collider physics", *Workshop on Physics at TeV Colliders, Les Houches, France, 26 May–6 June 2003*, arXiv:hep-ph/0403045.

[11] G. Corcella *et al.*, "HERWIG 6.5 release note", arXiv:hep-ph/0210213.

[12] E. Boos *et al.*, "Generic user process interface for event generators", *Workshop on Physics at TeV Colliders, Les Houches, France, 21 May–1 June 2001*, arXiv:hep-ph/0109068.

[13] S. Frixione and B.R. Webber, JHEP **0206** (2002) 029 [arXiv:hep-ph/0204244];
S. Frixione and B.R. Webber, "The MC@NLO event generator", arXiv:hep-ph/0207182;
S. Frixione, P. Nason and B.R. Webber, JHEP **0308** (2003) 007 [arXiv:hep-ph/0305252].

[14] http://hepwww.rl.ac.uk/theory/seymour/herwig/ also with a link to Jimmy's web page.

[15] G. Corcella and S. Moretti, Phys. Lett. B **590** (2004) 249 [arXiv:hep-ph/0402146];
G. Corcella and S. Moretti, "Matrix-element corrections to $\mathrm{gg}/\mathrm{q}\bar{\mathrm{q}} \to$ Higgs in HERWIG", *Workshop on Physics at TeV Colliders, Les Houches, France, 26 May–6 June 2003*, arXiv:hep-ph/0402149.

[16] J.M. Butterworth, J.R. Forshaw and M.H. Seymour, Z. Phys. C **72** (1996) 637 [arXiv:hep-ph/9601371].

[17] C.M. Buttar et al., these proceedings.
J.M. Butterworth and M.H. Seymour, in preparation.

[18] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **65** (2002) 092002.

[19] J.M. Butterworth and S. Butterworth, Comput. Phys. Commun. **153** (2003) 164 [arXiv:hep-ph/0210404].

[20] I. Borozan and M.H. Seymour, JHEP **0209** (2002) 015 [arXiv:hep-ph/0207283].

Herwig++

Stefan Gieseke
Institut für Theoretische Physik, Universität Karlsruhe, 76128 Karlsruhe, Germany

Abstract
I briefly review the status of the Herwig++ event generator. Current achievements are highlighted and a brief summary of future plans is given.

1 Introduction

Herwig++ [1] is a new Monte Carlo event generator for simulating collider physics, written in the object oriented programming language C++. The idea is to rewrite the well-established multi–purpose event generator HERWIG [2] and to improve it where necessary [3]. The Lund event generators PYTHIA [4] and ARIADNE [5] are also being rewritten at the moment. Herwig++ and ARIADNE will both be based on a common event generation framework, called ThePEG [6] which will make it possible to exchange single modules of the event generation and allows us to have a common, or at least a very similar user interface. PYTHIA8, the rewrite of PYTHIA (6.3) will be written independently of this project but may become integrated into the structure later [7]. A further object oriented event generator, SHERPA [8], is established as an independent project.

2 Event Generation

In its present version (1.0) Herwig++ is capable of simulating e^+e^- annihilation events. The physics simulation consists of several steps, going from small (perturbative) to large (non–perturbative) distance scales. First, the effective CM energy of the annihilating e^+e^- pair is selected according to some model structure function of the electron, thereby radiating photons that carry some fraction of the original energy. Next, we set up the $q\bar{q}$ final state and a hard matrix element correction is applied [9]. In the next step, parton showers are radiated from the coloured final state particles. These effectively resum large soft and collinear logarithms. The parton shower is modelled in terms of new evolution variables with respect to the FORTRAN program [10]. This, and the use of splitting functions for massive particles allow us to simulate the suppression of soft and collinear radiation from heavy particles dynamically (dead cone effect) which has previously only been modelled in a crude way. Parton showers from initial state particles in a hard scattering and from decays of heavy particles (particularly t–quarks) have been formulated for various situations in [10]. The next stage of the simulation is the hadronization of the outgoing coloured particles. First, remaining gluons are split into non–perturbative $q\bar{q}$–pairs. Colour connected particles are paired into colourless clusters. The invariant mass spectrum of these clusters contains a long high–mass tail that still contains a large scale. These heavy clusters are further split into pairs of lighter clusters. Once all clusters are below a certain mass threshold they decay into pairs of hadrons. The hadron species are selected only according to a handful of parameters. It is this stage where it has been observed in previous versions of the FORTRAN program that the meson/baryon number ratio in e^+e^- annihilation events was difficult to obtain when a large number of highly excited mesons is available in the program [11]. In the current version the hadron selection is reorganised and we obtain more stable results. Finally, the produced hadrons decay into stable hadrons according to some models. In version 1.0 the hadronic decays were modelled similarly to the decays in the FORTRAN version.

We have tested the simulation of e^+e^- annihilation events in very great detail [1]. We considered event shape variables, jet rates, hadron yields and many more observables. One observable of special interest has been the b–quark fragmentation function that we found to be well–described on the basis of the parton shower only. This is a result of the new shower algorithm for heavy quarks. The overall result was that we are capable of simulating e^+e^- events at least as well good as with the FORTRAN version.

3 Current and Future Developments

Many new features are currently being implemented for the event simulation at hadron colliders. The list of hard matrix elements will be slightly extended in the next version in order to cover some basic processes. In principle we can also rely on parton level event generators and read in event files that follow the Les Houches Accord [12]. The parton shower will include initial state radiation and gluon radiation in the perturbative decay of heavy particles. Some related aspects of estimating uncertainties from initial state parton showers were addressed in [13]. A large effort went into remodelling and updating the secondary hadronic decays. A future version should also be able to simulate hard jets in deep inelastic scattering. Exhaustive tests of our generator output against current data from the experiments at HERA and the Tevatron will be made in order to validate and understand our program. In the long–term we plan to include a larger number of simple processes, mainly $2 \to 2$ and some $2 \to 3$, both Standard Model processes and some BSM processes as well. The modelling of the underlying event will at first only be on the basis of the simple so–called UA5 model that is also available in the FORTRAN version. A refinement towards a more sophisticated multiple interaction model [14, 15] is planned.

References

[1] S. Gieseke, A. Ribon, M. H. Seymour, P. Stephens and B. Webber, JHEP **0402**, 005 (2004) [arXiv:hep-ph/0311208].

[2] G. Corcella *et al.*, arXiv:hep-ph/0210213;
G. Corcella *et al.*, JHEP **0101** (2001) 010 [arXiv:hep-ph/0011363].

[3] S. Gieseke, in Proc. *Hadron Collider Physics 2002*, eds. M. Erdmann, Th. Müller [arXiv:hep-ph/0210294].

[4] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands, arXiv:hep-ph/0308153; T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun. **135** (2001) 238 [arXiv:hep-ph/0010017].

[5] L. Lonnblad, Comput. Phys. Commun. **71** (1992) 15.

[6] M. Bertini, L. Lonnblad and T. Sjostrand, Comput. Phys. Commun. **134** (2001) 365 [arXiv:hep-ph/0006152].

[7] T. Sjöstrand, this workshop.

[8] T. Gleisberg, S. Hoche, F. Krauss, A. Schalicke, S. Schumann and J. C. Winter, JHEP **0402** (2004) 056 [arXiv:hep-ph/0311263].

[9] M. H. Seymour, Comput. Phys. Commun. **90** (1995) 95 [arXiv:hep-ph/9410414].

[10] S. Gieseke, P. Stephens and B. Webber, JHEP **0312**, 045 (2003) [arXiv:hep-ph/0310083].

[11] A. Kupco, in Proc. *Monte Carlo generators for HERA physics*, (Hamburg 1998–1999), eds. A.T. Doyle, G. Grindhammer, G. Ingelman, H. Jung. Hamburg, DESY-PROC-1999-02, 292 [arXiv:hep-ph/9906412].

[12] E. Boos *et al.*, arXiv:hep-ph/0109068.

[13] S. Gieseke, JHEP **0501**, 058 (2005) [arXiv:hep-ph/0412342].

[14] J. M. Butterworth, J. R. Forshaw and M. H. Seymour, Z. Phys. C **72** (1996) 637 [arXiv:hep-ph/9601371].

[15] I. Borozan and M. H. Seymour, JHEP **0209**, 015 (2002) [arXiv:hep-ph/0207283].

The Event Generator SHERPA

T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann, J. Winter
Institut für theoretische Physik, TU Dresden, D-01062 Dresden, Germany
E-mail: steffen@theory.phy.tu-dresden.de

Abstract
In this contribution the multi-purpose event generation framework SHERPA is presented and the development status of its physics modules is reviewed. In its present version, SHERPA is capable of simulating lepton-lepton, lepton-photon, photon-photon and fully hadronic collisions, such as proton-proton reactions.

SHERPA [1] is an independent approach for a framework for event generation at high energy collider experiments. The program is entirely written in the object-oriented programming language C++. This is reflected in particular in the structure of the program. In SHERPA, the various tasks related to the generation of events are encapsulated in a number of specific modules. These physics modules are initialized and steered by the SHERPA framework. This structure facilitates a high modularity of the actual event generator and allows for the easy replacement/modification of entire physics models, e.g. the parton shower or the fragmentation model. The current version SHERPA-1.0.6 incorporates the following physics modules:

- ATOOLS-2.0.6: This is the toolbox for all other modules. ATOOLS contain classes with mathematical tools like vectors and matrices, organization tools such as read-in or write-out devices, and physics tools like particle data or classes for the event record.
- BEAM-1.0.6: This module manages the treatment of the initial beam spectra for different colliders. At the moment two options are implemented for the beams: they can either be monochromatic, and therefore require no extra treatment, or, for the case of an electron collider, laser backscattering off the electrons is supported leading to photonic initial states.
- PDF-1.0.6: In this module the handling of initial state radiation (ISR) is located. It provides interfaces to various proton and photon parton density functions, and to the LHAPDFv3 interface. In addition, an analytical electron structure function is supplied.
- MODEL-1.0.6: This module comprises the basic physics parameters (like masses, mixing angles, etc.) of the simulation run. Thus it specifies the corresponding physics model. Currently three different physics models are supported: the Standard Model (SM), its Minimal Supersymmetric extension (MSSM) and the ADD model of large extra dimensions. For the input of MSSM spectra a run-time interface to the program Isasusy 7.67 [2] is provided. The next release of SHERPA will in addition support the SLHA format of spectrum files [3].
- EXTRA_XS-1.0.6: In this module a collection of analytic expressions for simple $2 \to 2$ processes within the SM and the corresponding classes embedding them into the SHERPA framework are provided. This includes methods used for the definition of the parton shower evolution, such as color connections and the hard scale of the process. The classes for phase space integration, which are common with AMEGIC, are located in a special module called PHASIC.
- AMEGIC++-2.0.6: AMEGIC [4] is SHERPA's own matrix element generator. It works as a generator-generator: during the initialization run the matrix elements for a set of given processes within the SM, the MSSM or the ADD model, as well as their specific phase space mappings are created by AMEGIC and stored in library files. In the initialization of the production run, these libraries are linked to the program. They are used to calculate cross sections and to generate single events based on them.

- PHASIC++-1.0.6: Here all classes dealing with the Monte Carlo phase space integration are located. As default the adaptive multi-channel method of [5] together with a Vegas optimization [6] for the single channels is used for the evaluation of the initial state and final state integrals.
- APACIC++-2.0.6: APACIC [7] contains classes for the simulation of both the initial and the final state parton shower. The sequence of parton emissions in the shower evolution is organized in terms of the parton's virtual mass as ordering parameter. Coherence effects are accounted for by explicit ordering of the opening angles in subsequent branchings. This treatment is similar to the Pythia [8] parton shower approach. All features for a consistent merging with matrix elements [9] are included.
- AMISIC++-1.0.6: AMISIC contains classes for the simulation of multiple parton interactions according to [10]. SHERPA extends this treatment of multiple interactions by allowing for the simultaneous evolution of an independent parton shower in each of the subsequent (semi-)hard collisions. This shower evolution is done by the APACIC module.
- SHERPA-1.0.6: Finally, SHERPA is the steering module that initializes, controls and evaluates the different phases in the entire process of event generation. Furthermore, all necessary routines for combining the parton showers and matrix elements, which are independent of the specific parton shower are found in this module. In addition, this subpackage provides an interface to the Lund String Fragmentation of Pythia 6.214 including its hadron decay routines.

SHERPA is publicly available from http://www.sherpa-mc.de. It has successfully been tested for various processes of great relevance at future colliders [11]. Present activities of developing SHERPA cover the modeling of the underlying event and an alternative fragmentation model [12].

References

[1] Gleisberg, T. and others, JHEP **02**, 056 (2004).

[2] Paige, F. E. and Protopescu, S. D. and Baer, H. and Tata, X., hep-ph/0312045.

[3] Skands, P. and others, JHEP **07**, 036 (2004).

[4] Krauss, F. and Kuhn, R. and Soff, G., JHEP **02**, 044 (2002).

[5] Kleiss, R. and Pittau, R., Comput. Phys. Commun. **83**, 141 (1994);
Berends, F. A. and Pittau, R. and Kleiss, R., Nucl. Phys. **B424**, 308 (1994).

[6] Lepage, G. P. CLNS-80/447.

[7] Kuhn, R. and Krauss, F. and Ivanyi, B. and Soff, G., Comput. Phys. Commun. **134**, 223 (2001);
Krauss, F. and Schälicke, A. and Soff, G., hep-ph/0503087.

[8] Sjöstrand, T., Comput. Phys. Commun. **82**, 74 (1994);
Sjöstrand, T. and Lönnblad, L. and Mrenna, S. and Skands, P., hep-ph/0308153.

[9] Catani, S. and Krauss, F. and Kuhn, R. and Webber, B. R., JHEP **11**, 063 (2001);
Krauss, F., JHEP **08**, 015 (2002).

[10] Sjöstrand, T. and van Zijl, M., Phys. Rev. **D36**, 2019 (1987).

[11] Gleisberg, T. and others, JHEP **09**, 001 (2003);
Gleisberg, T. and others, Eur. Phys. J. **C34**, 173 (2004);
Krauss, F. and Schälicke, A. and Schumann, S. and Soff, G., Phys. Rev. **D70**, 114009 (2004);
Krauss, F. and Schälicke, A. and Schumann, S. and Soff, G., hep-ph/0503280;
Gleisberg, T. and Krauss, F. and Schälicke, A. and Schumann, S. and Winter, J., hep-ph/0504032.

[12] Winter, J. and Krauss, F. and Soff, G., Eur. Phys. J. **C36**, 381 (2004).

ARIADNE at HERA and at the LHC

Leif Lönnblad
Department of Theoretical Physics, Lund University, Sweden

Abstract
I describe briefly the status of the ARIADNE program implementing the Dipole Cascade Model and comment both on its performance at HERA, and the uncertainties relating to the extrapolation to LHC energies.

1 Introduction

ARIADNE [1] is a Fortran subroutine library to be used with the PYTHIA event generator [2]. By simply adding a few lines to a PYTHIA steering routine, the PYTHIA parton shower is replaced by the dipole cascade in ARIADNE. For lepton–hadron DIS it can also be used together with the LEPTO [3] generator in a similar fashion. However, even if it thus simple to use ARIADNE also for the LHC, there are a few caveats of which the user should be aware. In this brief presentation of the program, I will first go through the main points of the final-state dipole shower relevant for $\mathrm{e}^+\mathrm{e}^-$-annihilation, then I will present the extention of the model to lepton–hadron DIS, and finally describe how the model works for hadron–hadron collisions.

2 The Basic Dipole Model

In the Dipole Cascade Model (DCM) [4, 5], the bremsstrahlung of gluons is described in terms of radiation from colour dipoles between gluons and quarks. Thus, in an $\mathrm{e}^+\mathrm{e}^- \to q\bar{q}$ event, a gluon, g_1 may be emitted from the colour-dipole between the q and $\bar{q}$. In this emission the initial dipole is replaced by two new ones, one between q and g_1 and one between g_1 and $\bar{q}$. These may then continue radiating independently in a cascade where each step is a $2 \to 3$ partonic splitting or, equivalently, a splitting of a dipole into two. The splittings are ordered in a transverse momentum variable, $\mathrm{p}_\perp$, defined in a Lorentz-invariant fashion, which also defines the scale in α_S.

There are several advantages of this model. One is that the coherence effects approximated by angular ordering [6] in eg. the HERWIG [7] parton cascade, are automatically included. Another is that the first order $\mathrm{e}^+\mathrm{e}^- \to qg\bar{q}$ matrix element correction is in some sense built-in. A major *dis*advantage is that the $g \to q\bar{q}$ splitting does not enter naturally in this formalism. Final-state $g \to q\bar{q}$ splittings are, however, easy to add [8] and for final-state cascades in $\mathrm{e}^+\mathrm{e}^-$-annihilation the description is complete. Ariadne is generally considered to be the program which best reproduces event shapes and other hadronic final-state observables at LEP (see eg. [9]).

3 ARIADNE at HERA

While for $\mathrm{e}^+\mathrm{e}^-$-annihilation, the DCM is formally equivalent to conventional angular ordered parton showers to modified leading logarithmic accuracy, the situation for collisions with incoming hadrons is quite different. In a conventional shower the struck quark in eg. lepton–hadron DIS is evolved backwards with an initial-state cascade according to DGLAP [10, 11, 12, 13] evolution. In contrast, the DCM model for DIS [14] describes all gluon emissions in terms of final-state radiation from colour-dipoles, in a similar way as in $\mathrm{e}^+\mathrm{e}^-$-annihilation, with the initial dipole now being between the struck quark and the hadron remnant. Contrary to $\mathrm{e}^+\mathrm{e}^-$-annihilation, the remnant must now be treated as an extended object and, since radiation of small wavelengths from an extended antenna is suppressed, the emission of high-$p_\perp$ gluons in the DCM is suppressed in the remnant direction.

Despite this suppression, which is modeled semi-classically, the net result is that gluon emissions are allowed in a much larger phase space region than in a conventional parton shower, especially for limited Q^2 values. Although the emissions are ordered in $p_\perp$, they are not ordered in rapidity (or x). Hence, if tracing the emissions in rapidity, they will be unordered in $p_\perp$, and there are therefore qualitative similarities between the DCM and BFKL evolution [15, 16, 17]. This is in contrast to conventional showers which are purely DGLAP-based and where the emissions are ordered both in scale and in x. One of the striking phenomenological consequences of this is that ARIADNE is one of the few programs which are able to describe the high rate of forward (in the proton direction) jets measured in small-x DIS at HERA [18, 19, 20], an observable which conventional parton showers completely fail to reproduce. In fact, ARIADNE is in general considered to be the program which best describe hadronic final-state observables at HERA [21].

