Collins-Soper kernel determination in gluon-iniciated processes

Oris Suárez Eng, Havana University, Cuba

September 7, 2022

Abstract

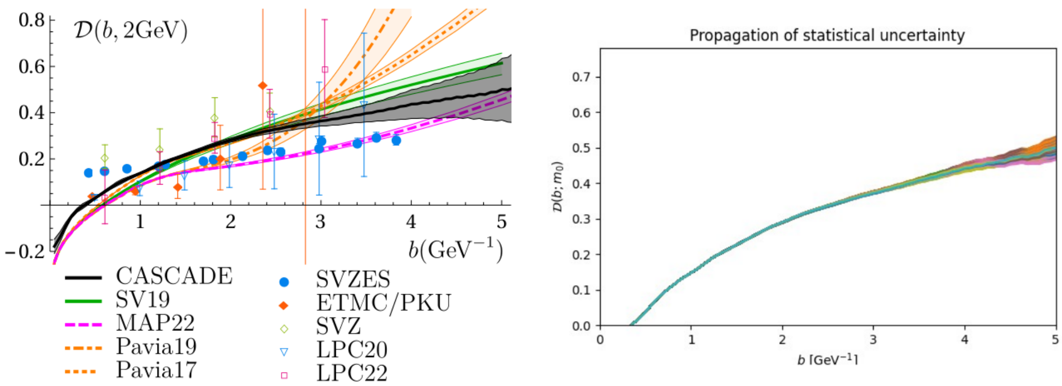
The rapidity anomalous dimension (RAD), or Collins-Soper (CS) kernel, defines the scaling properties of TMD distributions and can be extract from the experimental data, from the direct comparison of differential cross-sections measured a different energies. In this work we test this new method of determination of the CS kernel for gluon iniciated processes, explicitly for pair production of Higgs boson processes.

Contents

1 Introduction

The QCD vacuum is still a big thing. Its complex structure origins lots of fundamental and unanswered questions such as thoses regarding dark matter and confinament mechanism. The confinament mechanism is the phonemenon that color-charged particles (such as quarks and gluons) cannot be isolated, and therefore prevents us from any direct exploration of the hadron's insides. Even though direct explorations are not possible, physicist still manage to come up with indirect approaches using, mainly, the analysis of the differential cross-section of particle scattering. Let's just say that in a scattering process we can always split the total differential cross-section into two terms, one perturbative term, which means that we can use perturbation theory to determinate its value; and a second non-perturbative term, but universal, it does not depend on the specific process. This non-perturbative, universal, term is known as RAD (rapidity anomalous dimension), or Collins-Soper (CS) kernel, which emerges from the factorization theorem for the transverse-momentum differential cross-section, and was first introduced by [1].

The CS is not a characteristic of a hadron, it provides us information about the long range forces acting on quarks that are imposed solely by the non-trivial structure of the QCD vacuum. Despite being a part of the factorization theorem, the CS kernel is conceptually different from other distributions. The CS kernel contains information about the soft-gluon exchanges between parton and it also dictates the evolution properties for the TMD distribution functions.



(a) CS determination for different approaches

(b) CS determination for Drell-Yan process

Figure 1: Previous results of CS kernel determinationn for Drell-Yan process

At present, there are different approaches to the determination of the CS kernel 1a. In this work, we use a direct way of extracting the CS kernel from the scattering data, using lattice computation to form proper cross-sections ratios. This approach is a recent method suggested by [2]-[3]. Figures 1b shows the results obtained for the CS kernel in Drell-Yan process using this approach. The results obtained were good enough to encourage us to test this method in new processes [3]. This calculations were done using

