



Study of parton distributions in π^- within the `xFitter` framework

Ivan Novikov, Joint Institute for Nuclear Research, Dubna

Supervisor: Alexander Glazov

September 5, 2018

Abstract

In this work the `xFitter` tool is used to determine parton distribution functions in pions using data from E615 experiment, which studied Drell-Yan $\mu^+\mu^-$ -production in π^- -beam scattering on a tungsten target. Theoretical predictions are calculated up to next-to-leading order. A nuclear PDF set is used for the tungsten target to account for nuclear corrections. The extracted valence quark distributions with estimated error bands are presented. Obtained distributions are compared to available pion PDFs from LHAPDF library.

Contents

1	Introduction	3
2	Experimental data	3
3	Theoretical predictions	3
4	Fit methodology	5
5	Results	7

1 Introduction

Parton distribution functions (PDFs) remain a crucial component for description of hadron collisions. PDFs of the proton and related hadrons — antiproton and neutron, — are currently known with high precision thanks to HERA deep inelastic scattering and LHC data. On the other hand, much less is known about PDFs of light mesons — the charged pion and the kaon. The goal of this work is to study PDFs of charged pions using `xFitter`. While `xFitter` (former `HERAFitter`) is a mature open-source fitting framework and played a significant role in determination of proton PDFs, it could not be used to study PDFs of mesons. As such, this work also includes necessary code developments, which were committed to the `xFitter` repository.

2 Experimental data

The present analysis is based on Drell-Yan cross-section data measured by the E615 experiment[1]. The experiment studied scattering of pion beam with energy $E_\pi = 252\text{GeV}$ on a tungsten target. The data are available at `hepdata.net`¹. Events with a muon pair in the final state were selected:

$$\pi^- + \tau_6 W \rightarrow \mu^+ \mu^- + X$$

The cross-section is reported differentially with respect to dimensionless observable kinematic variables x_F and $\sqrt{\tau}$, defined as:

$$x_F = \frac{2p_L}{\sqrt{s}} \qquad \sqrt{\tau} = \frac{m_{\mu\mu}}{\sqrt{s}},$$

where $m_{\mu\mu}$ is the invariant mass of the muon pair, p_L is the longitudinal component of momentum of the muon pair in the center-of-mass frame. For a 252GeV pion center-of-mass energy is $\sqrt{s} = 21.8\text{GeV}$.

No information on bin-to-bin correlations in the data is available. Consequently, all data errors are treated as uncorrelated.

One may notice that the available kinematic region includes energies of Υ -resonances. The measured cross section in their region is expected to differ from Drell-Yan predictions. Therefore, this region is excluded from the analysis. In terms of $\sqrt{\tau}$, this corresponds to exclusion of bins between 0.415 and 0.484.

3 Theoretical predictions

In the following a fit procedure is performed for the pion PDF. Each iteration of the fit procedure requires a calculation of theoretical prediction of cross sections, which is

¹it should be noted that one of points, in table 14, was copied from the paper to `hepdata` incorrectly

performed according to the main formula of leading-twist QCD

$$d\sigma(Q^2, x_F, \sqrt{\tau}) = \sum_{\substack{i,j \in \\ \text{flavors}}} \int \int_0^1 dx_1 dx_2 f_i^\pi(x_1, Q^2) f_j^{\text{target}}(x_2, Q^2) d\sigma_{ij \rightarrow \mu\mu}(x_1, x_2, Q^2, x_F, \sqrt{\tau})$$

Here pion PDFs $f_i^\pi(x_1, Q^2)$ are parameterised as described in section 4, and are varied to minimize χ^2 .

The LHAPDF set `nCTEQ15FullNuc_184_74`[2] was used as the PDF f_j^{target} of nucleons in the tungsten target. As this PDF set was obtained specifically for a tungsten target, no additional nuclear corrections are necessary. Tungsten PDFs are kept constant during the fit.

The perturbative part of cross section $d\sigma_{ij \rightarrow \mu\mu}$ is independent of PDFs, does not change between iterations, and therefore can be calculated only once. This idea is realized in `APPLgrid` library [3], which calculates cross sections by convoluting PDFs with precalculated coefficients. The latter are calculated only once and stored in a grid file.