This does not mean that the DCM is perfect in any way. Most notably, the initial-state $g \to q\bar{q}$ and $q \to g^\star q$ (where the q is emitted into the final-state) splittings are not easily included. While the former process has been included as an explicit initial-state splitting step [22], the latter is currently absent in the the ARIADNE program. In addition, the treatment of the initial-state $g \to q\bar{q}$ splitting has been found to be somewhat incomplete, as it by construction imposes ordering in both $p_\perp$ and rapidity, thus excluding certain regions of the allowed phase space. At HERA, the incomplete treatment of the $g \to q\bar{q}$ and $q \to g^\star q$ splittings can be shown to be a small effect. However, this is not always the case at the LHC.

4 ARIADNE at LHC

Given the great success of ARIADNE at LEP and HERA, it is natural to assume that it also would do a good job at the Tevatron and the LHC. In principle, the extention of the DCM to hadron–hadron collisions is trivial, and indeed it is simple to run ARIADNE together with PYTHIA for hadron–hadron collisions. Whichever hard sub-process, PYTHIA generates, the relevant dipoles between hard partons and hadron remnants are constructed and are allowed to radiate. In addition, the initial-state $g \to q\bar{q}$ splittings are included from both sides. However, for many processes there are modifications needed.

The most obvious processes are Drell-Yan and vector boson production, where a quark from one hadron annihilates with an anti-quark from the other. The gluon radiation is then initiated by the dipole between the two remnants, and we have a suppression in both directions. However, it is unphysical to give the remnants a large transverse momentum from the recoil of the gluon emission. In DIS, this is resolved by introducing so-called recoil gluons [14], but here it is clear that the recoil should be taken by the vector boson or the Drell-Yan lepton pair. Such a procedure was introduced in [23], and together with a correction where the first emission is matched to the $qg \to qZ$ and $q\bar{q} \to gZ$ matrix elements, it describes well eg. the Z^0 $p_\perp$ spectrum measured at the Tevatron [24, 25]. There are still some differences wrt. conventional parton showers. Eg. the rapidity correlation between the vector boson and the hardest jet is more flat in ARIADNE due to the increased phase space for emissions [26]. Although W and Z^0 production at the Tevatron is not a small-x process, the effect is related to higher rate of forward jets in ARIADNE for DIS. Such correlations have not yet been measured at the Tevatron, but another related effect is the somewhat harder $p_\perp$-spectrum of the Z^0 for low $p_\perp$ in ARIADNE, which is compatible with Tevatron measurements [26]. For a conventional cascade to be able to describe the low-$p_\perp$ spectrum, a quite substantial "non-perturbative" intrinsic transverse momentum must be added to the incoming quarks [27, 28].

Going from the Tevatron to the LHC, there is a substantial increase in phase space for QCD radiation, and it can be argued that W and Z^0 production at the LHC is a small-x process with $x \propto m_Z/\sqrt{S} < 0.01$. Indeed ARIADNE predicts a harder $p_\perp$-spectrum for the W at the LHC as compared to conventional showers [29].

Also Higgs production can be argued to be almost a small-x process at the LHC, if the Higgs is found with a mass around the "most likely" value of $\approx$ 120 GeV. However, Higgs production is a gluon-initiated process, and the absence of the $q \to g^\star q$ splitting is a serious deficiency giving a much softer $p_\perp$-spectrum for the Higgs in ARIADNE as compared to conventional showers [30]. Hence the predictions from ARIADNE for this and similar processes can currently not be trusted. Furthermore, the increased phase-space at the LHC means that predictions also for quark-initiated processes may become affected by the deficiencies in the treatment of initial-state $g \to q\bar{q}$ mentioned above.

5 Conclusion

The success of the DCM as implemented in ARIADNE in describing hadronic final-state observables as measured at LEP and HERA makes it tempting to use it also to make predictions for the LHC. The temptation is even more difficult to resist as it is so simple to run ARIADNE together with PYTHIA for any LHC process. Currently, this must be done with great care. As explained above, it is possible to obtain reasonable predictions for vector boson production. Also standard jet-production should be fairly safe. However, for Higgs production, one of the most interesting processes at LHC, ARIADNE in its current state turns out to be quite useless.

ARIADNE is currently being rewritten in C++ within the framework of THEPEG [31, 32]. The planned features includes a remodeling of initial-state $g \to q\bar{q}$ splittings as well as the introduction of the $q \to g^\star q$ process. In addition the matching to fixed-order tree-level matrix elements à la CKKW [33, 34, 26, 35] will be implemented for the most common sub-processes. When this version is released, hopefully during 2006, it should therefore be safe to use ARIADNE to produce LHC predictions.

References

[1] L. Lönnblad, Comput. Phys. Commun. **71**, 15 (1992).

[2] T. Sjöstrand, and others, Comput. Phys. Commun. **135**, 238 (2001). `arXiv:hep-ph/0010017`.

[3] G. Ingelman, A. Edin, and J. Rathsman, Comput. Phys. Commun. **101**, 108 (1997). `hep-ph/9605286`.

[4] G. Gustafson, Phys. Lett. **B175**, 453 (1986).

[5] G. Gustafson and U. Pettersson, Nucl. Phys. **B306**, 746 (1988).

[6] G. Marchesini and B. R. Webber, Nucl. Phys. **B310**, 461 (1988).

[7] G. Corcella *et al.*, JHEP **01**, 010 (2001). `hep-ph/0011363`.

[8] B. Andersson, G. Gustafson, and L. Lönnblad, Nucl. Phys. **B339**, 393 (1990).

[9] K. Hamacher and M. Weierstall (1995). `hep-ex/9511011`.

[10] V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15**, 781 (1972).

[11] L. N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975).

[12] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).

[13] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[14] B. Andersson, G. Gustafson, L. Lönnblad, and U. Pettersson, Z. Phys. **C43**, 625 (1989).

[15] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **44**, 443 (1976).

[16] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **45**, 199 (1977).

[17] I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[18] H1 Collaboration, C. Adloff *et al.*, Nucl. Phys. **B538**, 3 (1999). hep-ex/9809028.

[19] ZEUS Collaboration, J. Breitweg *et al.*, Eur. Phys. J. **C6**, 239 (1999). hep-ex/9805016.

[20] ZEUS Collaboration, S. Chekanov *et al.* (2005). hep-ex/0502029.

[21] N. Brook, R. G. Waugh, T. Carli, R. Mohr, and M. Sutton. Prepared for Workshop on Future Physics at HERA (Preceded by meetings 25-26 Sep 1995 and 7-9 Feb 1996 at DESY), Hamburg, Germany, 30-31 May 1996.

[22] L. Lönnblad, Z. Phys. **C65**, 285 (1995).

[23] L. Lönnblad, Nucl. Phys. **B458**, 215 (1996). hep-ph/9508261.

[24] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **84**, 845 (2000). hep-ex/0001021.

[25] D0 Collaboration, B. Abbott *et al.*, Phys. Rev. **D61**, 032004 (2000). hep-ex/9907009.

[26] N. Lavesson and L. Lonnblad, JHEP **07**, 054 (2005). hep-ph/0503293.

[27] E. Thome (2004). hep-ph/0401121.

[28] E. L. Nurse. FERMILAB-THESIS-2005-05.

[29] E. Richter-Was and B. Kersevan, *The Monte Carlo Event Generator AcerMC and package AcderDET*. These Proceedings.

[30] Z. Czyczula *et al.*, *Multi-jet production and Multi-scale QCD*. These Proceedings.

[31] L. Lönnblad, *ThePEG: Toolkit for high energy physics event generation*. These Proceedings.

[32] M. Bertini, L. Lönnblad, and T. Sjöstrand, Comput. Phys. Commun. **134**, 365 (2001). hep-ph/0006152.

[33] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, JHEP **11**, 063 (2001). hep-ph/0109231.

[34] L. Lönnblad, JHEP **05**, 046 (2002). hep-ph/0112284.

[35] S. Höche *et al.*, *Matching Parton Showers and Matrix Elements*. These Proceedings.

The Monte Carlo Event Generator AcerMC and package AcerDET

B. Kersevan$^{(1),(2)}$, *E. Richter-Was*$^{(3),(4),*}$
(1)Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000, Slovenia
(2)Experimental Particle Physics Department, Jozef Stefan Institute, P.P. 3000, Jamova 39, SI-1000 Ljubljana, Slovenia.
(3)Institute of Physics, Jagiellonian University, 30-059 Krakow, ul. Reymonta 4, Poland.
(4)Institute of Nuclear Physics PAN, 31-342 Krakow, ul. Radzikowskiego 152, Poland.

Abstract
The **AcerMC** Monte Carlo Event Generator is dedicated for the generation of Standard Model background processes at *pp* LHC collisions. The program itself provides a library of the massive matrix elements and phase space modules for generation of selected processes. The hard process event, generated with one of these modules, can be completed by the initial and final state radiation, hadronisation and decays, simulated with either `PYTHIA`, `ARIADNE` or `HERWIG` Monte Carlo event generator and (optionally) with `TAUOLA` and `PHOTOS`. Interfaces to all these packages are provided in the distribution version. The matrix element code has been derived with the help of the `MADGRAPH` package. The phase-space generation is based on the multi-channel self-optimising approach using the modified Kajantie-Byckling formalism for phase space construction and further smoothing of the phase space was obtained by using a modified `ac-VEGAS` algorithm.

1 Introduction

The physics programme of the general purpose LHC experiments, ATLAS [1] and CMS [2], focuses on the searches for the *New Physics* with the distinctive signatures indicating production of the Higgs boson, SUSY particles, exotic particles, etc. The expected environment will in most cases be very difficult, with the signal to background ratio being quite low, on the level of a few percent after final selection in the signal window.

Efficient and reliable Monte Carlo generators, which enable one to understand and predict background contributions, are becoming key elements in the discovery perspective. As the cross-section for signal events is rather low, even rare Standard Model processes might become the overwhelming backgrounds in such searches. In several cases, generation of such processes is not implemented in the general purpose Monte Carlo generators, when the complicated phase space behaviour requires dedicated (and often rather complex) pre-sampling, whilst the general purpose Monte Carlo generators due to a large number of implemented processes tend to use simpler (albeit more generic) phase space sampling algorithms. In addition, the matrix element expressions for these processes are often quite lengthy and thus require complicated calculations. Only recently, with the appearance of modern techniques for automatic computations, their availability *on demand* became feasible for the tree-type processes. With the computation power becoming more and more easily available even very complicated formulas can now be calculated within a reasonable time frame.

* Supported in part by the Polish Government grant KBN 1 P03 091 27 (years 2004-2006) and by the EC FP5 Centre of Excellence "COPIRA" under the contract No. IST-2001-37259.

2 The Monte Carlo Event Generator AcerMC

The physics processes implemented in **AcerMC** library [5, 3, 4] represent such a set of cases. They are all being key background processes for the discovery in the channels characterised by the presence of the heavy flavour jets and/or multiple isolated leptons. For the Higgs boson searches, the $t\bar{t}H, ZH, WH$ with $H \to b\bar{b}$, the $gg \to H$ with $H \to ZZ^* \to 4\ell$, the $b\bar{b}h/H/A$ with $h/H/A \to \tau\tau, \mu\mu$ are the most obvious examples of such channels.

Let us shortly discuss the motivation for these few Standard Model background processes which are implemented in the **AcerMC 2.x** library.

The $t\bar{t}b\bar{b}$ production is a dominant irreducible background for the Standard Model (SM) and Minimal Supersymmetric Standard Model (MSSM) Higgs boson search in the associated production, $t\bar{t}H$, followed by the decay $H \to b\bar{b}$. Proposed analysis [1] requires identifying four b-jets, reconstruction of both top-quarks in the hadronic and leptonic mode and visibility of the peak in the invariant mass distribution of the remaining b-jets. The irreducible $t\bar{t}b\bar{b}$ background contributes about 60-70% of the total background from the $t\bar{t}$ events ($t\bar{t}b\bar{b}$, $t\bar{t}bj$, $t\bar{t}jj$).

The $Wb\bar{b}$ production is recognised as a substantial irreducible background for the Standard Model (SM) and Minimal Supersymmetric Standard Model (MSSM) Higgs boson search in the associated production, WH, followed by the decay $H \to b\bar{b}$.

The $Wt\bar{t}$ production is of interest because it contributes an overwhelming background [7] for the measurement of the Standard Model Higgs self-couplings at LHC in the most promising channel $pp \to HH \to WWWW$.

The $Z/\gamma^*(\to \ell\ell)b\bar{b}$ production has since several years been recognised as one of the most substantial irreducible (or reducible) backgrounds for the several Standard Model (SM) and Minimal Supersymmetric Standard Model (MSSM) Higgs boson decay modes as well as for observability of the SUSY particles. There is a rather wide spectrum of *regions of interest* for this background. In all cases the leptonic Z/γ^* decay is asked for, but events with di-lepton invariant mass around the mass of the Z-boson mass or with the masses above or below the resonance peak could be of interest. The presented process enters an analysis either by the accompanying b-quarks being tagged as b-jets, or by the presence of leptons from the b-quark semi-leptonic decays in these events, in both cases thus contributing to the respective backgrounds.

The $Z/\gamma^*(\to \ell\ell, \nu\nu, b\bar{b})t\bar{t}$ production is an irreducible background to the Higgs search in the invisible decay mode (case of $Z \to \nu\nu$) in the production with association to the top-quark pair [8]. With the $Z/\gamma^*(\to b\bar{b})$ it is also an irreducible resonant background to the Higgs search in the $t\bar{t}H$ production channel but with the Higgs boson decaying to the b-quark pair.

The complete **EW production** of the $gg, q\bar{q} \to (Z/W/\gamma^* \to)b\bar{b}t\bar{t}$ final state is also provided. It can be considered as a benchmark for the previous process, where only the diagrams with resonant $gg, q\bar{q} \to (Z/\gamma^* \to)b\bar{b}t\bar{t}$ are included. It thus allows the verification of the question, whether the EW resonant contribution is sufficient in case of studying the $t\bar{t}b\bar{b}$ background away from the Z-boson peak, like for the $t\bar{t}H$ with Higgs-boson mass of 120 GeV.

The $gg, q\bar{q} \to t\bar{t}t\bar{t}$ production , interesting process per se, is a background to the possible Higgs self-coupling measurement in the $gg \to HH \to WWWW$ decay, [7].

The $gg, q\bar{q} \to (WWbb \to)f\bar{f}f\bar{f}b\bar{b}$ and $gg, q\bar{q} \to (t\bar{t} \to)f\bar{f}bf\bar{f}\bar{b}$ processes give possibility to study spin correlations in the top-quark pair production and decays as well as the effect from the off-shell production. Those are important for the selection optimisation eg. in the $gg \to H \to WW$ channel, see the discussion in [9]. As an example, Fig. 1 illustrates spin correlation effects in the top-pair production and decays, namely asymmetry in the correlations between lepton and antilepton direction in the rest frame of top-quark, for events generated with $2 \to 6$ matrix element. Such correlation is absent if only $2 \to 2$ matrix element is used for events generation, followed by the independent decays of each top-quark.

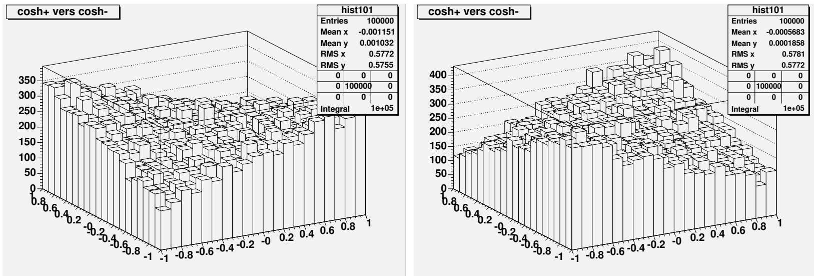


Fig. 1: *The correlations between* $\cos\Theta$ *(azimuthal angle) of lepton and antilepton from* $t\bar{t} \to \ell\bar{\nu}b\bar{\ell}\nu\bar{b}$ *decays measured in the rest frame of the top-quark with respect to the anti-top quark direction. Left plot is for* $gg \to (WWb\bar{b} \to) f\bar{f}f\bar{f}b\bar{b}$ *process, right plot for* $q\bar{q} \to (WWb\bar{b} \to) f\bar{f}f\bar{f}b\bar{b}$ *process.*

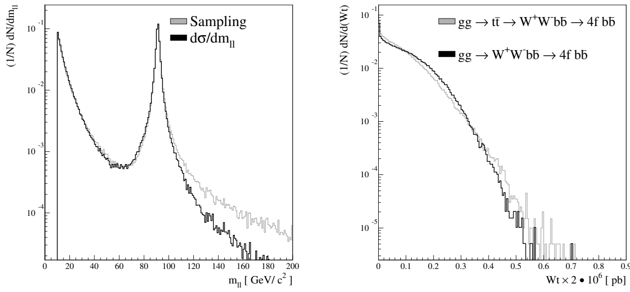


Fig. 2: *Left: A representative invariant mass distribution comparisons between the (normalised) sampling functions and the normalised differential cross-section for the* $\ell\bar{\ell}$ *pair in the process* $gg \to Z^0/\gamma^* b\bar{b} \to \ell\bar{\ell}b\bar{b}$ *process. Right: The weight distribution of the sampled events for the* $gg \to t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ *(light gray histogram) and* $gg \to b\bar{b}W^+W^- \to b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$ *(black histogram) processes. One can observe the well defined weight range for the two processes; as it turns out the weight distribution is even marginally better for the (more complex) second process, possibly because the higher number of sampling channels manage to cover the event topologies in phase space to a better extent.*

A set of **control channels**, i.e. the $\boldsymbol{q\bar{q} \to Z/\gamma^* \to \ell\ell, gg, q\bar{q} \to t\bar{t}}$, $\boldsymbol{q\bar{q} \to W \to \ell\nu}$ and $\boldsymbol{gg \to (t\bar{t} \to)WbW\bar{b}}$ processes, have been added to **AcerMC** in order to provide a means of consistency and cross-check studies.