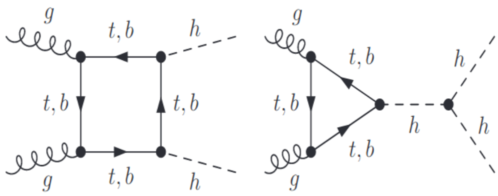


Figure 2: Feyman diagrams contributing to Higgs boson pair production via gluon fusion at leading order

for CS of quark TMD distributions but in this work we decided to study the CS for gluon TMD distributions for double Higgs Bosons production processes.
One of the primary goals of the LHC programme in the next decade is the detailed study of Higgs boson properties. In particular, the high luminosity upgade of the LHC is expected to provide direct constraints on the higgs boson trilinear coupling from Higgs boson pair production, which may reveal whether the Higgs potential is indeed Standard Model-like. A detailed theoretical understanding of the Higgs boson pair production processes is thus mandatory.
The phenomenology of multi-Higgs boson final states will provide complementary information to that found from single Higgs physics at the LHC. Due to generically small inclusive cross sections and a difficult signal vs. background discrimination, the best motivated multi-Higgs final states at theLarge Hadron Collider are Higgs boson pair final states, of which gluon fusion $gg \to hh$ is the dominant production mode 2.
Hence, in this letter, we decided to test this new method's power for gluon TMD distribution functions using the pseudodata generated by the CASCADE event generator.

2 Method

This method is founded on the leading power transverse momentum dependent (TMD) factorization theorem. We consider the Di-Higgs production process $g_1 + g_2 \to h_1 + h_2$. The cross section for the Di-Higgs pair production at small transverse momentum, described by the TMD factorization formula is

$$\frac{d\sigma}{dQ^2 dy dq_T^2} = \frac{2\pi}{9}\frac{\alpha_{em}^2(Q)}{sQ^2}|C_v(Q,\mu_Q\ )|^2 \int_0^\infty db b J_0(bq_T) \times \sum_q e_q^2 f_{q,h_1}(x_1, b; \mu_Q, Q^2) f_{\bar{q},h_2}(x_2, b; \mu_Q, Q^2) \quad (1)$$

where Q, y and q_t are the invariant mass, rapidity and tranverse momentum of the double Higgs system, $\mu_Q \sim Q$ is the factorization scale, s is the invariant mass of the initial state, e_q are the electric charges of quarks, α_{em} is the fine-structure constant and J_0 is the Bessel function of the first kind. The variables x_1 and x_2 are longitudinal momentum fractions.

$$x_1 = \frac{Q}{\sqrt{s}} e^y, \qquad x_2 = \frac{Q}{\sqrt{s}} e^{-y} \tag{2}$$

The hard coefficient function C_V is entirely perturbative and known up to next-to-next-to-next-to-leading order (N^3LO). The functions f are non-perturbative unpolarized TMD distributions.
The CS kernel is hidden in the Q-dependence of TMD distributions that is described by a pair of evolution equations

$$\frac{df_{q,h}(x,b;\mu,\zeta)}{d\ln\mu^2} = \frac{\gamma_F(\mu,\zeta)}{2} f_{q,h}(x,b;\mu,\zeta) \tag{3}$$

$$\frac{df_{q,h}(x,b;\mu,\zeta)}{d\ln\zeta} = -\mathcal{D}(b,\mu) f_{q,h}(x,b;\mu,\ zeta). \tag{4}$$

Here, γ_F is the TMD anomalous dimension which is perturbative and known up N^3LO, and ζ is the non-perturbative CS kernel. Thus, to extract the CS kernel one must explore the Q-dependence of the cross-section.
The essential complication of any phenomenological analysis with the TMD factorization is that all non-perturbative functions are defined in the position space. We perform the inverse Hankel transform of the cross-section

$$\sum(s,y,Q,b) = \int_0^\infty dq_T q_T J_0(q_T b) \frac{d\sigma}{dQ^2 dy dq_T^2} \tag{5}$$

