PROCEEDINGS OF SCIENCE

Parametrization sampling and the pion PDF in a phenomenological analysis

Aurore Courtoy[a,*]

[a] *Instituto de Física, Universidad Nacional Autónoma de México, Apartado Postal 20-364, 01000 Ciudad de México, Mexico*

E-mail: aurore@fisica.unam.mx

In these proceedings, we extend the discussion of the pion PDF obtained by NLO QCD analysis in the Fantômas4QCD framework. Our pion analysis uses a state-of-the-art statistical methodology that accounts for the epistemic uncertainty. Fantômas4QCD, designed to handle multiple functional forms for solving the inverse problem, systematically explores a variety of solutions for PDFs, thereby improving estimates of epistemic uncertainty. Through this novel approach, we interpret our results for the valence sector, considering various non-perturbative methods available for predicting the pion PDF. We emphasize the distinctions between (global) QCD analyses and QFT-based calculations.

10th International Conference on Quarks and Nuclear Physics (QNP2024)
8-12 July, 2024
Barcelona, Spain

*Speaker

PoS(QNP2024)050

1. Goals of data-based QCD analyses

Parton Distribution Functions (PDFs) are non-perturbative objects. They encapsulate the behavior of quarks and gluons inside hadrons for a given configuration of flavor, spin, and other relevant quantum numbers. PDFs follow a quantum-field theoretic definition, which is formally obtained through factorization theorems, at a given order in perturbation theory. Due to their very nature, PDFs cannot be evaluated from first principles. Since PDFs are of the utmost importance for predicting processes which involve hadron targets and/or beams, these functions have been determined through two main approaches, which we will call *phenomenological extractions* and *low-energy predictions*, nowadays complemented by lattice QCD approaches.

In the former approach, PDFs are extracted from the observables. Factorization theorems tell us how to implement the perturbative, calculable parts of a given observable. Then, the extraction of the PDFs is performed by solving an inverse problem, sketched as follows for structure functions

$$F(x_{\rm B}, Q^2) = \sum_a \int_{x_{\rm B}}^1 \frac{dx}{x} f_{a/p}(x, \mu^2)\, H_a\left(\frac{x_{\rm B}}{x}, \frac{\mu^2}{Q^2}\right) + O(M/Q) \ , \tag{1}$$

with H_a the hard part of the observable, and $f_{a/p}(x)$ the PDF for a parton a carrying a longitudinal momentum fraction x of a hadron p. DGLAP equations relate distribution functions at different resolution scales, μ^2. Higher orders in perturbative QCD can be implemented in both factors, *i.e.* H_a and $f_{a/p}$. Some higher-twist corrections are contained in the last term, $O(M/Q)$. From the factorization point of view, all of the unknown, long-distance behavior is embedded into the PDF, up to power-suppressed corrections. Since this methodology starts from observables to extract knowledge on a function by solving an inverse problem, we call it a top-down approach. The PDFs determined in the top-down approach are then used to make predictions for further observables. In the case of the proton PDFs, this is of high relevance for hadron colliders, such as the LHC, where PDFs are the baseline of most calculations. More recently, global analyses of the pion PDFs have been resumed, thanks to theoretical and computational efforts of the past two decades. The pion PDFs shed light on the meson structure formation, the key aspect of understanding the low-energy regime of QCD. In these proceedings, we will discuss the top-down approach to the pion PDF realized by the Fantômas4QCD project [1], complementing the panorama of data-based, top-down analyses of the pion PDFs by the JAM [2] and the xFitter collaborations [3].

On the other hand, low-energy predictions provide an evaluation of the QFT definition of the PDFs,

$$f_{a/p}(x, Q_0^2) = \int \frac{dz^-}{4\pi} e^{-i(x-\frac{1}{2})P^+z^-} \left\langle \pi^+(p) \right| \bar{\psi}_a\left(0, \frac{y^-}{2}, \mathbf{0}\right) \gamma^+ \psi_a\left(0, -\frac{y^-}{2}, \mathbf{0}\right) \left| \pi^+(p) \right\rangle \ , \tag{2}$$

here for the π^+ and in the light-cone gauge ($A^+ = 0$), incorporating (state-of-the-art) calculations of dressed propagators and vertex functions that reflect both the pseudo-Goldstone and bound-state nature of the pion, within accessible hypotheses, *e.g.*, [4–8]. The variable Q_0^2 is a low, hadronic scale at pre-factorization values, where the degrees of freedom of the non-perturbative calculation are expected to connect to the RGE spectrum in the $\overline{MS}$ scheme. We call this the bottom-up approach.