As perturbative part of cross sections depends on dimensionless kinematic observables x_F and $\sqrt{\tau}$, a different set of these coefficients has to be used for each point of the experimental data. While `APPLgrid` is able to use coefficients that depend on a kinematic variable, it only allows one variable, while this analysis requires two. To overcome this limitation, we let `APPLgrid` handle x_F - dependence, and treat set of datapoints with different values of $\sqrt{\tau}$ as independent datasets, and use a separate `APPLgrid` grid file for each dataset.

In order to generate a grid file one needs to integrate cross section of partonic subprocess over all kinematic variables, except the observed x_F and $\sqrt{\tau}$. A modified version of the `MCFM` [4] generator was used for this purpose. It should be noted that `MCFM` cannot be used for pion-on-tungsten collisions, and was used to simulate proton-proton collisions instead. This, however, does not pose a problem, as only hadron-independent information on partonic subprocesses was extracted from `MCFM`.

Although the final analysis is performed at NLO, we began by verifying the predictions of `APPLgrid` with LO grids by comparing them to our own calculation, based on leading-order formula:

$$\frac{d^2\sigma}{dx_\pi dx_N}(x_\pi, x_N) = \frac{4\pi\alpha_e^2}{9s} \sum_i Q_i^2 (f_i^\pi(x_\pi) \bar{f}_i^N(x_N) + \bar{f}_i^\pi(x_\pi) f_i^N(x_N))$$

For this initial test an existing pion PDF, namely `GRVPI0`[5]² from LHAPDF[6], was used. The results of these two methods were found to be in good agreement.

Since leading-order formulae are very simple it was possible to prepare LO grids via a single run of `MCFM` on a personal laptop. NLO calculations, on the other hand, are significantly more complicated, and their integration requires so much more Monte-Carlo simulation time, that running it on a single machine is no longer practical. Instead, grids

²due to an unusual convention, u and \bar{u} flavor numbers in this set are -2 and 2, instead of the other way around

were generated on DESY batch computing cluster BIRD. Of 1000 jobs submitted via HTCondor 986 finished successfully, others failed due to various network and filesystem errors. The generated 986 sets of grids were merged into one using a tool provided with APPLgrid.

In order to decide if the number of Monte-Carlo events was sufficient, as well as to perform a test of the generation procedure, grids consistency check is performed. The grids are convoluted with PDFs of protons that were used by MCFM and the resulting cross-sections are compared to reference cross-sections, which were calculated by MCFM. For the final grids used in the analysis the relative difference is at least an order of magnitude smaller than experimental errors, and smaller than 1% for a majority of bins.

To sum up, the theoretical predictions for each of the datapoints are obtained by APPLgrid convoluting a grid with the tungsten target PDF and the fitted pion PDF. It should be noted that some code modifications to the corresponding xFitter theory module were required to support a target described by a separate non-fitted PDF.

4 Fit methodology

In order to describe the pion structure by a small number of parameters, a number of assumptions is made. Negative pion quark composition in the quark model is $(d\bar{u})$, therefore, neglecting small $SU(3)$ -breaking effects, $d = \bar{u}$ and $u = \bar{d} = s = \bar{s}$ at the QCD evolution starting scale. Heavy quark densities are set to zero at the starting scale. In the end, there are three independent distributions to parameterise — valence v , sea S and gluon g .

Valence and sea distributions in the negatively charged pion are defined as:

$$v = \frac{d_v - u_v}{2} = d - u$$

$$S = \frac{u + \bar{d}}{2} = u,$$

where

$$u_v = u - \bar{u} \qquad d_v = d - \bar{d}$$

The distribution functions in pion must satisfy the valence and momentum sum rules:

$$\int_0^1 u_v dx = -1 \quad \int_0^1 d_v dx = 1$$

$$\int_0^1 x(u + \bar{u} + d + \bar{d} + s + \bar{s} + g) dx = 1$$

Under the assumptions above, the two valence sum rules are equivalent and reduce to only one condition. Or, in terms of the parameterised distributions:

$$\int_0^1 v dx = 1 \qquad \int_0^1 x(2v + 6S + g) dx = 1$$

As these definitions and sum rules differ significantly from their proton counterparts, they had to be implemented as a separate `xFitter` module.