This completes the list of the native **AcerMC** processes implemented so far. The hard process event, generated with one of these modules, can be completed by the initial and final state radiation, hadronisation and decays, simulated with either `PYTHIA`, `ARIADNE` or `HERWIG` Monte Carlo event generator and (optionally) with `TAUOLA` and `PHOTOS`. Interfaces to all these packages are provided in the distribution releases. The matrix element code has been derived with the help of the `MADGRAPH` package. The phase-space generation is based on the multi-channel self-optimising approach [3] using the modified Kajantie-Byckling formalism for phase space construction and further smoothing of the phase space was obtained by using a modified `ac-VEGAS` algorithm.

The improved and automated phase space handling provided the means to include the $2 \to 6$ processes like e.g. $gg \to t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}\ell\bar{\nu}_\ell q_1\bar{q}_2$ which would with the very complicated phase space topologies prove to be too much work to be handled manually. The studies show that the overall unweighting efficiency which can be reached in the $2 \to 6$ processes by using the recommended phase space structuring is on the order of 10 percent. An example of the implemented sampling functions and the actual differential distributions for the $gg \to Z^0/\gamma^* b\bar{b} \to \ell\bar{\ell}b\bar{b}$ process and of the weight distribution for the $gg \to t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}\ell\bar{\nu}_\ell q_1\bar{q}_2$ process are shown in Fig.2

In its latest version, the **AcerMC-2.4** package is interfaced also to `ARIADNE 4.1` [12] parton shower model and the LHAPDF structure functions [13].

It is not always the case that the matrix element calculations in the lowest order for a given topology represent the total expected background of a given type. This particularly concerns the heavy flavour content of the event. The heavy flavour in a given event might occur in the hard process of a much simpler topology, as the effect of including higher order QCD corrections (eg. in the shower mechanism). This is the case for the b-quarks present in the inclusive Z-boson or W-boson production, which has a total cross-section orders of magnitude higher than the discussed matrix-element-based $Wb\bar{b}$ or $Zb\bar{b}$ production. Nevertheless, the matrix-element-based calculation is a very good reference point to compare with parton shower approaches in different fragmentation/hadronisation models. It also helps to study matching procedures between calculations in a fixed α_{QCD} order and parton shower approaches. For very exclusive hard topologies matrix-element-based calculations represent a much more conservative approximation than the parton shower ones [6].

3 The AcerDET package

The package **AcerDET** [14] is designed to complete the **AcerMC** generator framework with the easy-to-use simulation and reconstruction algorithms for phenomenological studies on high p_T physics at LHC The package provides, starting from list of particles in the event, the list of reconstructed jets, isolated electrons, muons and photons and reconstructed missing transverse energy. The **AcerDET** represents a simplified version of the package called `ATLFAST` [15], used since several years within ATLAS Collaboration. In the **AcerDET** version some functionalities of the former one have been removed, only the most crucial detector effects are implemented and the parametrisations are largely simplified. Therefore it is not representing in details neither ATLAS nor CMS detectors. Nevertheless, we believe that the package can be well adequate for some feasibility studies of the high p_T physics at LHC.

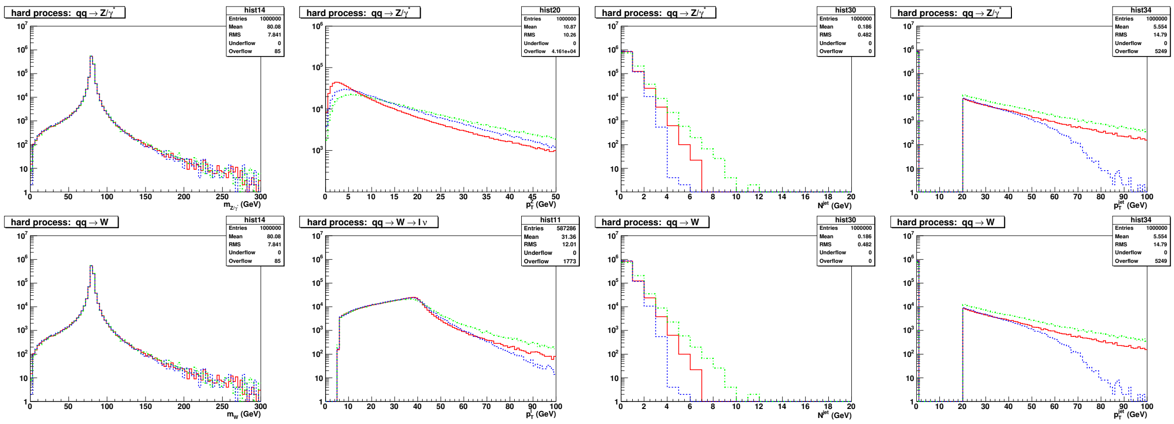


Fig. 3: *A few examples of theoretical systematic uncertainties from parton shower model on experimentally observable distributions from Drell-Yan W and Z boson production at LHC (see text).*

Fig. 3 shows possible application of the **AcerMC** control processes and **AcerDET** package for studying theoretical systematic uncertainties on the experimentally observed distributions from the parton shower model. The control channels $q\bar{q} \to Z/\gamma^* \to \ell\ell$, and $q\bar{q} \to W \to \ell\nu$ were processed with parton shower model as implemented in `PYTHIA` (red), `HERWIG` (blue) and `ARIADNE` (green). The comparison includes the distributions of the invariant mass of the di-lepton or lepton-neutrino system, transverse momenta of the Z boson, transverse mass of the W, multiplicity of jets from ISR parton shower and transverse momenta of the hardest jets reconstructed with **AcerDET** package. Perfect agreement on the most left plots confirms consistent starting point for the evolution of the ISR QCD parton shower. The differences observed on remaining plots should be attributed to the systematic theoretical uncertainties of the parton shower models.

References

[1] ATLAS Collaboration, *ATLAS Detector and Physics Performance TDR*, CERN-LHCC/99-15 (1999).

[2] CMS Collaboration, Technical Proposal, report CERN/LHCC/94-38 (1994).

[3] B. Kersevan and E. Richter-Was, Eur. Phys. J. **C39** (2005) 439.

[4] B. Kersevan and E. Richter-Was, *AcerMC version 2.4 with interfaces to PYTHIA 6.2, HERWIG 6.5 and ARIADNE 4.1*, hep-ph/0405247, available from `http://borut.home.cern.ch/borut/`

[5] B. Kersevan and E. Richter-Was, Comput. Phys. Commun. **149** (2003) 142.

[6] B. Kersevan and E. Richter-Was, *What is the $Wb\bar{b}$ background at LHC?...*, ATLAS Physics Note, ATL-PHYS-2003-018 (2001).

[7] A. Blondel, A. Clark and F. Mazzucato, ATLAS Physics Note, ATL-PHYS-2002-029 (2002).

[8] J. Gunion, Phys. Rev. Lett. **72** (1994) 199.

[9] N. Krauer and D. Zeppenfeld, Phys.Rev. **D65** (2002) 014021.

[10] E. Barberio and Z. Was, Comp. Phys. Commun. **79** (1994) 291.

[11] S. Jadach, J. H. Kuhn, Z. Was, Comput. Phys. Commun. **64** (1990) 275; M. Jezabek, Z. Was, S. Jadach, J. H. Kuhn, Comput. Phys. Commun. **70** (1992) 69; R. Decker, S. Jadach, J. H. Kuhn, Z. Was, Comput. Phys. Commun. **76** (1993) 361.

[12] L. Lönnblad, Computer Phys. Commun. **71** (1992) 15.
Manual for ARIADNE version 4 available with the distributed ARIADNE code

[13] LHAPDF documentation and code available from:
`http://hepforge.cedar.ac.uk/lhapdf/`

[14] E. Richter-Was, *AcerDET: A particlelevel fast simulation and reconstruction package for phenomenological studies on high p(T) physics at LHC*, hep-ph/0207355.

[15] E. Richter-Was, D. Froidevaux and L. Poggioli, ATLAS Internal Note ATL-PHYS-98-131 (1998).

RAPGAP

Hannes Jung
Deutsches Elektronen-Synchroton DESY Hamburg, FRG

RAPGAP originally developed to describe rapidity gap event in ep collisions has evolved into a multi-purpose Monte Carlo event generator for diffractive and non-diffractive processes at ep colliders both for high Q^2 and in the photoproduction regime ($Q^2 \sim 0$) as well as hard (single diffractive and non-diffractive) processes for pp and $p\bar{p}$ colliders. A detailed description of the program as well as the source code can be found under [1]. In the following only new developments are described.

1 NLO and Order α_s matrix elements

The $\mathcal{O}(\alpha_s)$ matrix elements for light quarks are divergent for $p_T^2 \to 0$, and usually a p_T^2 cutoff is applied. The $\overline{MS}$ factorization scheme provides a description which finite parts of the matrix elements are treated explicitely and which parts are included in the parton distribution functions. A consistent implementation of the NLO formalism for F_2 in DIS including initial state parton showering is described in detail in [2]. The LO (α_s^0) and the NLO (α_s) part are treated according the $\overline{MS}$ subtraction scheme, reformulated such that it properly can be used together with initial and final state parton showers, avoiding any double counting [3]. When using this scheme, the NLO parton densities calculated in the $\overline{MS}$ scheme should be selected. The program then transforms the parton densities from the $\overline{MS}$ to the BS scheme for parton showers. However, at present only the BGF part is implemented.

2 Les Houches interface

A generic format for the transfer of parton level event configurations from matrix element event generators (MEG) to showering and hadronization event generators (SHG) [4] is provided by the *Les Houches interface*. RAPGAP gives the possibility to write the full parton level events to the file `rapgap.gen`, which can be read in directly by the PYTHIA and HERWIG programs to perform the hadronization. This option is best suited to estimate the uncertainty coming from hadronization correction.

3 Proton dissociation ala DIFFVM

Dissociation of the proton according to the model in DIFFVM [5] can be included for diffractive events. The proton dissociation part of the cross section is given by

$$\frac{d\sigma}{dM_Y^2\, dt\, dx_{I\!P}} \sim \frac{1}{M_Y^{2(1+\epsilon_Y)}} \exp\left(-B_{diss}|t|\right)$$

with ϵ_Y describing the dependence on the dissociation mass M_Y and B_{diss} the t-dependence. The dissociative system Y is split into a $quark - gluon - diquark$ system for masses $M_Y > 2$ GeV whereas for masses $0.939 < M_Y < 2$ GeV the system is fragmented according to the nucleon resonances as implemented in DIFFVM [5].

4 Future Plans

In the next future it is planned to include double diffractive scattering for pp collisions to allow simulation of diffractive Higgs production.

References

[1] H. Jung, Comp. Phys. Comm. **86**, 147 (1995);
H. Jung, *The RAPGAP Monte Carlo for Deep Inelastic Scattering, version 3.1.* DESY Hamburg, 2005. `http://www.desy.de/~jung/rapgap/`.

[2] Collins, John C. and Zu, Xiao-Min, JHEP **06**, 018 (2002);
Collins, John C. and Zu, Xiaomin (2004).

[3] Schilling, S., *Implementation of bgf-processes in monte carlo generators for electron proton scattering*. DESY-THESIS-2000-040.

[4] Boos, E. and others, *Generic user process interface for event generators*, 2001.

[5] B. List, *Diffraktive J/ψ Produktion in Elektron - Proton Stößen am Speicherring HERA* (unpublished). Diploma thesis, Techn. Univ. Berlin, H1 note: H1-10/93-319, 1993;
List, B. and Mastroberardino, A., *Diffvm: A monte carlo generator for diffractive processes in ep scattering*, in *Proceedings of the Workshop on Monte Carlo generators for HERA physics*, eds. A. Doyle and G. Grindhammer and G. Ingelman and H. Jung, p. 396. DESY, Hamburg, 1999.

CASCADE

Hannes Jung
Deutsches Elektronen-Synchroton (DESY), Hamburg, FRG

CASCADE is a full hadron level Monte Carlo event generator for ep, γp, pp and $p\bar{p}$ processes, which uses the unintegrated parton distribution functions convoluted with off - mass shell matrix elements for the hard scattering. The CCFM [1] evolution equation is an appropriate description valid for both small and moderate x which describes parton emission in the initial state in an angular ordered region of phase space. For inclusive quantities it is equivalent to the BFKL and DGLAP evolution in the appropriate asymptotic limits. The angular ordering of the CCFM description makes it directly applicable for Monte Carlo implementation. The following processes are available: $\gamma^* g^* \to q\bar{q}(Q\bar{Q})$ $\gamma g^* \to J/\psi g$, $g^* g^* \to q\bar{q}(Q\bar{Q})$ and $g^* g^* \to h^0$.

A detailed description of CASCADE, the source code and manual can be found under [2]. A discussion of different unintegrated gluon densities can be found in [3, 4, 5].

The unintegrated gluon density $x\mathcal{A}_0(x, k_\perp, \bar{q})$ is a function of the longitudinal momentum fraction x the transverse momentum of the gluon $k_\perp$and the scale (related to the angle of the gluon) $\bar{q}$. Given this distribution, the generation of a full hadronic event is separated into three steps:

- The hard scattering process is generated,

$$\sigma = \int dk_{t\,1}^2 dk_{t\,2}^2 dx_1 dx_2 \mathcal{A}(x_1, k_{t\,1}, \bar{q}) \mathcal{A}(x_2, k_{t\,2}, \bar{q}) \sigma(g_1^* g_2^* \to X)\,, \tag{1}$$

  with X being $q\bar{q}$, $Q\bar{Q}$, $J/\psi g$ or h^0 states. The hard cross section is calculated using the off-shell matrix elements given in [6] for $q\bar{q}$ and $Q\bar{Q}$, $\gamma g^* \to J/\psi g$ in [7] and for Higgs production $g^* G^* \to h^0$ in [8]. The gluon momentum is given in Sudakov representation:

$$k = x_g p_p + \bar{x}_g p_e + k_t \simeq x_g p_p + k_t\,. \tag{2}$$

  where the last expression comes from the high energy approximation ($x_g \ll 1$), which then gives $-k^2 \simeq k_t^2$.
- The initial state cascade is generated according to CCFM in a backward evolution approach.
- The hadronization is performed using the Lund string fragmentation implemented in PYTHIA [9].

The backward evolution there faces one difficulty: The gluon virtuality enters in the hard scattering process and also influences the kinematics of the produced quarks and therefore the maximum angle allowed for any further emission in the initial state cascade. This virtuality is only known after the whole cascade has been generated, since it depends on the history of the gluon evolution (as $\bar{x}_g$ in eq.(2) may not be neglected for exact kinematics). In the evolution equations itself it does not enter, since there only the longitudinal energy fractions z and the transverse momenta are involved. This problem can only approximately be overcome by using $k^2 = k_t^2/(1 - x_g)$ for the virtuality which is correct in the case of no further gluon emission in the initial state. This problem is further discussed in [10, 5]

The CCFM evolution equations have been solved numerically [11] using a Monte Carlo method. Several sets of un-integrated gluon densities are available which have the input parameters were fitted to describe the structure function $F_2(x, Q^2)$ in the range $x < 5 \cdot 10^{-3}$ and $Q^2 > 4.5$ GeV^2 as measured at H1 [12] and ZEUS [13].

Also the unintegrated gluon densities described in [5] including non-linear terms [14] are available within CASCADE.

References

[1] M. Ciafaloni, Nucl. Phys. **B 296**, 49 (1988);
S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. **B 234**, 339 (1990);
S. Catani, F. Fiorani and G. Marchesini, Nucl. Phys. **B 336**, 18 (1990);
G. Marchesini, Nucl. Phys. **B 445**, 49 (1995).

[2] H. Jung, Comp. Phys. Comm. **143**, 100 (2002). `http://www.desy.de/~jung/cascade/`.

[3] Bo Andersson *et al.*, Eur. Phys. J. **C25**, 77 (2002).

[4] J. R. Andersen *et al.*, Eur. Phys. J. **C35**, 67 (2004).

[5] J. Collins, M. Diehl, H. Jung, L. Lönnblad, M. Lublinsky and T. Teubner, *Unintegrated parton density functions*. These proceedings, 2005.

[6] S. Catani, M. Ciafaloni and F. Hautmann, Nucl. Phys. **B 366**, 135 (1991).

[7] V. A. Saleev and N. P. Zotov, Mod. Phys. Lett. **A 9**, 151 (1994).

[8] F. Hautmann, Phys. Lett. **B535**, 159 (2002).

[9] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[10] J. Collins and H. Jung, *Need for fully unintegrated parton densities*, 2005.

[11] H. Jung and G. Salam, Eur. Phys. J. **C 19**, 351 (2001).

[12] S. Aid *et al.* [H1 Collaboration], Nucl. Phys. **B 470**, 3 (1996);
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. **C 21**, 33 (2001).

[13] M. Derrick *et al.* [ZEUS Collaboration], Z. Phys. **C72**, 399 (1996);
S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. **C21**, 443 (2001).

[14] K. Kutak and A. M. Stasto, Eur. Phys. J. **C41**, 343 (2005);
M. Lublinsky, *Parameterization of the dipole cross section and updf.*
`http://www.desy.de/~lublinm/`.

Leading proton production in ep and pp experiments: how well do high-energy physics Monte Carlo generators reproduce the data?

G. Bruni, G. Iacobucci, L. Rinaldi, M. Ruspa
INFN Bologna and University of Eastern Piedmont

Abstract
The simulation of leading-proton production at high-energy colliders as obtained by the HERWIG, LEPTO and PYTHIA Monte Carlo generators is analysed and compared to the measurements of HERA and fixed-target experiments. The discrepancies found between real data and Monte Carlo events could be responsible for inaccurate simulation of particle multiplicities and hadronic final states, which could eventually generate problems in computing the Standard-Model backgrounds to new physics at the LHC collider.

1 Introduction

The production of final state baryons carrying a large fraction of the available energy but a small transverse momentum (leading baryons) is crucial for a deep understanding of strong interactions beyond the perturbative expansion of QCD. Indeed, in high-energy collisions, the QCD-hardness scale decreases from the central, large p_T region, to the soft, non-perturbative hadronic scale of the target-fragmentation region. Therefore, the measurement of leading baryons in the final state of high-energy collisions allows to gather information on the non-perturbative side of strong interactions.

Another reason of interest in leading-baryon production comes from the fact that the energy carried away by the leading baryon(s) produced in a high-energy collision is not available for the production of the central-hadronic system. Therefore, the leading-baryon spectra should be well simulated for a proper accounting of the hadronic multiplicities and energies, e.g. at the LHC collider where an appropriate simulation of these quantities will be the ground for a reliable calculation of the Standard-Model backgrounds to new physics.