Formula 1 is valid at small q_T/Q, and the corrections to it are estimated as $\sim 1\%$ at $q_T = 0.1Q$. Consequently, $\sum$ is accurately (up to 1%) described in the terms of TMD distributions for $b \gtrsim (0.1Q)^{-1}$.
The main idea of the method is to compare $\sum$'s measured at different Q's (Q_1 and Q_2), such that the TMD distributions f exactly cancel in the ratio. To perfom the cancellation we adjust the values of s such that the variables $x_{1,2}$ are identical. We compute

$$\frac{\sum(s_1,y,Q_1,b)}{\sum(s_2,y,Q_1,b)} = \left(\frac{Q_2}{Q_1}\right)^4 Z(Q_1,Q_2) e^{2\delta(b;Q_1\to Q_2)}, \tag{6}$$

where $s_1/s_2 = Q^1/Q_2^2$. The function Z is entirely perturbative

$$Z(Q_1,Q_2) = \frac{\alpha_{em}^2(Q_1)|C_v(Q_1,\mu_{Q_1})|^2}{\alpha_{em}^2(Q_2)|C_v(Q_2,\mu_{Q_2})|^2}. \tag{7}$$

The function Δ is resulted from the evolution of TMD distribution to the same scale by equations 3 and 4,

$$\Delta(b, Q_1 \to Q_2) = \int_P \left(\gamma_F(\mu,\zeta)\frac{d\mu}{\mu} - \mathcal{D}(b,\mu)\right), \tag{8}$$

where P is a path connecting points (μ_{Q_1}, Q_1^2) and (μ_{Q_2}, Q_2^2) in the (μ, ζ)-plane. Thus, the only non-perturbative content in the formula 5 is the CS kernel.
To invent the formula 5 we use the rectangular contour for the integrations path in eqn.7 and find

$$(D)(b, \mu_0) = \frac{\ln\left(\frac{\sum_1}{\sum_2}\right) - \ln Z(Q_1, Q_2) - 2\Delta_R(Q_1, Q_2; \mu_0)}{4\ln(Q_2/Q_1)} - 1, \tag{9}$$

where $\sum_1 / \sum_2$ is shorthand notation for the ratio 5, and

$$\Delta_R(Q_1, Q_2; \mu_0) = \int_{\mu_{Q_2}}^{\mu_{Q_1}} \frac{d\mu}{\mu}\gamma_F(\mu, Q_1) - 2\ln\left(\frac{Q_1}{Q_2}\right)\int_{\mu_0}^{\mu_{Q_2}} \frac{d\mu}{\mu}\Gamma_{cusp}(\mu), \tag{10}$$

with Γ_{cusp} being the cusp anomalous dimension. The last term in eqn.9 evolves CS kernel to the scale μ_0, which is used to compare different extractions. Apart of the ratio $\sum_1 / \sum_2$ all terms in equation 8 are perturbative, and nowadays known up to N^3LO. Therefore, the formula 8 can be used to extract CS kernel directly from the data without any further approximation.
Practically, the experimental measurements for differential cross-sections are presented by a collection of points in bins of (Q, y, q_T). Therefore, the transformation 4 cannot be computed analytically but by the discrete Hankel transform. Herewith, one should find a balance between the statistical precision of $d\Sigma$ (which usually decreases at low-q_T) and the range of b (larger b requires lower q_T). Alternatively, the experimental curve can be fit by an analytical form, and the transformation 4 is performed analytically. This path, however, introduces uncertainty due to the curve parameterization.
The integration over q_T leaves intact the dependence on Q and y, which can be used o increase the statistical precision. We introduce

$$\Sigma(s, Q, b) = \int_{Q-\delta Q}^{Q+\delta Q} dQ^2 \int_{\delta y}^{\delta y} dy d\Sigma(s, y, Q, b) \tag{11}$$