The parameterisation at the QCD evolution starting scale are chosen as follows

$$v = \Lambda_1 x^{B_v} (1-x)^{C_v} \quad (1)$$

$$S = A_S x^{B_S} (1-x)^{C_S} \quad (2)$$

$$g = \Lambda_2 x^{B_g} (1-x)^{C_g} \quad (3)$$

Here Λ_1 and Λ_2 are not independent parameters, but are constrained by the valence and momentum sum rules, respectively. The starting scale $Q_0^2 = 1.9\text{GeV}^2$ was chosen to be a little bit below the charm mass threshold.

A choice of particular functional form for parameterisations of PDFs may introduce a bias to their determination. In order to control for this problem, a series of fits with a more free form of distribution v were performed. During these fits, the parameterisation was modified to

$$v = \Lambda_1 x^{B_v} (1-x)^{C_v} (1 + A_1 \sqrt{x} + A_2 x + A_3 x^{\frac{3}{2}} + A_4 x^2 + \dots)$$

with additional free parameters A_i added one-by-one. However, the decrease in χ^2 per additional parameter was less than 1, which means that this relaxed functional form does not improve fit quality. Consequently, the final fit was performed using the basic form 1.

During preliminary fit attempts the sea and gluon parameters were left free. However, it was discovered that these parameters can vary within a wide range with little to no effect on χ^2 , as can be seen in figure 1. This result shows that the data are not sensitive to sea and gluon distributions. A decision was made to fix sea and gluon parameters and focus on the valence distribution. Following the idea described in [7] we fix the shape of sea and gluon distributions to be similar to corresponding distributions in proton, using HERAPDF NLO set as reference:

$$B_S = 8 \quad C_S = -0.2 \quad B_g = 0 \quad C_g = 9$$

Initial fits produced a very weak constraint on the sea normalization parameter $A_S = 0.0818 \pm 0.0618$. In the following fits A_S was fixed at this value.

As mentioned above, the tungsten PDF is fixed during the fitting procedure. Uncertainties in this PDF might affect the resulting pion PDF. In order to estimate these uncertainties, we repeated the analysis using error PDF sets, that are provided with `nCTEQ15FullNuc_184.74`.

Statistical errors of number of events in a bin are typically estimated as a square root of number of events. As measured number of events fluctuates arounds its mean value, so does the estimate of statistical error of measurements. Measurements with random fluctuation down are assigned a lower statistical error, and, therefore, have a higher weight in calculation of χ^2 . This introduces an overall negative bias. In order to correct

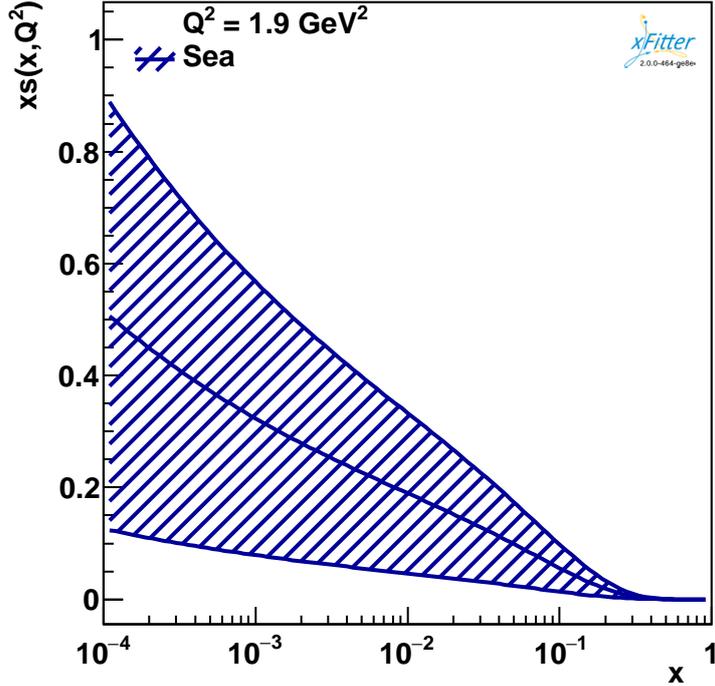


Figure 1: Sea $S = u = \bar{d} = s = \bar{s}$ distribution in pion at the starting scale. Error band indicates uncertainty at 60% confidence level, corresponding to $\Delta\chi^2 = 1$