Here we will review the data on the production of leading protons and compare them to the most popular Monte Carlo generators available.

2 The data and the Monte Carlo generators used for the comparison

2.1 The proton-proton data

Although the experimental data on leading-proton production are scarce, a few measurements in a large x_L range are available, where x_L represents the fractional longitudinal momentum of the proton. In proton-proton collisions, leading-proton production has been studied both at the ISR [2, 3] and in fixed-target experiments [4, 5, 6]. The x_L spectra measured in fixed-target experiments are shown in Fig. 1a-c,e.

2.2 The ep data

Cross sections for the production of leading protons were also measured at the HERA collider [7, 8, 9]. More recently, the ZEUS Collaboration made a new measurement [10] of the cross-section for the semi-inclusive reaction $ep \to eXp$ in deep-inelastic scattering using $12.8\,\mathrm{pb}^{-1}$ of data collected during 1997. The single-differential cross sections, $d\sigma_{ep\to eXp}/dx_L$ and $d\sigma_{ep\to eXp}/p_T^2$, and the double-differential cross section, $d^2\sigma_{ep\to eXp}/dx_L dp_T^2$, were measured in the kinematic range $Q^2 > 3\,\mathrm{GeV}^2$ and $45 < W < 225$ GeV, where W is the total mass of the hadronic system. The protons were measured using the

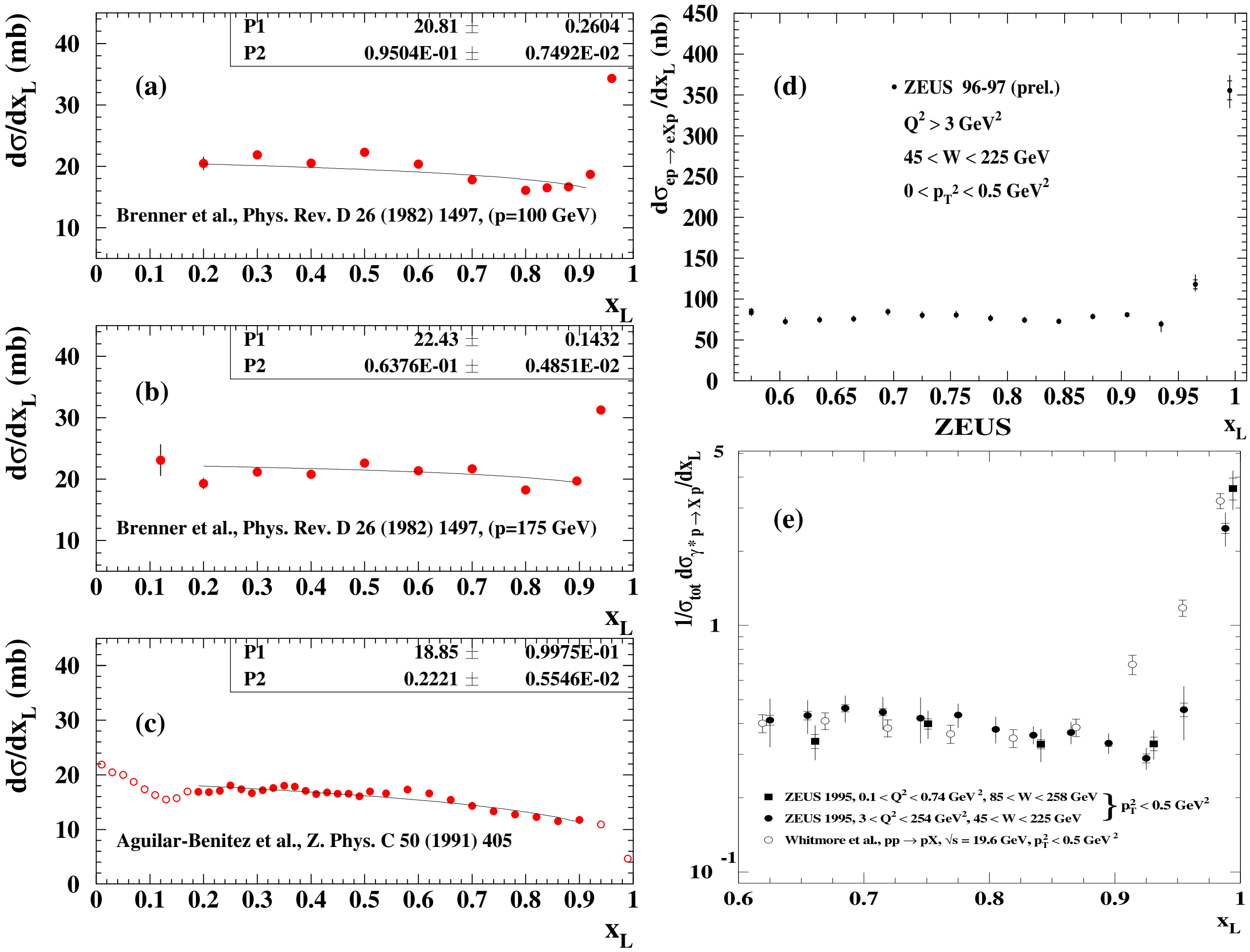


Fig. 1: Differential cross sections $d\sigma/dx_L$ measured in fixed-target experiments (a, b, c and e) and by the ZEUS Collaboration [10, 9] (d, e).

leading-proton spectrometer (LPS) [11] in the range $x_L > 0.56$ and $p_T^2 < 0.5$ GeV^2, where p_T is the scattered-proton transverse momentum.

2.3 The Monte Carlo generators

Large samples of Monte Carlo *ep* events were generated to be compared to the data. The LEPTO generator was used either with the MEPS or the ARIADNE packages; in the latter case the diffractive component of the cross section was simulated using the Soft Color Interaction model. Events were also generated with HERWIG. Since this Monte Carlo does not simulate diffractive events, the POMWIG generator was used to account for the single diffractive events, and the SANG generator to account for the diffractive events in which the scattered-proton dissociates in a higher-mass hadronic system.

pp events at the LHC center of mass energy (14 TeV) were generated with PYTHIA.

3 Discussion

3.1 The x_L spectrum

Figure 1a, b and c show the $d\sigma/dx_L$ obtained by the fixed target experiments [4, 5, 6] which measured leading protons in a wide range of x_L. The cross section for such events shows a peak for values of the

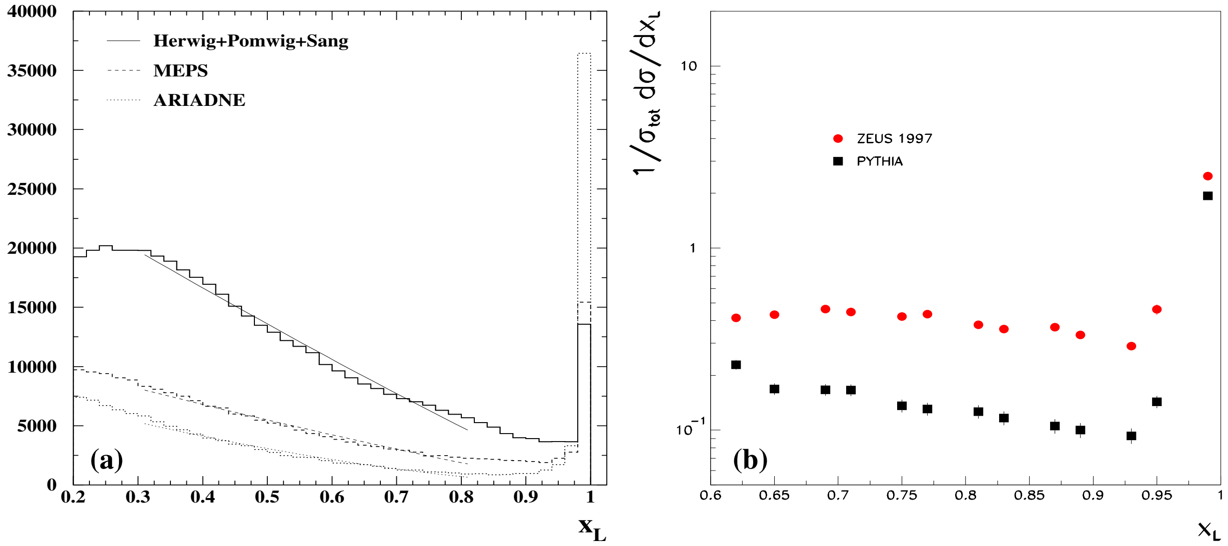


Fig. 2: (a) Differential cross section $d\sigma_{ep\to eXp}/dx_L$ generated with LEPTO, with MEPS or ARIADNE, and with HERWIG+POMWIG+SANG Monte Carlos; (b) Comparison between the normalised differential cross section $1/\sigma_{tot}\ d\sigma_{pp\to Xp}/dx_L$ simulated with PYTHIA and $1/\sigma_{tot}\ d\sigma_{ep\to eXp}/dx_L$ measured by the ZEUS Collaboration [9].

final-state proton momentum close to the maximum kinematically allowed value, the so-called diffractive peak. Below the diffractive peak the cross section is lower and consistent with a flat one. In this region, under the assumption of Regge factorisation [12], the fraction of events with a leading proton is expected to be approximately independent of the energy and type of the incoming hadron. The lines superimposed to the data are the results of fits in the range $0.1 < x_L < 0.9$ to the function $(1-x_L)^\alpha$ that is commonly used to characterise the longitudinal distributions of leading particles. The values of α obtained are 0.1, 0.06 and 0.22 respectively for Fig. 1a, b and c. Fig. 1d shows the preliminary $d\sigma/dx_L$ obtained by ZEUS. Below the diffractive peak the cross section is again consistent with a flat one, i.e. $\alpha \sim 0$. A comparison between the normalised cross section $1/\sigma_{tot}\ d\sigma/dx_L$ obtained by the fixed-target data [6] and by the ep data is shown in Fig. 1e. For $x_L < 0.9$ the fraction of events with a leading proton is indeed consistent for the pp and the ep data set.

The x_L distributions of the simulation of the HERA events are shown in Fig. 2a. Already at first glance, the difference w.r.t. the data is evident, since the spectra are much more populated at low x_L in the Monte Carlo than they are in the data. Indeed, the fits to the same functional form as for the data give α = 1.0 for LEPTO-MEPS, α = 1.4 for LEPTO-ARIADNE and α = 1.0 for HERWIG+POMWIG+SANG.

In Fig. 2b the x_L distribution obtained from the simulation with PYTHIA of pp events at the LHC center of mass energy (14 TeV) is compared to the ZEUS data. As discussed previously, according to the vertex factorisation hypothesis, the fraction of events with a leading proton is expected to be consistent in the ep and in the pp case. The simulation appears to approximately agree with the data in the diffractive peak region but is not able to describe the data neither in shape nor in normalisation in the region outside the diffractive peak.

In general we conclude that the fraction of beam energy carried away by the leading proton in the Monte Carlo is on average much smaller than in the data, with the consequence that the energy available in the simulation for the production of the central-hadronic system is correspondingly larger than in nature.

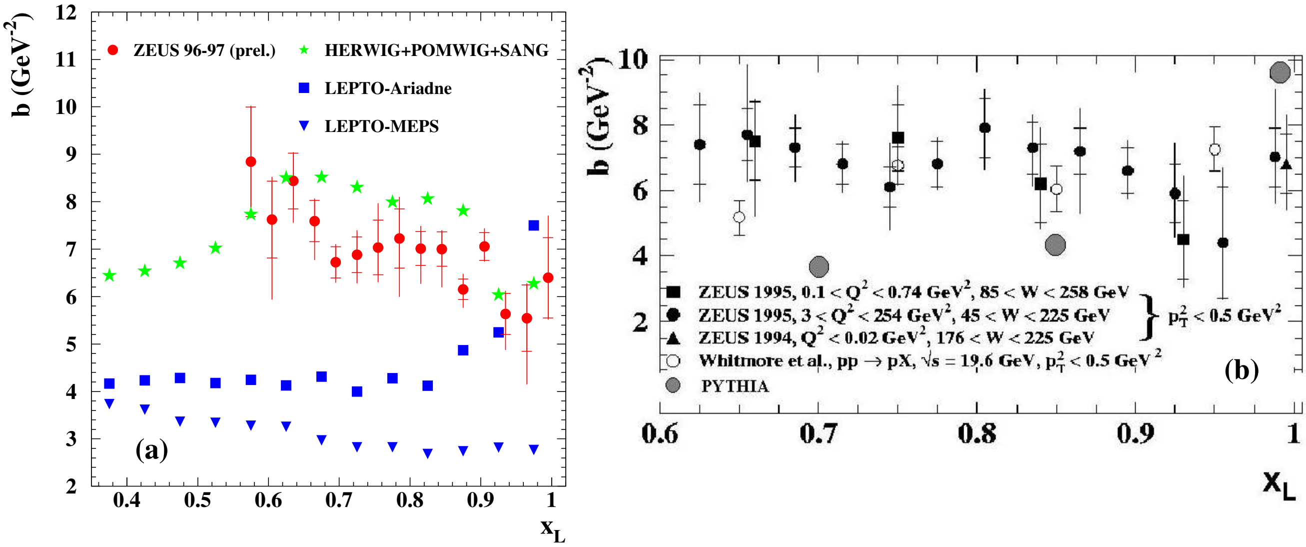


Fig. 3: The slope-parameters b obtained from a single-exponential fit to the differential cross section $d\sigma_{ep\rightarrow eXp}/dp_T^2$ ($d\sigma_{pp\rightarrow Xp}/dp_T^2$ for PYTHIA pp sample) to the function $A \cdot e^{-b \cdot p_T^2}$ in each of the x_L bins shown. (a) The red dots are the ZEUS preliminary data [10], while the other symbols represent the results of the Monte Carlos described in the picture; (b) The grey dots are the pp events simulated with PYTHIA, while the other symbols are the data described in the picture.

3.2 The p_T^2 spectrum

Although the p_T^2 distribution of the leading proton is less important for the hadronic final states than the x_L distribution, it is interesting to investigate how well the generators can reproduce it.

In Fig. 3a the red dots show the values of the slope-parameter b obtained from a single-exponential fit to the function $e^{-b \cdot p_T^2}$ in each of the x_L bins of the ZEUS $d\sigma/dp_T^2$ measurement. The b slopes obtained by a similar fit performed on the simulated events in the ep case and in the pp case are also reported in Fig. 3a and Fig. 3b, respectively. In the pp case the extracted b-slopes have been corrected for the expected shrinkage of the diffractive peak.

The b-slopes values resulting from the fit to the p_T^2 distribution of the HERWIG+POMWIG+SANG sample appear to be in the right ball park.

The LEPTO generator shows too small b-slope values, smaller than those of the data by approximately 3 GeV^{-2}. In the case the matrix-element parton showers are used to generate the events, since the dependence of b on x_L is similar to that of the data, it is conceivable to fix the difference by tuning the primordial k_T of the generation. If the ARIADNE package is used instead, it seems quite difficult to improve the situation in a similar way, since the generated b values increase with x_L, a feature that is not seen in the data.

The b-slopes values resulting from the fit to the p_T^2 distribution of the PYTHIA pp sample are approximately consistent with the ZEUS data in the diffractive peak region, but lower than the data again by approximately 3 GeV^{-2} in the region outside the peak.

3.3 Reweighting of the PYTHIA leading proton spectrum

A sample of pp proton events generated with PYTHIA has been used to simulate the many interactions per bunch crossing (pile-up events) occurring at the LHC luminosities in a recent study on the diffractive production of a Higgs boson at the LHC [13]. The simulated leading proton spectrum has been reweighted both in x_L and in p_T^2 with the following function, calculated for each x_L bin of Fig. 2b:

$$f(x_L) = [-37.22 + 135.1 \cdot x_L - 148.5 \cdot x_L^2 + 54.3 \cdot x_L^3] \cdot \frac{e^{-b_{ZEUS}t}}{e^{-b_{PYTHIA}t}} \cdot \frac{b_{ZEUS}}{b_{PYTHIA}},$$

where $b_{ZEUS} = 7$ GeV^{-2} and $b_{PYTHIA} = 4.4$ GeV^{-2}. The polynomial form in x_L is the result of a fit to the ratio ZEUS/PYTHIA of the differential cross sections $1/\sigma_{tot}\ d\sigma/dx_L$ of Fig. 2b; thus $f(x_L)$ provides the number of simulated leading protons to be consistent with the ZEUS data.

3.4 The fraction of diffractive –large-rapidity gap– events w.r.t. the total

One way to identify a diffractive event produced in ep or pp interactions is to search for a large-rapidity gap (LRG) in the pseudorapidity distribution of the particles produced. In ZEUS, a diffractive-LRG event was tagged by $\eta_{max} < 2$, where η_{max} corresponds to the pseudorapidity of the most forward (i.e. proton direction) energy deposit in the calorimeter exceeding 400 MeV. The η_{max} distribution for ZEUS DIS events with $Q^2 > 3$ GeV^2 is shown by the dots in Fig. 4. The two regions of non-diffractive events (with η_{max} between 2 and 8) and of diffractive events (which distribute at η_{max} values below 2) are clearly distinguishable.

The LEPTO-MEPS events were passed through the standard simulation of the ZEUS trigger and detector, and through the same reconstruction and analysis programs as the data. The η_{max} distribution of the MC events after such processing is shown by the dashed histogram in Fig. 4. We note that the diffractive events with $\eta_{max} < 2$ generated with the Soft Color Interaction algorithm in LEPTO-MEPS are more than twice those found in the data. Therefore, the Soft Color Interaction algorithm, as implemented in the LEPTO generator, fails to describe the data in the range of the ZEUS-LPS detector, i.e. $x_L > 0.56$.

4 Summary

The data on leading-proton production in ep and pp scattering have been compared to the most popular Monte Carlo generators available to simulate high-energy physics events. This exercise has revealed that the simulation of the leading-proton momenta, both longitudinal and transverse to the beams, does not reproduce the properties of the data. In particular, the x_L distribution of the leading protons would be made more close to that of the data if a proper accounting of the energy available for the hard-scattering process could be achieved.

Although the HERWIG generator has been successful in simulating many features of high-energy physics final states, it does not contain the diffractive component of the cross section, and the x_L spectrum it produces is far from being almost flat, as seen in the data.

The Soft Color Interaction model in its standard implementation in LEPTO is producing twice the fraction of diffractive-LRG events seen in the ZEUS data at $x_L > 0.56$, therefore distorting in a significant way the multiplicities and hadronic energies present in real events.