The results from the bottom-up approach are frequently compared with those from the top-down approach. This raises the question of whether such a comparison is meaningful. This issue will be revisited shortly.

2. Fantômas in a nutshell

In the top-down narrative, we, the practitioners, face an inverse problem that involves the determination of a function of x from a finite number of discrete data points, through a convolution. This problem admits more than one solution for the PDFs achieving a good agreement with the data. In order to investigate the impact of multiple solutions, the Fantômas4QCD project was created, based on previous studies of functional mimicry [9] and representative sampling [10] in the context of PDFs. We parametrize the PDFs as the product of a carrier function, which describes the asymptotics, and a modulator, which we take to include a Bézier curve of degree N_m, $\mathcal{B}^{(N_m)}$,

$$x\, f_i(x, Q_0^2) = A_i x^{B_i} (1-x)^{C_i} \left[1 + \mathcal{B}^{(N_m)}(y(x))\right] . \tag{3}$$

The novelty in our metamorph function, Eq. (3), is that the Bézier curve is directly related to its values P_j at control points x_j, i.e. $P_j = \mathcal{B}^{(N_m)}(y(x_j))$, with $y(x)$ a function of x. If we use $N_m + 1$ such control points, the vector of coefficients of the polynomial can uniquely be determined by a simple matrix equation, see Ref. [1] and references therein. The metamorph captures the modulations in the PDF due to the shift of the control points. By varying the position and the number of fixed control points or the function $y(x)$, we are able to generate, on the fly, many different functional forms for the *l.h.s.* of Eq. (3).

This innovative parametrization was used to fit the pion PDF using the xFitter framework. xFitter already contained the pion-induced Drell-Yan data as well as prompt photon's, to which we have added a minimal set of leading-neutron data. The Drell-Yan data largely constrains the large-x valence and sea quark pion PDF, while the other two bring more information on the sea and the gluon, though not enough presently to disentangle them. The Fantômas framework allows, for the first time, for the accounting of the statistical (*aleatoric*) uncertainties coming from the data but also the systematics that now include methodological choices and the quality of various samplings, known as *epistemic* uncertainties. That is, sampling over the space of solutions for the function of interest contributes to the *epistemic uncertainty*. On the *l.h.s.* of Fig. 1, we show the resulting combination of the 5 most diverse solutions, after ~ 100 trials filtered through soft constraints, for the valence, sea and gluon PDFs of the pion.

3. The valence sector of the pion PDF– a critical view

Having tried approximately a hundred different functional forms, we can conclude on the size of the uncertainties in given regions of the (x, Q^2) plane. The very large-x and moderate Q^2 region can be compared against early-QCD predictions. Counting rules suggested a large-x behavior of the pion quark PDF of $\lim_{x\to 1} f_V^{\pi}(x) \propto (1-x)^{\beta}$ with $\beta \approx 2$, in the scaling region. This expectation may be modified by various radiative contributions at large momentum fractions that affect the interpretation of realistic measurements [9]. In this regard, the Fantômas4QCD analysis did not qualitatively differ from other recent analyses: the fall-off of the valence PDF at large x

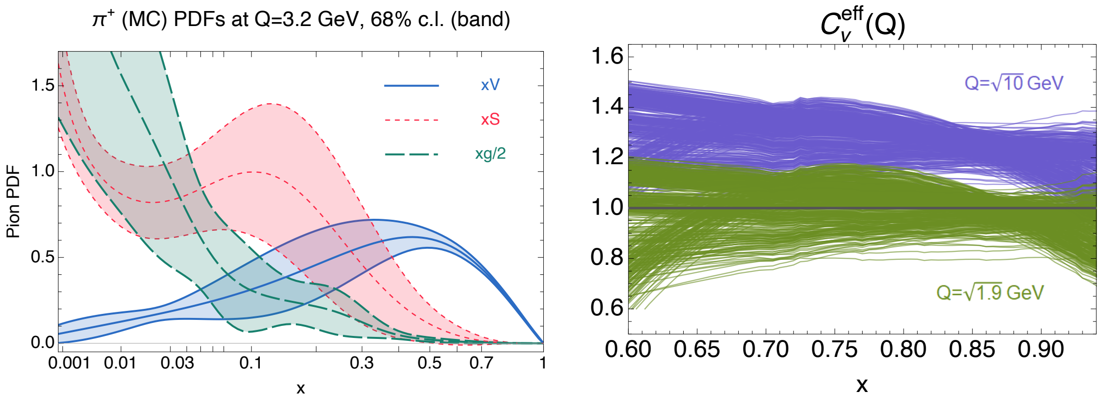

Figure 1: Left: FantôPDFs for the valence (blue, full curves), the sea (red, dotted curves) and gluon (green, dashed curves) at $Q = \sqrt{10}$ GeV. Right: The effective $(1-x)$ exponent of the valence PDF in the FantôPDF ensemble at $Q_0 = \sqrt{1.9}$ GeV (green), and at $\sqrt{10}$ GeV (blue).