for this bias, rescaling of data errors is performed. The following adjusted definition of χ^2 is used:

$$\chi^2 = \sum_i \left(\frac{D_i - T_i}{\tilde{\sigma}_i} \right)^2 \quad \tilde{\sigma}_i = \sqrt{\frac{T_i}{D_i}} \sigma_i, \quad (4)$$

where D_i , σ_i , T_i are data, error and calculated prediction for bin i , respectively.

An iterative Lagrange Multiplier method implemented in `xFitter`[8] was used to build error bands for pion PDF.

5 Results

B_v	C_v	$\chi^2/N_{DoF}(N_{points})$	Comment
0.6817 ± 0.013	0.9794 ± 0.023	$209.10/138 = 1.53$	main fit
0.6782 ± 0.013	0.9789 ± 0.023	$164.46/135 = 1.22$	3 points-outliers excluded
0.7006 ± 0.014	0.9623 ± 0.023	$469.60/166 = 2.83$	Υ region included
		$424.02/140 = 3.03$	GRVPI0 pion PDFs

Table 1: Results of the fit

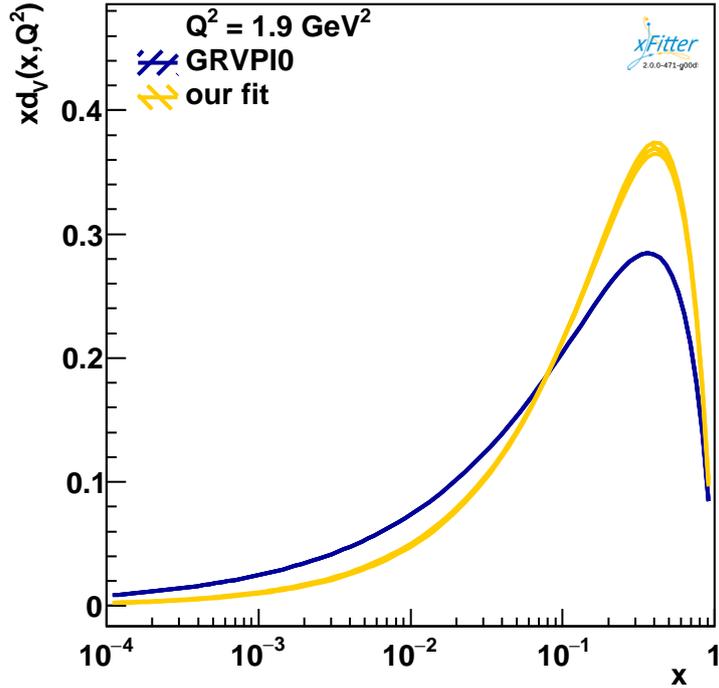


Figure 2: d -valence distribution at the starting scale for our fit (yellow line) compared to GRVPI0 set (blue line). For our fit, uncertainties are shown with a band.

Results of the final fit are presented in table 1. Careful examination of datapoints shows three outliers that have a significant contribution to χ^2 . These points can be seen in figure 3. Their exclusion results in improved χ^2 , while the values of fitted parameters do not change significantly.

As described above, region containing Υ -resonances was excluded from the main fit. Including these points in the fit leads to a significantly higher χ^2 .

For comparison, χ^2 was also calculated using LHAPDF set GRVPI0 for the pion. The obtained distribution and GRVPI0 are compared in figure 2. The two distributions appear to be in poor agreement, although uncertainties of GRVPI0 distribution are unknown.