The PYTHIA Monte Carlo has been used to simulate pp events at the LHC center of mass energy (14 TeV), and then also compared to ep data. The generator has been shown to reproduce the longitudinal and transverse momentum of the data in the diffractive peak region; however, it underestimates both the cross section and the p_T^2-slopes at lower values of the scattered proton momentum, contradicting the hypothesis of vertex factorisation, which is supported by the data.

All the above arguments generate some concern that the hadronic multiplicities of the MC generators taken into account here have been tuned consistently and that they can produce an accurate simulation of the final states of the Standard-Model processes at the LHC energies.

5 Acknowledgments

We would like to thank Torbjörn Sjöstrand for useful discussions.

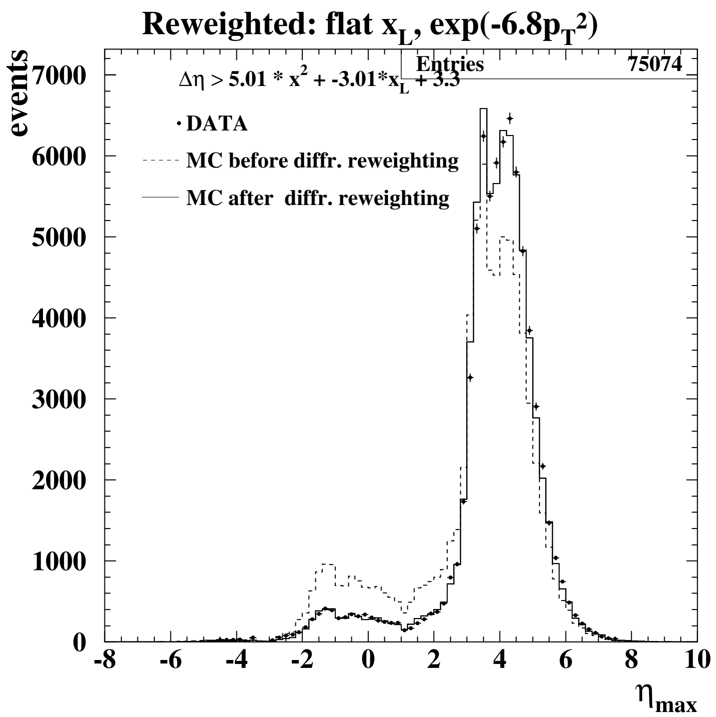


Fig. 4: The η_{max} distribution of the ZEUS data [10] (dots) and of the LEPTO-MEPS Monte Carlo obtained with the default x_L and p_T^2 distributions (dashed-line histogram). The MC distribution after reweighting to a flat x_L and to $e^{-6.8p_T^2}$ is shown by the full-line histogram.

References

[1] M. Basile et al., Nuovo Cimento 66A (1981) 129.

[2] M. Basile et al., Lett. Nuovo Cimento 32 (1981) 321.

[3] V. N. Gribov et al., "The creation of QCD and the effective energy", World Scientific Series in 20th Century Physics 25 (2001).

[4] M. Aguilar-Benitez et al., Z. Phys. C 50 (1991) 405.

[5] A. E. Brenner et al., Phys. Rev. D 26 (1982) 1497.

[6] J. Withmore et al., Phys. Rev. D 11 (1975) 3124.

[7] H1 Collab., C. Adloff et al., Nucl. Phys. B 619 (2001) 3.

[8] H1 Collab., C. Adloff et al., Eur. Phys. J. C 6 (1999) 587.

[9] ZEUS Collab., S. Chekanov et al., Nucl. Phys. B 658 (2002) 3.

[10] ZEUS Collab., S. Chekanov et al., Proceedings to the International Europhysics Conference on High Energy Physics (EPS03), Abstract 544.

[11] ZEUS Collab., M. Derrick et al., Z. Phys. C 3 (1997) 253.

[12] See e.g.: K. Goulianos and references therein, Phys. Rep. 102 (1983) 169.

[13] M. Grothe et al., "Diffractive Higgs: CMS/TOTEM Level-1 Trigger Studies", these proceedings.

[14] M. Basile et al., Nuovo Cimento 79A (1984) 1.

NLOLIB - A Common Interface for Fixed-Order Calculations

Klaus Rabbertz[†] *and Thomas Schörner-Sadenius*[o]
[†]University of Karlsruhe, EKP, Postfach 6980, D-76128 Karlsruhe, *klaus.rabbertz@cern.ch*
[o]University of Hamburg, IExpPh, Luruper Chaussee 149, D-22761 Hamburg, *schorner@mail.desy.de*

Abstract
We present the current status of the NLOLIB framework which provides an interface to various higher order perturbative calculations, thus allowing for simple comparisons of these calculations with each other and with measured distributions. We show, as a newly included example of the NLOLIB abilities, a comparison of calculations for jet production in deeply inelastic ep scattering.

1 Introduction

Progress in particle physics relies, to a large extent, on the comparison of data to theoretical predictions. Most commonly, the theoretical calculations are available to the experiments in the form of Monte Carlo event generators producing event records that contain all generated particles, their four-momenta, the decay trees etc. in a commonly adopted format. The events that represent the outcome of such a prediction can be used directly by experiment-specific detector simulation and analysis software in order to perform detailed comparisons with experimental data.

Due to complications in the involved mathematical techniques most programs providing higher order electroweak or QCD calculations, however, require large numbers of events to be generated with positive and negative weights. These can not be easily used in simulations because the meaning of the cancellation of positive and negative weights in combination with detector influences is not overly clear. In addition, producing the necessary numbers of events is extremely time-consuming. Nevertheless the results are very useful for comparisons to data and they are used in a variety of experiments to perform, for example, precision measurements of standard model parameters like the strong coupling parameter, α_s.

Although many such programs have been developed for a variety of physical processes, and sometimes even more than one program exist for the same purpose, the usage and the presentation of the results has not been unified so far. This leads to a number of unnecessary technical problems for the user who wants to compare the predictions of more than one program to some measured distribution. To be more specific, the user has to learn, for each single program he or she wants to use, how to install and compile the program, how to implement the specific distributions or processes of interest, and how to extract the results. It should be noted that the programs are also implemented in different programming languages, mostly *FORTRAN* and *C++*.

NLOLIB seeks to simplify the physicists' life by unifying the steering and providing a common framework to implement new quantities and to extract the results. A first version of NLOLIB was developed in the workshop on 'Monte Carlo Generators for HERA Physics' 1998/99 and aimed to integrate and compare three different programs for next-to-leading order calculations in electron-proton scattering [1]. Now this scope will be extended to include also proton-(anti)proton and electron-positron collisions.

2 Implemented programs

RacoonWW provides tree-level cross-sections to all processes where an electron-positron collision yields four fermions in the final state. In addition, it contains a number of higher order corrections, see [2] for details.

DISENT [3] is a next-to-leading order calculation for the production of one or two jets (processes up to $\mathcal{O}(\alpha_s^2)$) for deep-inelastic ep scattering. It uses a subtraction scheme [4] for the cancellation of divergencies. DISENT is the standard program for DIS jet production at HERA and has been used in a variety of analyses.

DISASTER++ [5] offers more possibilities to separate different terms in the derivation of the cross sections; otherwise it provides a functionality similar to DISENT. It also employs the subtraction method.

JetViP [6, 7] is a next-to-leading order calculation for jet production in deep-inelastic ep scattering and $\mathrm{e}^+\mathrm{e}^-$ collisions that contains processes up to $\mathcal{O}(\alpha_s^2)$ and implements both direct and resolved contributions to the cross-section. The cancellation of divergencies is performed using the phase-space slicing method [8] which leads to dependencies of the resolved contribution on the unphysical slicing parameter, y_{cut}. The direct predictions of JetViP have been shown to be compatible [9, 7] with those of DISENT and DISASTER++. So far, only the implementation into NLOLIB of the direct photon ep scattering part of JetViP has been thoroughly tested; the resolved photon part is implemented in principal but needs more testing. The implementation of the $\mathrm{e}^+\mathrm{e}^-$ part into NLOLIB has just started.

MEPJET [10] was the first complete next-to-leading order calculation available for deep-inelastic ep scattering and is based on the phase-space slicing method in combination with the technique of crossing functions [11]. The predictions of MEPJET show some discrepancies with respect to the results of DISASTER++, DISENT and JetViP [9].

NLOJET++ [12] incorporates next-to-leading order predictions for ep, pp and e+e- scattering using the subtraction method. Due to its very different way of allowing users to implement their favourite quantities, an integration into NLOLIB on the same footing as for the other programs seems not to be feasible. However, it will be tried to achieve an approach as similar as possible. Currently, only the original version of NLOJET++ in its unchanged form is included.

FMNR [13]: FMNR is a program for the calculation of next-To-leading order photoproduction jet cross-sections with heavy quarks in the final state. The implementation of FMNR into NLOLIB has only just begun.

3 Getting started

Once, a new release is finished, a compressed tar archive will be made available like it is done already now on `http://www.desy.de/~nlolib`.

Since the way in which NLOLIB is installed has changed considerably in the course of this workshop from the original version [1], some short instructions on how to get started are given here.

Originally, the `make` tool together with a set of `perl` scripts containing hardware-specific settings were used. However, this procedure was not easily maintainable, so it was decided to employ the GNU `autotools` [14]. For this to work `automake` versions 1.7 or higher and `autoconf` versions 2.57 or higher are needed. In addition, the CERN libraries including a version of PDFLIB [15] and the HzTool [16] libraries as available from our web page are required.

When these conditions are met, the following scheme should be followed for installing NLOLIB:

- Retrieve the NLOLIB source code from the web page at DESY (at a later stage it will also be downloadable from the CVS server in Karlsruhe) and copy it to your working directory (assumed to be `~/nlolib` in the following).
- In `nlolib.sh` (for c-type shells in `nlolib.csh`) set the correct paths to the PDFLIB and the CERNLIB libraries (`libpdflib.a` or `libpdflib804.a`, `libkernlib.a`, `libpacklib.a`), the HzTool libraries (`libhztool.a`, `libmyhztool.a`) and the directories where to put the NLOLIB binaries and libraries, typically `~/nlolib/bin` and `~/nlolib/lib`.
- Go to the working directory and source `nlolib.sh` (or `nlolib.csh`):
  $\sim \backslash nlolib >$ `source nlolib.sh`

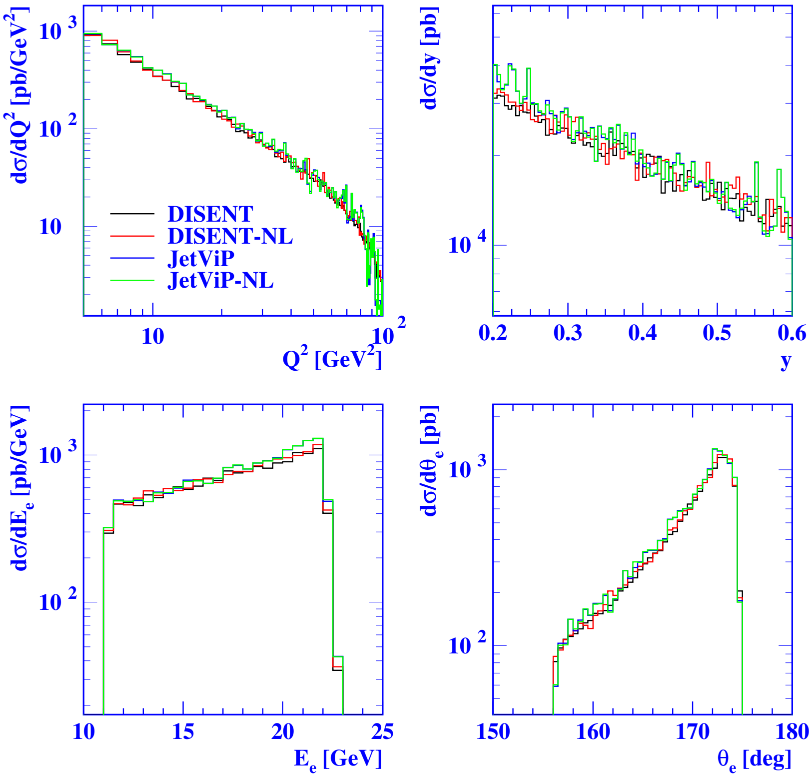


Fig. 1: Cross-sections as functions of Q^2, y, E_e and θ_e for the different DISENT and JetViP predictions.

- Usually, the following three steps can be skipped. But in case the `configure` script below fails, it has to be recreated by doing:
  - $\sim \backslash nlolib >$ `aclocal`
  - $\sim \backslash nlolib >$ `automake`
  - $\sim \backslash nlolib >$ `autoconf`
- $\sim \backslash nlolib >$ `./configure`
- This step can be skipped if a complete recompile is not necessary/wanted:
  $\sim \backslash nlolib >$ `make clean`
- $\sim \backslash nlolib >$ `make`
- $\sim \backslash nlolib >$ `make install`

Since the running of NLOLIB depends on the simulation program required by the user no general rules can be given here concerning the NLOLIB execution.

4 Jet cross-sections in ep NC DIS

As a new check of the NLOLIB framework we tested the predictions of the DISENT and JetViP programs for deep-inelastic ep scattering single-inclusive jet production against the stand-alone versions of the programs and against data published by the H1 collaboration [17]. The phase-space of the measurement is determined by two requirements on the scattered electron: the energy of the scattered electron E' must be larger than 10 GeV, and its polar angle must be larger than 156^o. In addition, two kinematic cuts are applied to select well-reconstructed low-Q^2 DIS events: $5 < Q^2 < 100$ GeV^2 and $0.2 < y < 0.6$.

Jet reconstruction for the selected events is performed in the Breit reference frame with the lon-

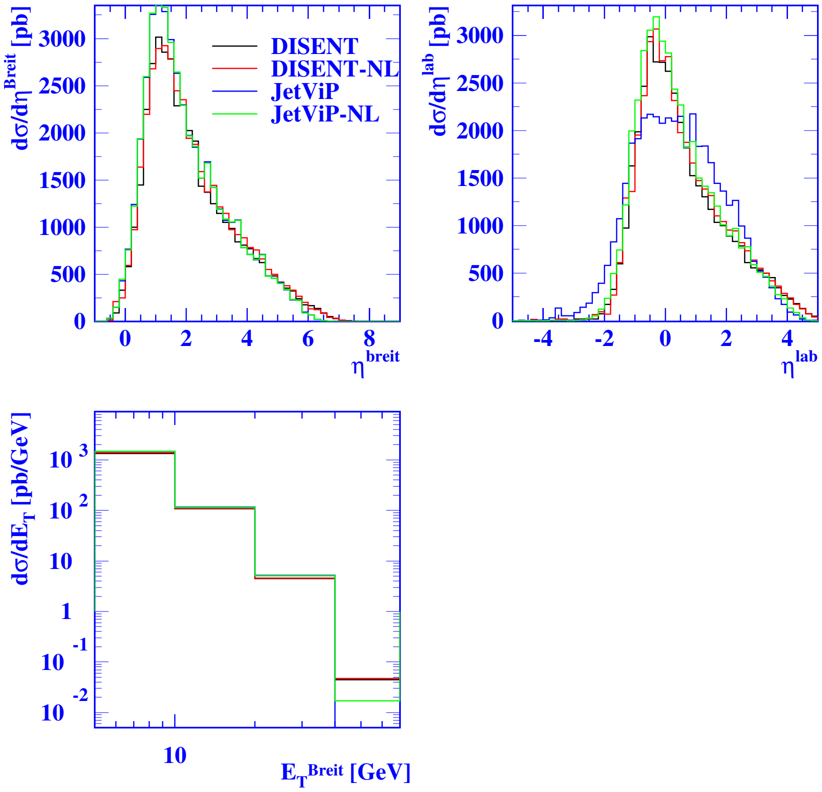


Fig. 2: Inclusive jet cross-sections as functions of η^{Breit}, η^{lab} and E_T^{Breit} for the different DISENT and JetViP predictions.

gitudinally invariant $k_\perp$ algorithm [18] in the inclusive mode [19]. Jets are selected by requiring their transverse energy in the Breit frame to be larger than 5 GeV, $E_T^{breit} > 5$ GeV, and their pseudorapidity in the laboratory frame, η^{lab}, to be between -1 and 2.5.

4.1 Comparison of calculations and data

We first present a comparison of event and jet quantities between the stand-alone versions and the versions implemented in NLOLIB of DISENT and JetViP. For this purpose 1 million events have been generated with each of the four programs using CTEQ4M as proton PDFs and Q^2 as renormalization and factorization scale. Figure 1 shows the differential cross-sections as functions of Q^2, y, E_e and θ_e for the four predictions. A very good agreement between the predictions is observed.

Also the comparison of the various predictions for the jet pseudorapidities in the Breit and laboratory reference frames, η^{Breit} and η^{lab}, and for the jet transverse energy in the Breit frame, E_T^{Breit}, shows satisfactory agreement. There are, however, small discrepancies between the two JetViP and the two DISENT predictions in the η^{Breit} distribution and a rather large discrepancy between the stand-alone JetViP prediction on the one hand side and the other three calculations on the other hand side for η^{lab}, see figure 2.

Figure 3 finally compares the four predictions to the published H1 data which are presented as inclusive jet cross-sections as functions of E_T^{Breit} in different ranges of η^{lab}. Also for these published observables the agreement of the various predictions is reasonable.