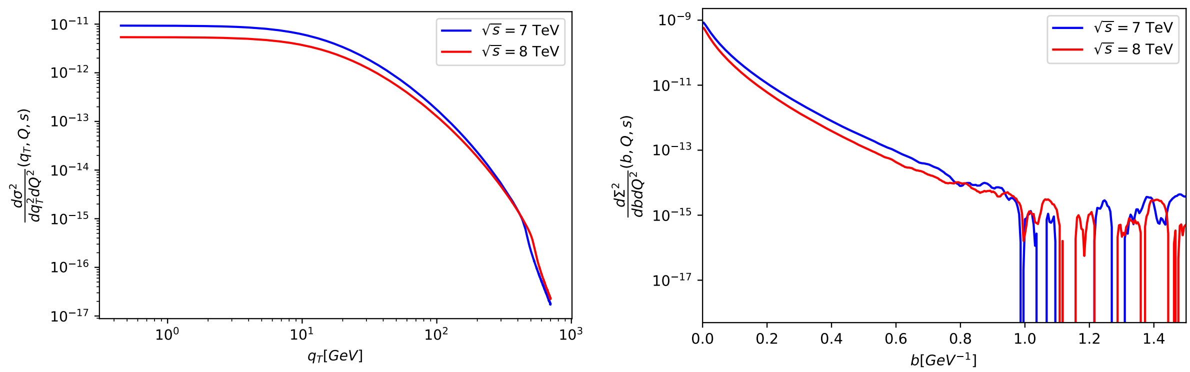
where δQ and δy are sizes of the bin. These function can be also used in the ratio Σ_1/Σ_2 with the only restriction that $\delta Q << Q$. In this case, the effects of variation of Q within the bin could be neglected. There is no limitation for δy, except that δy is the same for Σ_1 and Σ_2 .

3 Analysis of the results

To test the proposed approach, we study the pseudo-data generated with the CASCADE event generator. The PB algorithm does not explicity employ the TMD factorization theorem. There are no parameters specially dedicated to the CS kernel, and thus the CS kernel emerges via the interplay of the parameters controlling the PB-TMD evolution in the CASCADE generator.
The inverse Hankel transform has been perfomed using the algorithm. The algorithm

expects that the input function vanishes beyond the presented domain. It is a good approximation since the cross-section for the double Higgs process drops rapidly at large-q_T.

The effective range and accuracy of the discrete Hankel transform depend on the density and range of the input cross-section. So, to obtain a stable curve at the large-b, one needs a large number of points at small-q_T. At large values of b, the inverse function is sensitive to the finite-bin effects and becomes unstable. The examples of the cross-section in momentum and position spaces are shown 3.



(a) Differential cross-section defined in the momentum space

(b) Differential cross-section defined in the position space

Figure 3: Differential cross-section for double Higgs process, for $\sqrt{s} = 6 - 8TeV$

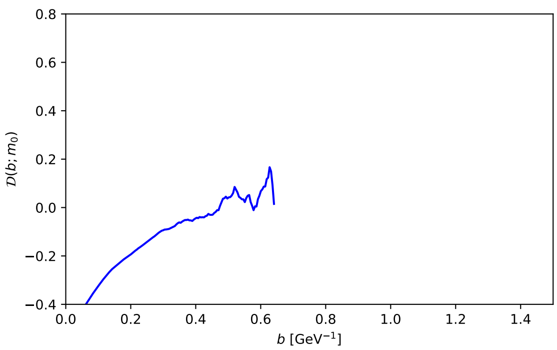
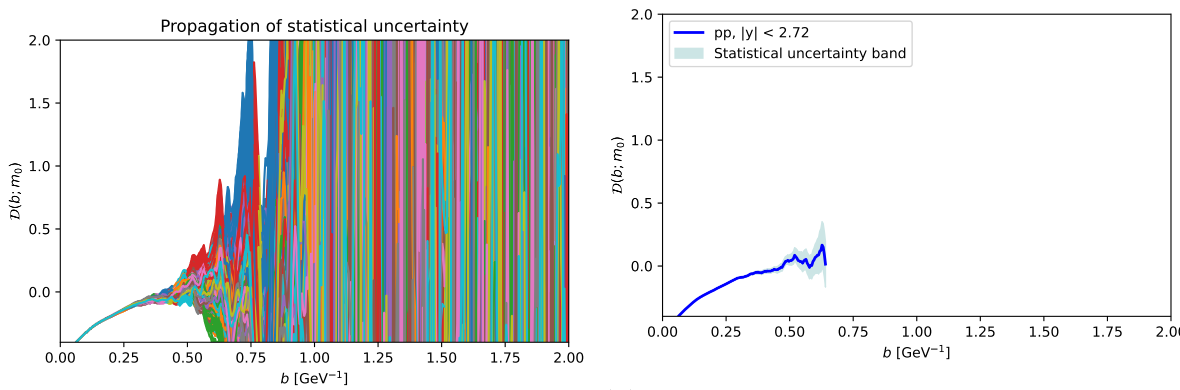