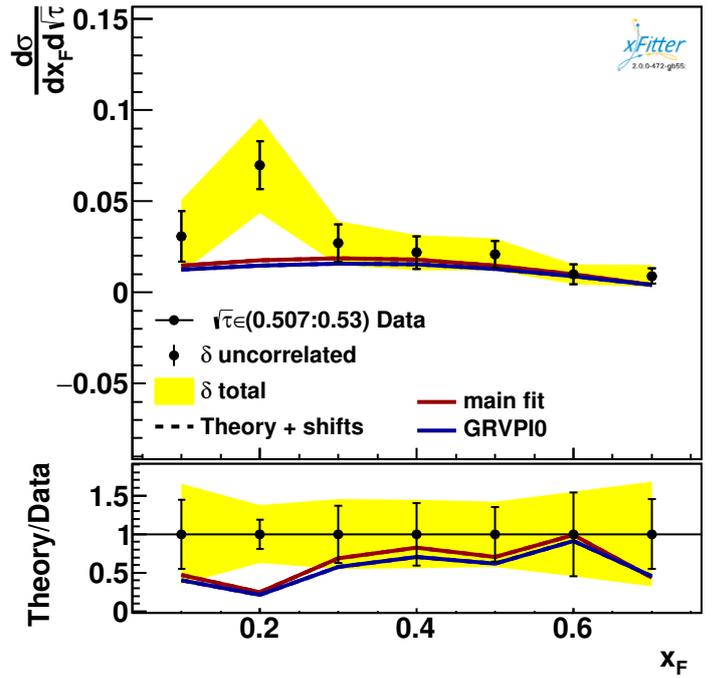
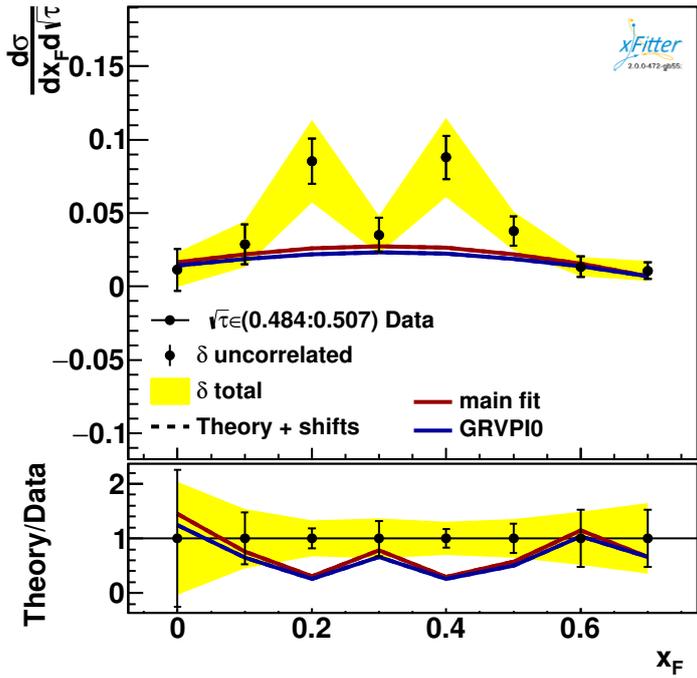
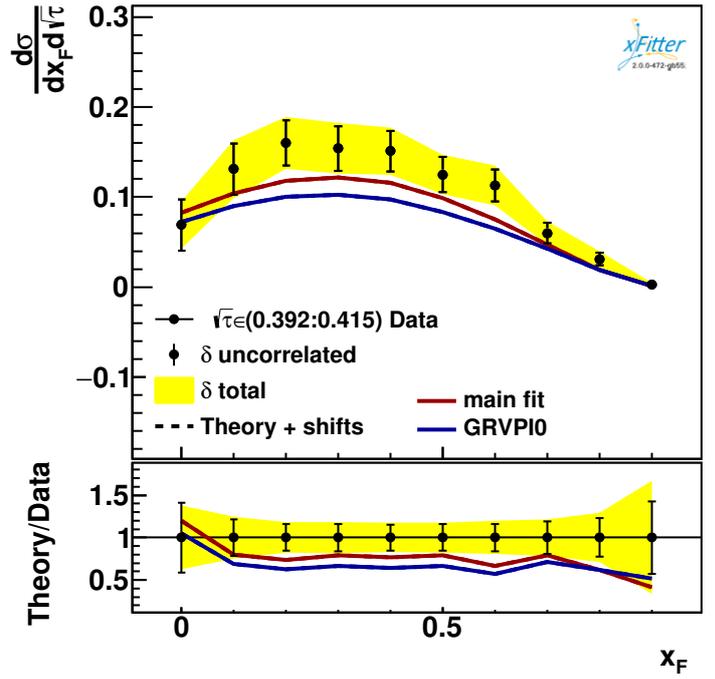
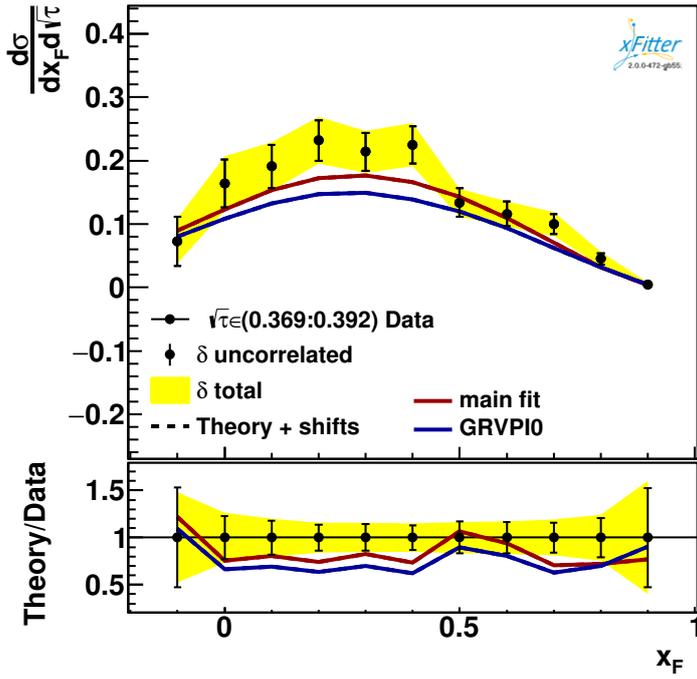