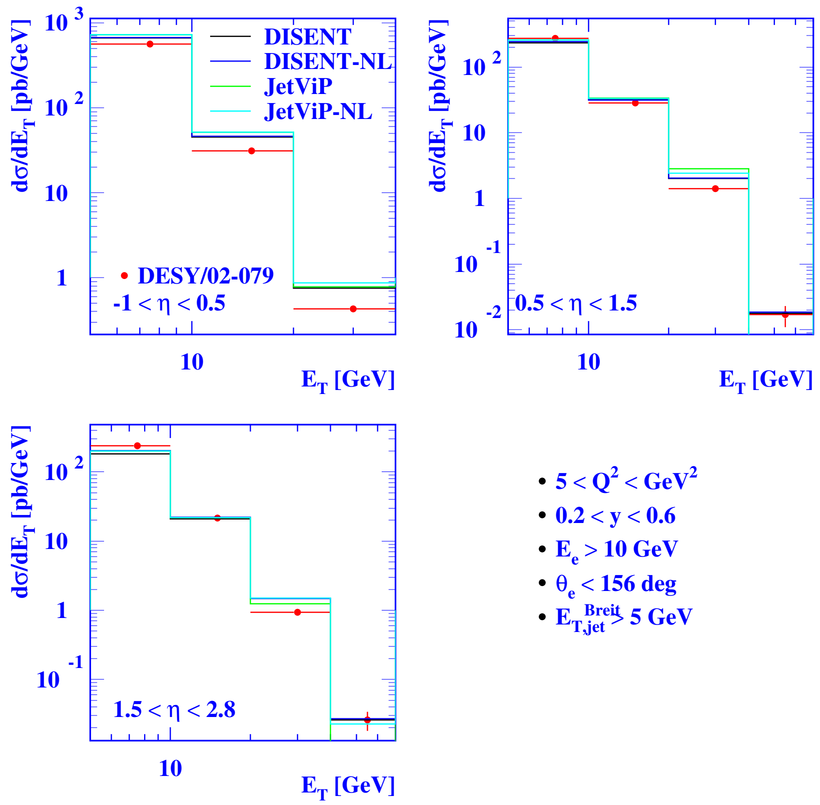


Fig. 3: Inclusive jet cross-sections as functions of E_T^{Breit} in different ranges of η^{lab}. The H1 data are compared to the different DISENT and JetViP predictions.

4.2 How to obtain the theoretical distributions

The NLOLIB calculations shown in this section have been obtained by running DISENT and JetViP via the HzTool interface in NLOLIB and using the HzTool routine for the H1 data analysis, `hz02079.f`. The steering files for the DISENT and JetViP job can be found in `nlolib/steering` and are called `dis02079.t` for DISENT and `jv02079.t` for JetViP. The command to run the DISENT job is thus (assuming the command is issued in the `nlolib` directory)

$\sim \backslash nlolib >$ `bin/hzttol < steering/dis02079.t`

for JetViP the command is

$\sim \backslash nlolib >$ `bin/hzttol < steering/jv02079.t`

In both cases a HBOOK file `test.hbook` is created that contains, in subdirectory 02079, the results of the calculation and the published H1 data points. A *PAW* macro `epjets.kumac` that creates the plots shown here will soon be available.

5 Summary

Some results with recently implemented higher order calculations have been shown, but clearly many items on our agenda unfortunately are still to be done.

For the implementation of JetViP into NLOLIB first, and most importantly, the e^+e^- mode has to be implemented and tested – so far only the ep mode has been done. Secondly, the resolved photon contribution has to be tested more thoroughly and the discrepancies between DISENT and JetViP in the pseudorapidity distributions need to be sorted out.

Concerning NLOJET++ a similar approach like for the other programs has to be set up and thoroughly tested in an all-program comparison, for example of event shapes or jet cross-sections in deep-inelastic scattering. Then, jet cross-sections in hadron-hadron collisions can be derived with NLOJET++. In addition, the use of PDFLIB will be replaced by LHAPDF [20].

Finally, the work on the implementation of more programs, for example FMNR or further proton-(anti)proton programs, needs to be followed up.

References

[1] T. Hadig and G. McCance, Proceedings of the *Workshop on Monte Carlo Generators for HERA Physics*, eds. G. Grindhammer, G. Ingelman, H. Jung and T. Doyle, DESY, Hamburg (1998), 125.

[2] A. Denner *et al.*, Phys. Lett. B**475** (2000) 127.

[3] S. Catani and M. Seymour, Nucl. Phys. B**485** (1997), 291; Erratum ibid. B**510** (1998) 503.

[4] S. Catani and M. Seymour, Proceedings of the Workshop *Future Physics at HERA*, eds. G. Ingelman, A. DeRoeck and R. Klanner, DESY, Hamburg (1996) 519.

[5] D. Graudenz, arXiv:hep-ph/9710244 (1997).

[6] B. Pötter, Comput. Phys. Commun. **119** (1999) 45.

[7] B. Pötter, Comput. Phys. Commun. **133** (2000) 105.

[8] M. Klasen and G. Kramer, Z. Phys. C**72** (1996)107;
M. Klasen, DESY report DESY-96-204;
M. Klasen, T. Kleinwort and G. Kramer, EPJdirect C**1** (1998) 1.

[9] C. Duprel *et al.*, Proceedings of the *Workshop on Monte Carlo Generators for HERA Physics*, eds. G. Grindhammer, G. Ingelman, H. Jung and T. Doyle, DESY, Hamburg (1998), 142.

[10] E. Mirkes and D. Zeppenfeld, Phys. Lett. B**380** (1996) 205;
E. Mirkes, preprint TTP-97-39 (1997), arXiv:hep-ph/9711224.

[11] W.T. Giele, E.W.N. Glover and D.A. Kosover, Nucl. Phys. B**403** (1993) 633.

[12] Z. Nagy and Z. Trocsanyi, Phys. Rev. Lett. **87** (2001) 082001.

[13] S. Frixione *et al.*, Phys. Lett. B**348** (1995) 633;
S. Frixione *et al.*, Nucl. Phys. B**454** (1995) 3.

[14] http://www.gnu.org.

[15] H. Plothow-Besch, CERN-ETT/TT 2000.04.17 (CERNLIB write-up W5051).

[16] J. Bromley *et al.*, Proceedings of the Workshop *Future Physics at HERA*, eds. G. Ingelman, A. DeRoeck and R. Klanner, DESY, Hamburg (1996) 611; http://www.desy.de/∼carli/hztool.html.

[17] H1 Collaboration, C. Adloff *et al.*, Phys. Lett. B**542** (2002) 193.

[18] S. Catani *et al.*, Nucl. Phys. B**406** (1993) 187.

[19] S.D. Ellis and D.E. Soper, Phys. Rev. D**48** (1993) 3160.

[20] M. Whalley, http://hepforge.cedar.ac.uk/lhapdf/.

.

HZTool

J. M. Butterworth [1], *H. Jung* [2], *V. Lendermann* [3], *B. M. Waugh* [1]
[1] Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, England
[2] Deutsches Elektronen-Synchrotron, Notkestr. 85, 22607 Hamburg, Germany;
[3] Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany

Abstract
HZTOOL is a library of Fortran routines used for tuning and validating of Monte Carlo and analytical models of high energy particle collisions. The library includes an extensive collection of routines reproducing ep, γp, $p\bar{p}$ and $\gamma\gamma$ data published by the HEP experiments, as well as calculations as they were performed in the experimental data analyses. During this workshop the library was further developed to include HERA and TeVatron results relevant for the studies at the LHC.

1 Introduction

Data from high-energy physics experiments have seen the triumph of the Standard Model both in precision electroweak measurements and in the verification of QCD to a reasonable degree of precision. However, a number of aspects of high energy collisions remain poorly understood due to technical difficulties in the calculation. This is particularly the case for measurements of the hadronic final state in high energy collisions, where the specific event shapes variables, jet algorithms and kinematic cuts may be rather complex.

Accurate models of the final state are often needed to design new experiments and to interpret the data from them. Simulation programs employing fits to existing data address these problems. However, consistent tuning of the parameters of these programs, and examination of the physics assumptions they contain, is non-trivial due to the wide variety of colliding beams, regions of phase space, and complex observables. Comparing a new calculation to a sensible set of relevant data is in practice extremely time consuming and prone to error.

HZTOOL [1] is created to improve this situation. It is a library of Fortran routines allowing reproduction of the experimental distributions and easy access to the published data. Basically, each subroutine corresponds to a published paper. If supplied with the final states of a set of simulated collisions, these routines will perform the analysis of the final state exactly as it was performed in the paper, providing simulated data points which may be compared to the measurement. HZTOOL currently contains measurements from ep, γp, $\gamma\gamma$ and $p\bar{p}$ collisions. Others may easily be added.

While it is designed to be used as simply as possible as a standalone library, HZTOOL is also a key component of JETWEB [2]. JETWEB is a Web-based facility for tuning and validating Monte Carlo models. A relational database of reaction data and predictions from different models is accessed via the Web, enabling a user to find out how well a given model agrees with a range of data and to request predictions using parameter sets that are not yet in the database.

HZTOOL can also be used together with RUNMC [3], which is an object-oriented frond-end for Monte Carlo programs incorporating a sophisticated graphical user interface.

The library was initially developed within the workshop "Future Physics at HERA" [4] but has since expanded to become a more general toolkit. The package was managed by *T. Carli* for quite some time and many people have contributed routines and general development to HZTOOL since it first appeared. Nowadays, HZTOOL and JETWEB are further developed within the CEDAR project [5]. The current maintainers are *J. Butterworth*, *H. Jung*, *E. Nurse* and *B. Waugh*[1]. It is planned to design a C++ equivalent, in order to provide a native interface to the new C++ Monte Carlo programs, to enable a straightforward implementation of new HEP data analyses performed in C++.[2]

2 HZTOOL Usage

Each analysis subroutine books, fills and outputs two sets of histograms: one reproducing the published data and another one filled by the chosen simulation program. The routine names relate to the publication. The preferred convention is:

`HZHyymmnnn` where yymmnnn is the arXiv:hep–ex preprint number.

Alternative naming schemes, used for older routines or when a hep–ex number is not available, are:

`HZDyynnn` where yynnn is the DESY preprint number.
`HZCyynnn` where yynnn is the CERN preprint number.
`HZFyynnnE` where yynnn is the FNAL preprint number.
`HZyynnn` where yynnn is the DESY preprint number[3].

If none of the above number schemes exist, the routine name is generally derived from the journal publication (e.g. `HZPRT154247`). Occasionally a single publication contains results taken under more than one set of beam conditions, in which case there will be a routine for each beam condition, distinguished by appending a letter to the expected name (e.g. `HZC88172A`, `HZC88172B`).

HZTOOL is a library, and the main program, which is usually a Monte Carlo event generator, must be provided by the user. The code of HZTOOL routines is basically independent of the program used to simulate the collisions. The Monte Carlo generators currently supported are: ARIADNE [6], CASCADE [7], HERWIG [8] (including JIMMY [9]), LEPTO [10], PYTHIA [11], PHOJET [12], QCDINS [13], RAPGAP [14], RIDI [15], and DJANGOH [19]. The production versions of these programs are all currently written in Fortran. The library can also be accessed within the NLOLIB framework for running NLO QCD programs [16]. Besides the HZTOOL library it is also necessary to link in the CERNLIB library routines and possibly PDFLIB [17] or LHAPDF [18]. Examples of the main programs can be found in the HZSTEER package [20] which provides the executable programs for JETWEB to submit from its backend.

To ease the implementation of the analysis code, HZTOOL provides the relevant jet finders (the cluster and cone algorithms with various options), as well as a number of utilities to calculate event shape variables, to perform Lorentz boosts etc.

3 Recent Developments

Within this workshop, an effort was made to include all results from HERA and other HEP experiments which can be helpful for tuning of MC models used for event simulations at the LHC. In particular, the models for multiple interactions and for heavy flavour production were considered.

The publications which may be relevant for the tuning of multiple interaction models and which are available in HZTOOL are listed in [21]. The names for the corresponding HZTOOL subroutines are also specified. From this list, the newly written routines are: `HZH9505001` (*J. M. Butterworth*, *B. M. Waugh*), `HZH9810020` (*S. Lausberg*, *V. Lendermann*), `HZH0006017` (*D. Beneckenstein*, *V. Lendermann*) and `HZH0302034` (*K. Lohwasser*, *V. Lendermann*). The routine `HZ95219` was extended by *A. Buniatian* to include the results from the corresponding H1 paper which are especially sensitive to underlying events in the photoproduction of jets. The models of multiple interactions and efforts of their tuning are reviewed

[1]The maintainers can be contacted at `hztool@cedar.ac.uk`. To receive announcements of new releases, send an e-mail to `majordomo@cedar.ac.uk` with `subscribe hztool-announce` in the body of the e-mail.

[2]More details may be found at `http://hepforge.cedar.ac.uk/rivet`.

[3]This naming should not be used for new routines; the HZD prefix is preferred.

in [22].

As for heavy flavour production, a number of HERA measurements of open charm and beauty production are included in the library [23]. From those, the newly written routines are: `HZH0108047` (*P. D. Thompson*), `HZH0312057` (*O. Gutsche*), `HZH0408149` (*A. W. Jung*). New publications [24] are to be implemented. The following routines for the TeVatron results [25] were also recently provided: `HZH9905024` (*O. Gutsche*), `HZH0307080` (*H. Jung*), `HZH0412071` (*H. Jung*, *K. Peters*). Furthermore a set of *Benchmark* cross sections have been defined for easy comparison of different calculations: `HZDISCC`, `HZDISBB` for charm and beauty production in DIS, `HZHERAC`, `HZHERAB` for photoproduction of charm and beauty and `HZLHCC`, `HZLHCBB` for charm and beauty production at the LHC [26].

References

[1] HZTOOL package, manual and tutorial can be downloaded from
`http://hepforge.cedar.ac.uk/hztool/`

[2] J. M. Butterworth and S. Butterworth, Comput. Phys. Commun. **153**, 164 (2003);
`http://jetweb.cedar.ac.uk`

[3] S. Chekanov, these proceedings, working group 5;
`http://hepforge.cedar.ac.uk/runmc/`

[4] J. Bromley *et al.*, Proceedings of the Workshop “Future physics at HERA”, Hamburg 1995/96, vol. 1, 611-612.

[5] J. M. Butterworth, S. Butterworth, B. M. Waugh, W. J. Stirling and M. R. Whalley, hep–ph/0412139;
A. Buckley *et al.*, these proceedings, working group 5;
`http://www.cedar.ac.uk/`

[6] L. Lönnblad, Comput. Phys. Commun. **71**, 15 (1992);
`http://www.thep.lu.se/~leif/ariadne/`

[7] H. Jung, G. Salam, Eur. Phys. J. C **19**, 351 (2001);
H. Jung, Comput. Phys. Commun. **143**, 100 (2002);
`http://www.desy.de/~jung/cascade/`

[8] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101**, 010 (2001);
`http://hepwww.rl.ac.uk/theory/seymour/herwig`

[9] J. M. Butterworth, J. R. Forshaw and M. H. Seymour, Z. Phys. C **72**, 637 (1996);
`http://hepforge.cedar.ac.uk/jimmy`

[10] G. Ingelman, A. Edin and J. Rathsman, Comput. Phys. Commun. **101**, 108 (1997);
`http://www3.tsl.uu.se/thep/lepto/`

[11] T. Sjöstrand, P. Eden, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun. **135**, 238 (2001);
`http://thep.lu.se/tf2/staff/torbjorn/`

[12] R. Engel, Z. Phys. C **66**, 203 (1995);
`http://www.physik.uni-leipzig.de/~engel`

[13] A. Ringwald and F. Schrempp, Comput. Phys. Commun. **132**, 267 (2000);
`http://www.desy.de/~t00fri/qcdins/qcdins.html`

[14] H. Jung, Comput. Phys. Commun. **86**, 147 (1995);
`http://www.desy.de/~jung/rapgap/`

[15] M. G. Ryskin and A. Solano,
`http://www-zeus.desy.de/~solano/RIDI2.0/`

[16] K. Rabbertz and T. Schörner-Sadenius, these proceedings, working group 5;
`http://www.desy.de/~nlolib/`

[17] H. Plothow-Besch, Comput. Phys. Commun. **75**, 396 (1993).

[18] W. Giele *et al.*, Proceedings of Workshop on Physics at TeV Colliders, Les Houches, France, 2001, [hep–ph/0204316];
M. Whalley, D. Bourilkov and R. C. Group, these proceedings, working group 5;
`http://hepforge.cedar.ac.uk/lhapdf/`

[19] G. A. Schuler and H. Spiesberger,
`http://wwwthep.physik.uni-mainz.de/~hspiesb/djangoh/djangoh.html`

[20] HZSTEER package can be downloaded from
`http://hepforge.cedar.ac.uk/hzsteer/`

[21] M. Derrick *et al.* [ZEUS Collaboration], Phys. Lett. B **342**, 417 (1995), `HZ94176`;
M. Derrick *et al.* [ZEUS Collaboration], Phys. Lett. B **348**, 665 (1995), `HZ95033`;
M. Derrick *et al.* [ZEUS Collaboration.], Phys. Lett. B **354**, 163 (1995), `HZH9505001`;
M. Derrick *et al.* [ZEUS Collaboration], Phys. Lett. B **369**, 55 (1996), `HZ95194`;
S. Aid *et al.* [H1 Collaboration], Z. Phys. C **70**, 17 (1996), `HZ95219`;
M. Derrick *et al.* [ZEUS Collaboration], Phys. Lett. B **384**, 401 (1996), `HZ96094`;
K. Ackerstaff *et al.* [OPAL Collaboration], Z. Phys. C **73**, 433 (1997), `HZC96132`;
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **1**, 97 (1998), `HZ97164`;
J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **2**, 61 (1998), `HZ97191`;
J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **1**, 109 (1998), `HZ97196`;
J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **4**, 591 (1998), `HZ98018`;
J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **6**, 67 (1999), `HZ98085`;
J. Breitweg *et al.* [ZEUS Collaboration], Phys. Lett. B **443**, 394 (1998), `HZ98162`;
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **10**, 363 (1999), `HZH9810020`;
G. Abbiendi *et al.* [OPAL Collaboration], Eur. Phys. J. C **10**, 547 (1999), `HZC98113`;
J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **11**, 35 (1999), `HZ99057`;
C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **483**, 36 (2000), `HZ00035`;
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **18**, 293 (2000) `HZH0006017`;
S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **23**, 615 (2002), `HZ01220`;
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **25**, 13 (2002), `HZ01225`;
T. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **65**, 092002 (2002), `HZF01211E`;
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **29**, 497 (2003), `HZH0302034`;
S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **35**, 487 (2004), `HZH0404033`.

[22] C. Buttar *et al.*, *Underlying events*, these proceedings, working group 2.

[23] C. Adloff *et al.* [H1 Collaboration], Z. Phys. C **72**, 593 (1996), `HZ96138`;
J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **6**, 67 (1999), `HZ98085`, `HZ98085p`;
C. Adloff *et al.* [H1 Collaboration], Nucl. Phys. B **545**, 21 (1999), `HZ98204`;
C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **467**, 156 (1999)
[Erratum-ibid. B **518**, 331 (2001)], `HZ99126`;

J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **18**, 625 (2001), `HZ00166`;
C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **528**, 199 (2002), `HZ01100`;
C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **520**, 191 (2001), `HZH0108047`;
S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **565**, 87 (2003), `HZ03015`;
S. Chekanov *et al.* [ZEUS Collaboration], Nucl. Phys. B **672**, 3 (2003), `HZ03094`;
S. Chekanov *et al.* [ZEUS Collaboration], hep–ex/0312057, `HZH0312057`;
A. Aktas *et al.* [H1 Collaboration], Eur. Phys. J. C **38**, 447 (2005), `HZH0408149`.