is compatible with $\beta = C_V^{\text{eff}} = 1$ at $Q_0 = \sqrt{1.9}$ GeV, in spite of the multiple functional forms that have been considered (Fig. 1 right). Threshold effects could affect the extraction of the pion PDF at large x. The JAM collaboration investigated various treatments of resummation and found that β could significantly vary, and chose to adopt the result that leads to $\beta \approx 1$, motivated by the most accepted resummation treatment [2]. Hence comes the question of comparison of predictions with data-driven analyses.

In Ref. [9], we argued that polynomial mimicry implies that there is no sufficient condition to confirm that one specific non-perturbative representation of a PDF is the true one when compared to a phenomenologically extracted polynomial, or even compared against data. A necessary condition is that the curves agree according to an agreed-upon metric, such as χ^2. For practical use of the bottom-up results, the hadronic scale Q_0^2 must be determined. It is customary to fix it by comparison of the *l.h.s.* of Eq. (2) with observables (see, *e.g.* Ref. [11]) or moments obtained in global analyses (see, *e.g.* Ref. [12]), by using the backward DGLAP evolution. Since Q_0^2 turns out to be of a few hundred MeV, this procedure relies on pushing the validity of DGLAP evolution to extremely low scales. Once the hadronic scale is determined, it is also used to make predictions on functions or observables, using forward DGLAP evolution. Predictions made using this method do not have a straightforward interpretation.

For illustration purposes, we present in Fig. 2 the central values of naïvely evolved bottom-up results alongside the comprehensive global top-down QCD analysis. In this figure, both the Nambu–Jona-Lasinio model [5, 6] (labeled “NJL") and the Dyson-Schwinger inspired analysis of Ref. [7] (labeled “DSE") are evolved from a very low hadronic scale ($Q_0 = 0.29$ and 0.33 GeV, respectively) to $Q = \sqrt{10}$ GeV using the LO DGLAP evolution as implemented in Refs. [14, 15]. Thus, in Fig. 2, $xV(x, Q_0)_{\text{bottom-up}}$ serves solely as an initial condition for the LO DGLAP equations. Additionally, we show the central value of the hybrid analysis using Light-Front Wave Functions [13] (labeled “MAP LFWA"), which is a NLO analysis providing its own LHAPDF grids.

The interpretation of Fig. 2 is as follows. The MAP LFWA free parameters are determined from fitting the Drell-Yan and prompt-photon data of xFitter's NLO framework. The mid- and

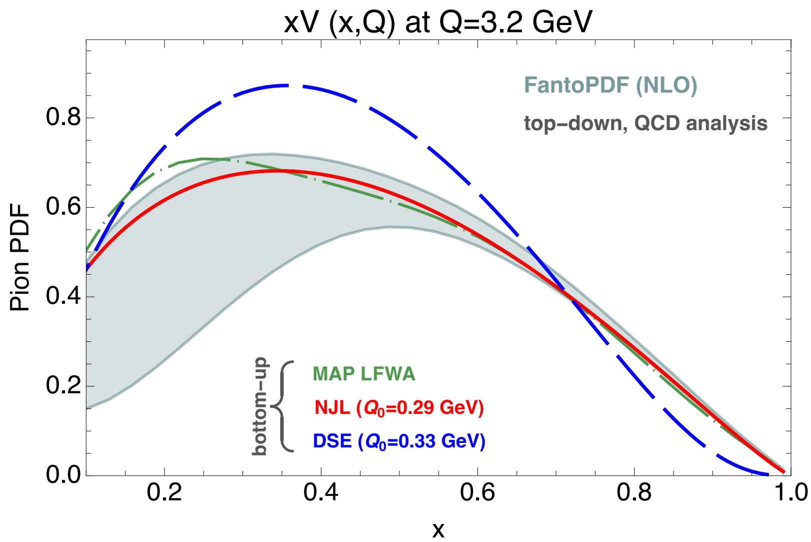

Figure 2: The valence of the pion at $Q = \sqrt{10}$ GeV, in linear scale. The central value of MAP LFWA result is shown in green (long-dashed–dotted curve), the NJL result evolved for ($Q_0 = 0.29, \Lambda_{\rm LO} = 0.2$) GeV, in red (full curve), and the DSE result of Ref. [7] for ($Q_0 = 0.33, \Lambda_{\rm LO} = 0.2$) GeV, in blue (dashed curve), and compared to the Fantômas PDFs, shown in dark cyan (solid band).