Figure 3: Experimental data and theory predictions in our fit compared to predictions obtained using GRVPI0 set. Yellow bands show original data uncertainties, error bars show uncertainties rescaled according to 4. Notice the three outliers in the two bottom plots. For illustration, only a partial region of $\sqrt{\tau}$ range is shown.

References

- [1] John Steven Conway et al. Experimental Study of Muon Pairs Produced by 252-GeV Pions on Tungsten. *Phys. Rev.*, D39:92–122, 1989.
- [2] Karol Kovarik et al. nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework. *Phys. Rev.*, D93(8):085037, 2016.
- [3] Tancredi Carli, Dan Clements, Amanda Cooper-Sarkar, Claire Gwenlan, Gavin P. Salam, Frank Siegert, Pavel Starovoitov, and Mark Sutton. A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project. *Eur. Phys. J.*, C66:503–524, 2010.
- [4] John M. Campbell and Richard Keith Ellis. An Update on vector boson pair production at hadron colliders. *Phys. Rev.*, D60:113006, 1999.
- [5] Moshe Gluck, Ewald Reya, and Andreas Vogt. Pionic parton distributions. *Z. Phys.*, C53:651–656, 1992.
- [6] Andy Buckley, James Ferrando, Stephen Lloyd, Karl Nordström, Ben Page, Martin Rfenacht, Marek Schnherr, and Graeme Watt. LHAPDF6: parton density access in the LHC precision era. *Eur. Phys. J.*, C75:132, 2015.
- [7] Moshe Gluck, Ewald Reya, and Ingo Schienbein. Pionic parton distributions revisited. *Eur. Phys. J.*, C10:313–317, 1999.
- [8] Jonathan Pumplin, Daniel Robert Stump, and Wu-Ki Tung. Multivariate fitting and the error matrix in global analysis of data. *Phys. Rev.*, D65:014011, 2001.