[24] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Rev. D **69**, 012004 (2004);
S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **599**, 173 (2004);
A. Aktas *et al.* [H1 Collaboration], Eur. Phys. J. C **40**, 349 (2005);
A. Aktas *et al.* [H1 Collaboration], Eur. Phys. J. C **41**, 453 (2005);
A. Aktas *et al.* [H1 Collaboration], Phys. Lett. B **621**, 56 (2005);
A. Aktas *et al.* [H1 Collaboration], Submitted to Eur. Phys. J. C , hep–ex/0507081;
S. Chekanov *et al.* [ZEUS Collaboration], hep–ex/0508019.

[25] B. Abbott *et al.* [D0 Collaboration], Phys. Lett. B **487**, 264 (2000), `HZH9905024`;
D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **91**, 241804 (2003), `HZH0307080`;
D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **71**, 032001 (2005), `HZH0412071`.

[26] O. Behnke, M. Cacciari, M. Corradi, A. Dainese, H. Jung, E. Laenen, I. Schienbein and H. Spiesberger, these proceedings, working group 3.

CEDAR

A Buckley [1], *J M Butterworth* [2], *S Butterworth* [2], *L Lönnblad* [3], *W J Stirling* [1], *M Whalley* [1], *B M Waugh* [2]
[1] Institute for Particle Physics Phenomenology, University of Durham, DH1 3LE, UK
[2] Department of Physics and Astronomy, UCL, Gower Street, London, WC1E 6BT
[3] Department of Theoretical Physics, Sölvegatan 14A, S-223 62 Lund, Sweden

Abstract
The CEDAR collaboration is developing a set of tools for tuning and validating theoretical models by comparing the predictions of event generators with data from particle physics experiments. CEDAR is also constructing resources to provide access to well defined versions of high-energy physics software and support for software developers. Here we give an overview of the CEDAR project and its status and plans.

1 Introduction

Despite the success of the Standard Model in accurately describing a wide range of phenomena in high-energy particle physics, there are aspects of high-energy collisions where technical difficulties in the relevant calculations make it hard to attain a good understanding This is particularly true where non-perturbative QCD is involved, as in the description of hadronic collisions, where the final state is influenced by the parton distribution functions (PDFs) of the colliding beams, by multiple soft interactions leading to an "underlying event" and by hadronisation of the outoing partons.

These theoretical uncertainties can limit the precision of new measurements, as well as hindering the planning of future experiments. Building accurate models of hadronic processes is important for these reasons as well as for the insight they may offer into the fundamental physics involved. However, the models that are constructed typically have a number of parameters that can be varied, constrained only by how well the resulting predictions agree with experiment.

Tuning these free parameters and testing the models against experimental data is a difficult task because the data are so varied, involving different beam particles, different regions of phase space and complex observables. Changing a single parameter in a model can affect the predictions for different measurements in very different ways, and tuning to a limited set of data may result in a contradiction with other data not taken into account.

It is thus important to compare models simultaneously with as wide a range as possible of experimental results, and the aim of CEDAR [1] is to simplify this task.

The rest of this contribution will describe in turn the projects making up CEDAR.

2 HZTool and Rivet

The first requirement is for a library of routines to enable, for each experimental measurement of interest, a comparable prediction to be produced from any given Monte Carlo generator. This role is currently filled by the Fortran library HZTool, described in more detail in another contribution to these proceedings [2]. The HZTool library is being maintained by CEDAR, with subroutines for various measurements contributed by a number of authors within and outside the CEDAR collaboration. A number of HERA routines were written within this workshop.

Work is underway to build a replacement for HZTool, to be called Rivet (Robust Independent Validation of Experiment and Theory). This will use an object-oriented design, implemented in C++, together with standard interfaces (such as HepMC [3] and AIDA [4]) to make the new framework more flexible and extensible than the Fortran HZTool. For example, it will be easier to incorporate new Monte Carlo generators into Rivet than into HZTool.

3 JetWeb

JetWeb [5] provides a web interface to HZTool, along with a relational database of both experimental data and model predictions generated using HZTool. The core of JetWeb is a set of Java servlets that manipulate an object model representing data and predictions. A user can use a web form to specify a model and choice of parameters, and the data they wish to compare to this model. If this model and parameter set are already in the database, a set of comparison plots and statistics is returned. Otherwise the user may request a set of Monte Carlo jobs to be run using their specified model, and the results will be added to the database.

The existing JetWeb database has been frozen, although it can still be searched, while the design and functionality are improved. JetWeb is being adapted to use the data already stored in HEPDATA rather than duplicating this in its own database, and to make the addition of further Monte Carlo models (beyond the currently supported HERWIG [6] and PYTHIA [7]) easier.

4 HEPDATA

HEPDATA [8] is a well established and widely used source of scattering data from HEP experiments. As part of CEDAR it has been converted from the existing hierarchical structure to a relational database using MySQL. The next steps in this part of the CEDAR project will provide front ends so that the data in the relational database can be accessed through a searchable web interface and also directly by JetWeb and other users.

5 HepML

In order to simplify the transfer of data between different parts of the CEDAR project and other software frameworks, an XML schema [9] is being developed to specify particle reactions and experimental results (as provided by HEPDATA) as well as generator programs and parameters.

This schema is separate from the HepML developed within the MCDB project [10], which is designed primarily as a format for event records. It may be that some parts of the schemas can be unified later or become parts of a more general schema.

6 HepCode

HepCode [11] aims to provide access to well defined versions of Monte Carlo programs, parton distribution functions and other high-energy physics calculation programs. Currently it is simply a list of codes with details for each of the processes calculated, the order of the calculation, the authors and the programming language used, along with a link to further information where this is available.

HepCode will eventually feature a search facility so that users can find a set of available programs simply by entering the details of a particular scattering process. It may also be possible to have links from matching data records in HEPDATA or from papers in bibliographic databases such as SPIRES.

7 HepForge

CEDAR also provides a development environment, HepForge, for authors of HEP software, including Monte Carlo generators. In addition to the core CEDAR projects (HZTool, HZSteer, JetWeb, HepML) other projects using HepForge are fastNLO [12], Herwig++ [13], Jimmy [14], KtJet [15], LHAPDF [16], RunMC [17] and ThePEG [18].

Facilities provided to developers include a code repository (using CVS or Subversion), a bug tracker (using Trac), a wiki for documentation and communication between project contributors, and mailing lists for project discussions, queries and announcements.

References

[1] J. M. Butterworth, S. Butterworth, B. M. Waugh, W. J. Stirling and M. R. Whalley, hep–ph/0412139, presented at CHEP'04, Interlaken, September 2004; `http://www.cedar.ac.uk/`

[2] J. M. Butterworth, H. Jung, V. Lendermann and B. M. Waugh, *HZTool*, these proceedings, working group 5;
HZTool package, manual and tutorial can be downloaded from `http://hepforge.cedar.ac.uk/hztool/`

[3] M. Dobbs and J. B. Hansen, Comput. Phys. Commun. **134** 41 (2001); `http://mdobbs.home.cern.ch/mdobbs/HepMC/`

[4] `http://aida.freehep.org/`

[5] J. M. Butterworth and S. Butterworth, Comput. Phys. Commun. **153**, 164 (2003); `http://jetweb.cedar.ac.uk/`

[6] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101**, 010 (2001); `http://hepwww.rl.ac.uk/theory/seymour/herwig/`

[7] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun. **135**, 238 (2001); `http://www.thep.lu.se/ torbjorn/Pythia.html`

[8] `http://durpdg.dur.ac.uk/hepdata/`

[9] `http://hepforge.cedar.ac.uk/hepml/`

[10] P. Bartalini, L. Dudko, A. Kryukov, I. Seluzhenkov, A. Sherstnev, A. Vologdin, hep-ph/0404241.

[11] `http://www.cedar.ac.uk/hepcode/`

[12] `http://hepforge.cedar.ac.uk/fastnlo/`

[13] `http://hepforge.cedar.ac.uk/herwig/`

[14] J. M. Butterworth, J. R. Forshaw and M. H. Seymour, Z. Phys. C **72**, 637 (1996); `http://hepforge.cedar.ac.uk/jimmy/`

[15] J. M. Butterworth, J. P. Couchman, B. E. Cox and B. M. Waugh, Comp. Phys. Comm. **153**, 85 (2003) `http://hepforge.cedar.ac.uk/ktjet/`

[16] M. Whalley and S. Bourikov, these proceedings, working group 5; `http://hepforge.cedar.ac.uk/lhapdf/`

[17] S Chekanov, these proceedings, working group 5; `http://hepforge.cedar.ac.uk/runmc/`

[18] L. Lönnblad, these proceedings, working group 5; `http://hepforge.cedar.ac.uk/thepeg/`

RunMC: an object-oriented analysis framework to generate Monte Carlo events for current and future HEP experiments

S. Chekanov
HEP division, Argonne National Laboratory, 9700 S.Cass Avenue, Argonne, IL 60439 USA
E-mail: chekanov@mail.desy.de

Abstract
RunMC is a C++ object-oriented framework aimed to generate and to analyze high-energy collisions using Monte Carlo simulations. The package provides a common interface to different Monte Carlo models using modern physics libraries developed for the LHC and NLC experiments. Physics calculations can easily be loaded and saved as external project modules. This simplifies the development of complicated physics calculations for high-energy physics in large collaborations.

1 Introduction

Monte Carlo models (MC) written in FORTRAN 77 are widely used in many high-energy physics laboratories worldwide. These models are known to be fast, robust and well tested. However, the main choice for future high-energy experiments is an object-oriented programming language, either C++ (the LHC experiments at CERN) or Java (the NLC project). Some steps towards converting the FORTRAN MC models to the C++ programming language have already been undertaken [1]. However, such models written in C++ will require a thoughtful verification to insure that their predictions are consistent with the original FORTRAN-based MC programs, as well as with the physics results obtained in the past. Such verifications will go over certain time, and a tool which allows to perform such comparisons is urgently needed.

A program which allows running of both FORTRAN-coded and C++ MC models using a common C++ programming environment should be valuable. This is important not only for comparisons and verifications of these MC models. Such a C++ framework can also extend the lifetime of FORTRAN-based models especially for the LHC, NLC and TEVATRON communities, and can provide compatibility of most popular MC models with the new software to be used for current and future HEP experiments.

The RunMC package [2] is a common C++ frond-end of Monte Carlo models which provides a unified approach to generate and analyze very different MC models independent of their native codes. In this approach, the MC output (typically the HEPEVT record) is converted to C++ classes for further analysis or graphical representation (histograms). The graphical user interface (GUI) of this program helps to initialize MC models and histograms, as well as to monitor the event generation.

The RunMC program fully complies with the change in the programming paradigm of data analysis, and meets the requirements of future high-energy experiments. Instead of FORTRAN-based analysis tools, such as PAW [3] and HBOOK [4], it uses the modern CERN C++ analysis packages, CLHEP [5] and ROOT [6].

In this respect, the RunMC program is similar to the JetWeb server [7], which also provides the ability to compare the existing MC models and to confirm the physics assumptions they contain. However, in contrast to JetWeb, the RunMC program was designed as a stand-alone desktop application. Therefore, the user has full access to his calculations and to the program itself.

The RunMC program also provides an interface to the popular HZTOOL package [8], thus many physics calculations from HERA, LEP and TEVATRAN can easily be accessed. In addition, within the RunMC approach, the concept of project modules was introduced (in fact, the HZTOOL package is

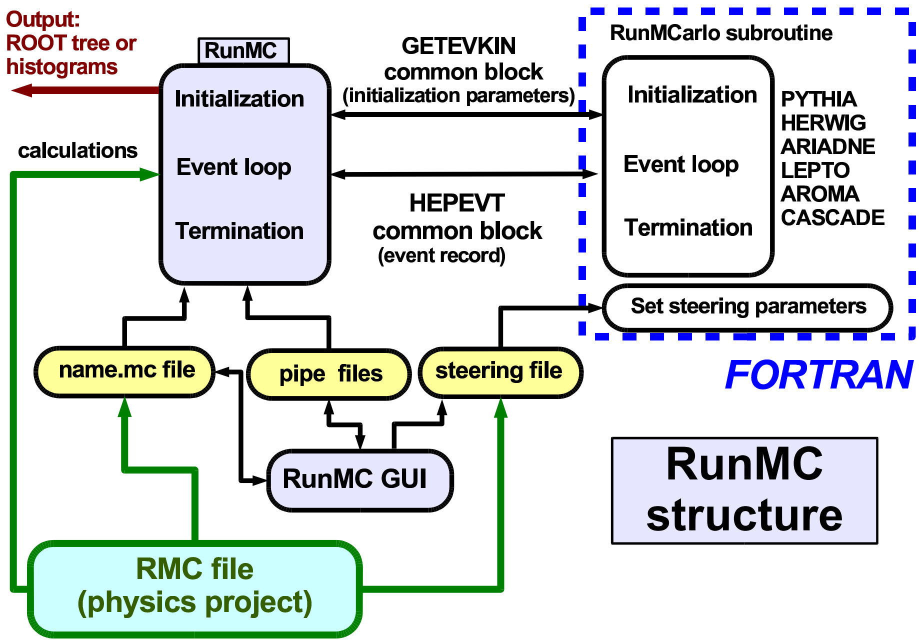


Fig. 1: The diagram shows the structure of the RunMC program.

one of such modules). A project file, which can contain external calculations, MC tunings, histogram definitions, etc., can be loaded to RunMC with the same ease as a document can be opened in the Microsoft Word program. The project files are small and platform independent, therefore, it is fairly simple to share complicated physics calculations between scientists in large collaborations.

2 Program structure

RunMC consists of a GUI and several RunMC MC programs. There are two implementations of the RunMC GUI: one is written using the Wide Studio C++ classes [9] and an alternative GUI is based on Java. Due to complete independence of RunMC GUI from RunMC MC programs, one can run jobs in the background without any GUI or pop-up window.

The RunMC programs integrate the C++ ROOT and CLHEP packages with native implementations of MC models. A schematic structure of RunMC is illustrated in Fig. 2. The following MC models are included to RunMC (version 3.3): PYTHIA 6.2 [10], HERWIG 6.5 [11], ARIADNE 4.12 [12], LEPTO 6.5 [13], AROMA 2.2 [14], CASCADE 1.2 [15], PHOJET 1.05 [16], RAPGAP 3.1 [17]. There are several executable RunMC MC programs corresponding to each MC model.

RunMC GUI communicates with the RunMC MC programs using pipe files located in the directory “$RUNMC/pipes”. Here, “$RUNMC” denotes the installation directory which has to be defined by the user. All the directories to be discussed below are assumed to be located in this directory.

2.1 RunMC GUI

RunMC GUI allows an interaction between the user and RunMC MC programs. At present, two types of RunMC GUI are available: a user interface based on C++ (can be executed with the command “runmc”) and that based on Java (the command “jrunmc”). Below we will describe only the C++ RunMC GUI.

The task of RunMC GUI is to generate the output file “project.mc”, where “project” is a user-defined name of the current calculation. This file contains the most important information for MC running, such as the type of MC model, the number of output events, the type of colliding particles, their energies, RunMC output (histograms or ntuples) etc..

RunMC GUI adopts the following strategy to define the histograms: The window "Variables" contains the names of the variables (with some additional comments) defined for a certain physics project. The user should select the appropriate variable and copy it to the window "Histograms" by clicking on the corresponding variable name. If two one-dimensional histograms are defined, a two-dimensional histogram can be build from these two histograms using RunMC GUI.

The variable names are divided into the three categories: event-based variables (characterizing the event as whole), single-particle densities (filled for each particle/jet; the variable name starts with "@") and two-particle densities (filled for each particle/jet pair; the name starts with "@@"). The histograms can also be filled in the user-defined subroutine "user-run.cpp"; in this case the naming convention for the variables discussed above is unnecessary.

During the event generation, the ROOT canvas can display the output histograms (up to eight in total) with filled events. The output log from RunMC MC is written to ".analmc.log" (a symbolic link to the "project.log" file). Possible errors are redirected to the file "project.err", which is constantly monitored by RunMC GUI.

The ROOT histograms are automatically modified at the end of the fill if they are required to be normalized to the total number of events or converted to differential cross sections. Note that there is no need to wait until the end of the current run: once the histogram statistics is sufficient, one can terminate the run by clicking "Stop" on the GUI window. Histograms should be saved in the ROOT file "project.root" for further studies. The style of the histograms can further be modified using the ROOT canvas editor.

2.2 The RunMC MC programs

The RunMC MC programs integrating Monte Carlo models with ROOT C++ classes have the generic names "analmc.MCname", where "MCname" denotes the MC name. The main C++ function of RunMC MC is located in the file "analmc.cpp" (in the directory "main/src"). The C++ code accesses the HEPEVT common block of a given MC program via a C-like structure. The RunMC MC program receives the initial parameters via the symbolic link ".analmc.ln" pointing to the file "project.mc".

Each MC model has its own FORTRAN subroutine "runmcarlo" which provides an interface to the native MC code. This interface program (in the file "RUNMC-MCname.f") is located in the directory "main/mcarlo/MCname". The task of the subroutine "runmcarlo" is to fill the HEPEVT common block. In addition, some initial settings are done by accessing a C/C++ structure with the initial parameters defined in the "project.mc" file. The main function in "analmc.cpp" calls this interface subroutine and fills the C/C++ structure which represents the complete HEPEVT event record. The output is copied to the class "HEPLIST" which can be accessed by external calculations. The HEPLIST class consists of several vectors based on the LorentzVector vector class (from the CLHEP library) which represents four-momentum of a particle (or a jet). The definition of the HEPLIST class, as well as other include files, can be found in the "main/inc" directory. Note that the user still can access more elementary event records (such as FORTRAN HEPEVT common block) which can be used to transform them to other event classes and physics calculations.

There are several physics packages available inside RunMC MC to transform the original four-momentum vector of particles/jets to the required observable:

- the transformations provided by the physics vector class "LorentzVector" from CLHEP can be used, since a particle or a jet is represented as a general four-vector based on this class;
- the event-shape calculations are available using the package developed by M. Iwasaki [18];
- the longitudinally-invariant k_T algorithm as implemented in C++ [19] can be used for the jet

reconstruction. In addition to this package, the JADE and Durham jet algorithms are implemented according to M. Iwasaki [18];

- the Breit frame calculations were implemented for ep deep inelastic scattering.