Figure 4: CS kernel determination for gluon iniciated process

Figure 4 shows the CS kernel function for double Higgs process. We can notice the function becomes unstable for large values of b, due to, as we mentioned before, the inverse Hankel/Fourier transform. We also see how the functions ends, for $b \sim 0.75GeV^{-1}$. It seems like the model restricts the shape of $\mathcal{D}$ for large values of b. The function also suggest the existence of a flat region, certain saturation point where b changes its value "faster" than $\mathcal{D}$, of course all this right before the instability region.

3.1 Statistical Uncertainty of the CS kernel function

We employ the boostrap method to estimate the propagation of statistical uncertainty from the momentum space to the position space (see figure 5). For example, consider the ratio $r = a/b$ between two measurements a and b, perfomed with partially overlapping data. The nominal measuremente r_0 is performed using the nominal dataset, while a series of boostrap measurements are perfomed using an ensemble of pseudo-experiments. A replica measurement $r_i = a_i/b_i$ is performed using replica datsets i. In such replica measurements, the shared events have the same fluctuations away from the nominal dataset, which will affect the measurements coherently: in the same direction if the measurements are positively correlated or in opposite direction if they are anti-correlated. The distribution of measurements r_i can be treated as the probability distribution function of r and an uncertainty on this quantity can be derived from this distribution, from its standard deviation if appropriate [5]. During the sampling, we also vary the central value of q_T within a bin, which allows us to estimate the uncertainty due to finite bin size at large-b.



(a) Replica measurements

(b) CS kernel for double Higgs production process with its statistical uncertainty band

Figure 5: Statistical uncertainty of the CS kernel function

4 Conclusions

In this work, we tested a new method of direct extracting of the CS kernel from the data, using the proper combination of cross-section with different kinematics, for gluon TMD distribution functions. We chose pair production of Higgs bosons processes to develop our research. Although, some considerations were taken into account to simplify the analysis, the model calculation puts a serious restriction on the shape of the CS kernel at large values of b. Still, figure 6 shows that this method can be use to evaluate the CS kernel behavoir for gluon iniciated processes.

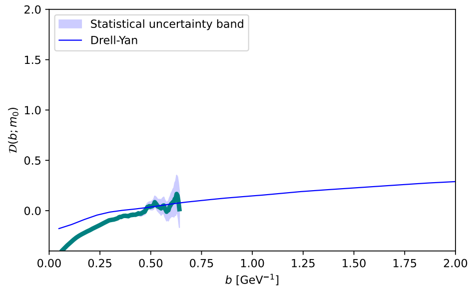


Figure 6: Comparison between the CS kernel function for Drell-Yan process vs double Higgs production process

References

[1] J.C.Collins, Davison E. Soper. Transverse momentum distribution in Drell-Yan pair and W and Z boson production. Nuclear Physics B250(1985) 199-224. 1984.

[2] Alexey A. Vladimirov. Self-contained definition of the Collins-Soper kernel. arXiv:2003.02288v2 [hep-ph]. 2020.

[3] Armando Bermudez Martínez, Alexey A. Vladimirov. Determination of Collins-Soper kernel from cross-sectios ratios. arXiv:2206.01105v1 [hep-ph]. 2022.

[4] The CMS Collaboration. Combination of seaches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13TeV$. arXiv:1811.09689v2 [hep-ex]. 2019.

[5] The ATLAS Collaboration. Evaluating statistical uncertainties and correlations using the bootstrap method. arXiv:1811.09689v2 [hep-ex]. 2019.