The physics packages with the corresponding documentation are located in the directory “main/physics”.

3 User calculations

For a new physics calculation, the directory “proj” should be modified. This user directory can contain external calculations, steering cards for MC initializations, as well as the standard RunMC functions which are necessary to initialize and fill the histograms.

The user directory should always contain the file “project.mc” created by RunMC GUI. This file can be edited manually without the RunMC GUI program using any text editor. On Linux/Unix, one can load this file to RunMC GUI by executing the command “runmc project.mc” from the shell prompt (or using the option “Projects→read MC” of RunMC GUI).

The directory “proj” can contain steering files “MCname.cards” to redefine initial MC parameters. Such files can be created via RunMC GUI (“MC settings” option). For more flexibility, the MC initialization parameters can also be overwritten by FORTRAN-coded subroutines located in the directory “proj/ini”. If this is not done, the default MC parameters will be selected according to the RunMC option.

To define histograms, the user-defined variables should be calculated in the file “proj/user_afill.cpp”. The output of this function is a pointer. The output variable name should always be associated with this pointer. The variable names should be specified in the file “user-name.txt”. It includes the variable names to the list “Variables” accessed by RunMC GUI. Finally, to compile the source codes and to rebuild all RunMC MC programs to take into account changes made in the project source files, one should type “make” in the “proj” directory. All MC programs will be recompiled and RunMC GUI will be updated with new histograms. Then, the command “runmc” (or “jrunmc”) should be executed from the directory “proj” to start RunMC GUI. The main advantage of this approach is that once a necessary variable is defined, new histogram definitions do not require the MC recompilation.

RunMC histograms can also be filled using the conventional method, i.e. in the function located in “user-run.cpp”. In this case, the initialization of histograms is not required, as long as the file “project.mc” defines which histograms should be filled and what presentation style should be used to fill the histograms. The histograms can be initialized in the file “user-init.cpp” using the standard ROOT procedure.

4 Physics calculations as external RMC projects

In order to share complicated analysis calculations or to store them for future use, the directory “proj” can be packed into an external RMC file with the extention “rmc”. For example, “project.rmc” is the RMC file which has the user-defined name “project”. The “proj” directory inside of it has a file “project.mc” with RunMC GUI settings, user-defined external functions, libraries, make files MC steering files, etc..

RunMC GUI can automatically load and recompile such project RMC files (see details in [2]). The user can also save his/her calculations into a RMC file for future analysis. As it was mentioned, the project files are compact and platform independent, therefore, it is fairly simple to share physics calculations between the users, as long as the RunMC package is installed.

At present, several RunMC project files are available on the Web [2] (they are also included in the directory “archive” of RunMC):

- the default project. Only pre-installed variables can be included in the calculations.
- HERA kinematic variables (Q^2, x, etc.);
- jets at HERA and LHC using the longitudinally-invariant k_T algorithm in the Breit frame. In

addition, the ratio of jet cross sections at the parton and hadron levels are calculated (the so-called hadronisation corrections);

- D^* cross sections in ep collisions at HERA;
- cross sections for strange-particle production in ep collisions at HERA;
- the HZTOOL package [8];
- the event-shape variables in e^+e^- at NLC energies;
- several examples of how to visualize tracks and k_T jets in 3D for a single MC event (e^+e^-, ep, pp collisions).

The RMC project files discussed above only illustrate how to set up and to develop new physics calculations in the RunMC framework. For practical applications, these examples should be modified.

5 RunMC ROOT tree analyzer

In addition to the standard functionality of the MC event simulation, RunMC GUI can also use ROOT trees as the input for physics calculations.

The ROOT tree can be generated by selecting the option "HEPEVT" or "RUNMC", in addition or instead of the ROOT histogram option. Then, the MC events should be generated as usual, but this time a ROOT tree with the extension ".rtup" or ".htup" will be created. Then, RunMC can run over this ROOT tree if, instead of the MC model, the option "RUNMC" or "HEPEVT" is selected. Several ROOT trees can automatically be included in the analysis, as long as they are in the same directory. The analysis of the ROOT trees is very similar to the standard run over MC events. External RMC files can be used to include new calculations, variables and histograms.

The main advantage of the RunMC ROOT tree analyzer is that physics calculations can be validated significantly faster than when RunMC is used to generate events and to fill histograms at the same time. In case of the ROOT trees, RunMC can fill histograms by a factor of $\sim$10–15 faster, thus the RMC project files can be validated and analyzed more efficiently.

With this additional functionality, the RunMC program can also be used to analyze experimental data if the event record is converted to the appropriate ROOT tree. The data analysis can be performed using exactly the same RMC project files as for the usual MC simulation runs.

References

[1] M. Bertini, L. Lönnblad and T. Sjöstrand, Comput. Phys. Commun. **134**, 365 (2001);
L. Lönnblad, THEPEG*: Toolkit for High Energy Physics Event Generation* (unpublished), available on `http://www.thep.lu.se/ThePEG/`.

[2] S. Chekanov, RUNMC - C++ *object-oriented framework for monte carlo models*. Preprint hep-ph/0411080, Comm. Phys. Comm. (in press), available on
`http://hepforge.cedar.ac.uk/runmc/`
`http://www.hep.anl.gov/chakanau/runmc`.

[3] Application-Software-Group, PAW: *Physics Analysis Workstation*, available on `http://wwwasd.web.cern.ch/wwwasd/paw/`.

[4] R. Brun and D. Lienart, CERN-Y250 (1987).

[5] L. Lönnblad, Comput. Phys. Commun. **84**, 307 (1994);
M. Fischler and A. Pfeiffer, CLHEP - *new developments and directions*, in *The proceedings of International Conference on Computing in High Energy Physics and Nuclear Physics (CHEP 2000), Padova, Italy, 7-11 Feb 2000.* Available on
`http://wwwasd.web.cern.ch/wwwasd/lhc++/clhep/`.

[6] R. Brun and F. Rademakers, Nucl. Instrum. Meth. **A389**, 81 (1997);
R. Brun, F. Rademakers, P. Canal and M. Goto, ECONF **C0303241**, MOJT001 (2003).

[7] J.M. Butterworth and S. Butterworth, Comput. Phys. Commun. **153**, 164 (2003).

[8] J. Bromley et al., HZTOOL: *a package for* MONTE CARLO - *data comparisons at hera*, in *Future Physics and HERA*, eds. Buchmuller, W. and Ingelman, G., Vol. 2, p. 611. Hamburg, Germany, 1996-1997, available on `http://hztool.hep.ucl.ac.uk/`.

[9] T. Hirabayashi, *The Wide Studio project* (unpublished), available on `http://www.widestudio.org`.

[10] T. Sjöstrand, L. Lönnblad and S. Mrenna, *Pythia 6.2: Physics and manual*. Preprint hep-ph/0108264, 2001.

[11] G. Corcella et al., JHEP **0101**, 10 (2001).

[12] L. Lönnblad, Comput. Phys. Commun. **71**, 15 (1992).

[13] G. Ingelman, A. Edin and J. Rathsman, Comput. Phys. Commun. **101**, 108 (1997).

[14] G. Ingelman, J. Rathsman and G. Schuler, Comput. Phys. Commun. **101**, 135 (1997).

[15] H. Jung, Comput. Phys. Commun. **143**, 100 (2002).

[16] R. Engel, PHOJET (unpublished), 1997, available on `http://www-ik.fzk.de/~engel/phojet.html`.

[17] H. Jung, Comm. Phys. Commun. **86**, 147 (1995).

[18] M. Iwasaki, *Jet-EventShape-Finders package* (unpublished).

[19] J.M. Butterworth, J.P. Couchman, B.E. Cox and B.M. Waugh, Comput. Phys. Commun. **153**, 85 (2003).

A C++ framework for automatic search and identification of resonances

S. Chekanov
HEP division, Argonne National Laboratory, 9700 S.Cass Avenue, Argonne, IL 60439 USA
E-mail: chekanov@mail.desy.de

Abstract
This paper describes the first proof-of-concept version of a C++ program designed for peak searches in invariant-mass distributions. The program can be useful for searches of new states as well as for the reconstruction of known resonances.

Presently, there is a considerable progress in understanding the rich spectrum of hadron resonances. However, even in case of known baryonic resonances, there are still many open questions. The Particle Data Group (PDG) quotes more than 100 baryonic states, but only half of them are reasonably well established. Recently, the revitalized interest to baryon spectroscopy was triggered by observation of narrow peaks in the $K^{\pm}n$ and $K^0_S p$ invariant-mass distributions which can be interpreted as pentaquarks. There is also longstanding interest to other exotic multiquark states, glueballs, hybrids, baryonia, etc. The HERA experiments have active program for searching such states [1]. At the LHC this physics program will continue.

In hadron spectroscopy, searches for new resonances and measurements of known states are based on the reconstruction of invariant-mass distributions of two or more tracks in order to find the rest mass of the originally decaying particle. The usual procedure for such studies is to assign certain masses to tracks, and then to combine their four-momenta to form the invariant-mass distributions. From the observed peaks, one can determine resonance masses, widths and cross sections.

Obviously, a particle identification is highly desirable for any experiment in order to reduce combinatorial background. Without the ability to identify tracks, the combinatorial background rises as $\sim 0.5n^2$, where n is the number of produced particles used in the searches. Thus high-energy experiments have to deal with a significant background for searches in fully inclusive events. For the reconstruction of three- and four-body decays, the combinatorial background is even higher than for the two-body decays.

Particle identification can be achieved using various approaches, such as the energy loss per unit length (dE/dx), silicon-strip detectors, Cherenkov counters etc.. In many cases, the particle identification cannot be perfect. For example, when the dE/dx method is used for pentaquark searches, it is difficult to obtain a high-purity sample of kaons or protons for tracks with large momenta ($p > 1$ GeV) due to significant overlap of the dE/dx bands for different particle species in this momentum range. In this case, several mass assumptions for invariant-mass distributions are needed to be checked in order to exclude peaks due to possible misidentification.

Searches for new resonances using various mass assumptions remain to be a tedious task since no much progress has been made so far to develop a tool which can automate this procedure. For example, the reconstruction of two-, three-, four-body decays involving only three mass assumptions leads to 36 possible non-identical invariant-mass distributions (6 distributions for two-body decays, 10 - for three-particle decays and 20 - for four-particle decays). All these invariant-mass distributions should be reconstructed, analyzed and possible reflections from known PDG states, when different mass assumptions are used, should be disregarded. Obviously, this time consuming work can be simplified and automated.

The program called "SBumps", which is still at the early stage of the development, attempts to accomplish this task. It rather represents the first proof-of-concept version of a program which helps to perform automatic searches of peaks in invariant-mass distributions. To perform the reconstruction of invariant-mass distributions, the user should specify:

- the event record, i.e. a list of tracks and (optionally) the probabilities that tracks belong to certain particle species. Such probabilities can be obtained using various particle-identification technics;
- the names of particle species used during the mass assignments;
- the statistical significance of expected peaks;
- instrumental invariant-mass resolution;
- how many tracks should be combined to the invariant-mass distributions.

Once the initial conditions are specified, the program runs over the event list and reconstructs all possible invariant-mass distributions with the mass assumptions as specified by the user. SBumps can reconstruct at the same time $2-$, $3-$ and $4-$ body decays in different combinations. At the end of the run, the program saves all the created invariant-mass distributions to the ROOT histograms [2].

At the second stage, the program analyses the created distributions and attempts to find statistically significant peaks. SBumps uses the build-in ROOT fast peak finder (from the TSpectrum class), which uses a fast deconvolution method [3] based on a Markov approach for peak searching in presence of a background and statistical noise. This algorithm was mainly developed for narrow, high-amplitude peaks which are characteristic for γ-ray physics. Therefore, this algorithm is not completely appropriate for peak searches in the invariant-mass spectra without tunings of the initial parameters of this algorithm. In case of the SBumps package, the deconvolution techniques is only used to identify the so-called seeds, i.e. bins with positions of possible peaks above a smooth background. The number of seeds can be rather large, and not all of them correspond to statistically significant peaks. At this stage, several adjustable parameters are available, such as the resolution of neighboring peaks, the peak sensitivity and the peak thresholds. Such parameters can be set using the steering file before each run.

Next, the program evaluates each seed by calculating the statistical significance, $N(S)/\Delta N$, where $\Delta N = \sqrt{N(B) + N(S)}$, with $N(S)(N(B))$ being the number of the signal (background) events. The calculation was done by comparing the values at the seed positions with the background level, which is determined from the neighboring bins. Then the peaks, which have the statistical significance above the level specified by the user, are considered for further analysis. This part of the program can be further improved, introducing more complicated algorithms. For example, sufficiently broad resonances might be overlooked by the present prescription.

The two-step procedure described above simplifies the peak search since the user normally does not need to deal with tunings of the initial parameters for the fast deconvolution method. At the same time, the method is sufficiently fast and can possibly be extended by introducing other algorithms to identify the peaks. For example, algorithms based on a fitting procedure can also be used; in this case, the seed positions can be used as the initial parameters of the fit functions.

At the third stage, the program attempts to identify the found peaks by comparing them with the masses of known PDG states. For this, a look-up table containing the information on established PDG resonances is used. Since the errors on the reconstructed peaks and the errors on the masses of known resonances taken from the PDG look-up table are known, the program matches the peaks using a simple criteria: $\mid L \mid /\Delta L < S$, where S is a free parameter given by the user, L is the distance between the PDG mass and the peak position and ΔL is the error on the reconstructed peak position combined with the error on the mass of known PDG resonances. For this matching procedure, the information on charge of decaying resonance is properly taken into account. In future, the information on specific decay channel can also be taken into account in order to reduce misidentification in the matching procedure.

Figure 1 shows the output of the SBumps program for events generated with PYTHIA Monte Carlo model. This example shows one histogram for two-body decays with two automatically identified peaks corresponding to known states. The required statistical significance for the final peaks was 4σ. This example shows that the program can easily find and identify rather broad resonances (K^*), as well

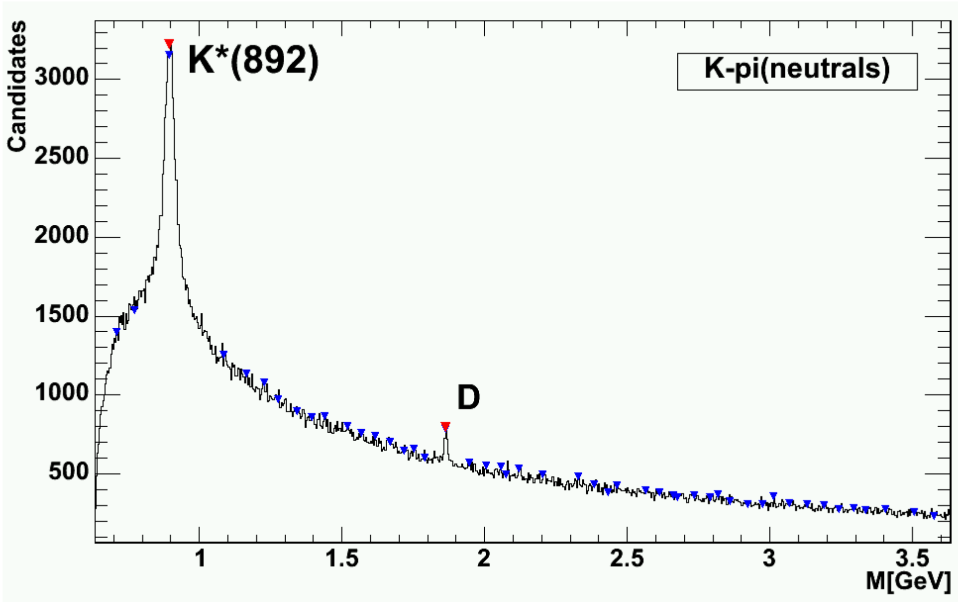


Fig. 1: The two-particle invariant-mass distribution calculated using the events generated with the PYTHIA Monte Carlo model (pp collisions at 1.4 TeV), when two mass hypothesis are used ($K^{\pm}$ and π mesons). The blue triangle symbols show the seeds used for the calculations of statistical significance. The seeds were found by using the fast deconvolution method. Most of the seed peaks were disregarded after the final calculation of the statistical significance. The program correctly identifies the known PDG states, K^* and D^0, after comparing the reconstructed peak positions with the PDG look-up table.

as narrow peaks which do not have high statistical significance.

The proof-of-concept version of the SBumps program is available as a loadable RMC file of the RunMC analysis framework [4]. After loading the "sbumps.rmc" module, RunMC first creates a ROOT tree with Monte Carlo events. The initial conditions are given in the file "sbumps.cards". The executable file "sbumps.exe" can be used to run over the events. At the end of the calculations, the ROOT browser should display the reconstructed histograms with invariant-mass distributions. The peaks which have the statistical significance above that specified by the user should be labeled.

References

[1] HERMES Coll., A. Airapetian et al., Phys. Lett. **B 585**, 213 (2004);
ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C38**, 29 (2004);
ZEUS Coll., S. Chekanov et al., Phys. Lett. **B591**, 7 (2004);
ZEUS Coll., S. Chekanov et al., Phys. Lett. **B610**, 212 (2005);
H1 Coll., A. Aktas et al., Phys. Lett. **B588**, 17 (2004).

[2] R. Brun and F. Rademakers, Nucl. Instrum. Meth. **A389**, 81 (1997);
R. Brun, F. Rademakers, P. Canal and M. Goto, ECONF **C0303241**, MOJT001 (2003).

[3] M. Morhac et al., Nucl. Instrum. Meth. **A401**, 113 (1997).

[4] S. Chekanov, RUNMC - C++ *object-oriented framework for monte carlo models.* Preprint hep-ph/0411080, Comp. Phys. Comm. (in press), available on
`http://hepforge.cedar.ac.uk/runmc/`,
`http://www.hep.anl.gov/chakanau/runmc`.

6 Acknowledgements

We are grateful to the CERN and DESY directorate for financial support of this workshop and for the hospitality which they extended to all the participants of the workshop. We are grateful D. Denise, A. Grabowksi and S. Platz for their continuous help and support during all the meeting weeks. We would like to thank also B. Liebaug for the design of the poster for this first HERA-LHC workshop.