large-x behavior of the central value of MAP valence PDF follows that of the FantôPDFs's. As anticipated, the predictions of NJL and DSE cannot be directly compared to the data-driven, fixed-order $\overline{MS}$ PDFs of the Fantômas analysis. The uncertainty on the formers is unknown, due to the non-perturbative nature of the approaches. What the red and blue curves reflect is the evolution of a flat and a quadratic PDF as initial conditions, respectively, with LO DGLAP equation. Variations of the value of Q_0, in our working hypotheses, do not allow for a transition from one initial condition to the other.

4. Conclusions

In this contribution to the proceedings, we have discussed the recent results on the PDFs of the pion with epistemic uncertainties. A representative sampling of the parametrization, a solution to the inverse problem of interest, increases the size of the uncertainties of the valence, sea and gluon PDFs in most of the x range. However, since the Drell-Yan data constrains the large-x region greatly, our fixed-order analysis agrees with previous extractions in the limit $x \to 1$.

We commented on the difference between global QCD top-down analyses and non-perturbative bottom-up approaches. For phenomenological PDFs to meet the predictions from non-perturbative methodologies, further studies must be undertaken.

Acknowledgments

AC would like to thank P. Nadolsky for fruitful discussions, S. Noguera for providing the code for the DGLAP evolution of Ref. [14], as well as useful discussions on the NJL results, J. Segovia for useful exchanges, and S. Venturini for providing the MAP LFWA [13] grids. AC is supported by the UNAM Grant DGAPA-PAPIIT IN111222 and CONACyT Ciencia de Frontera 2019 No. 51244 (FORDECYT-PRONACES).

References

[1] L. Kotz, A. Courtoy, P. Nadolsky, F. Olness and M. Ponce-Chavez, Phys. Rev. D **109** (2024) no.7, 074027 doi:10.1103/PhysRevD.109.074027 [arXiv:2311.08447 [hep-ph]].

[2] P. C. Barry *et al.* [Jefferson Lab Angular Momentum (JAM)], Phys. Rev. Lett. **127** (2021) no.23, 232001 doi:10.1103/PhysRevLett.127.232001 [arXiv:2108.05822 [hep-ph]].

[3] I. Novikov, H. Abdolmaleki, D. Britzger, A. Cooper-Sarkar, F. Giuli, A. Glazov, A. Kusina, A. Luszczak, F. Olness and P. Starovoitov, *et al.* Phys. Rev. D **102** (2020) no.1, 014040 doi:10.1103/PhysRevD.102.014040 [arXiv:2002.02902 [hep-ph]].

[4] M. B. Hecht, C. D. Roberts and S. M. Schmidt, Phys. Rev. C **63** (2001), 025213 doi:10.1103/PhysRevC.63.025213 [arXiv:nucl-th/0008049 [nucl-th]].

[5] R. M. Davidson and E. Ruiz Arriola, Acta Phys. Polon. B **33** (2002), 1791-1808 [arXiv:hep-ph/0110291 [hep-ph]].

[6] L. Theussl, S. Noguera and V. Vento, Eur. Phys. J. A **20** (2004), 483-498 doi:10.1140/epja/i2003-10174-3 [arXiv:nucl-th/0211036 [nucl-th]].

[7] M. Ding, K. Raya, D. Binosi, L. Chang, C. D. Roberts and S. M. Schmidt, Phys. Rev. D **101** (2020) no.5, 054014 doi:10.1103/PhysRevD.101.054014 [arXiv:1905.05208 [nucl-th]].

[8] J. Lan, C. Mondal, X. Zhao, T. Frederico and J. P. Vary, [arXiv:2406.18878 [hep-ph]].

[9] A. Courtoy and P. M. Nadolsky, Phys. Rev. D **103** (2021) no.5, 054029 doi:10.1103/PhysRevD.103.054029 [arXiv:2011.10078 [hep-ph]].

[10] A. Courtoy, J. Huston, P. Nadolsky, K. Xie, M. Yan and C. P. Yuan, Phys. Rev. D **107** (2023) no.3, 034008 doi:10.1103/PhysRevD.107.034008 [arXiv:2205.10444 [hep-ph]].

[11] F. A. Ceccopieri, A. Courtoy, S. Noguera and S. Scopetta, Eur. Phys. J. C **78** (2018) no.8, 644 doi:10.1140/epjc/s10052-018-6115-3 [arXiv:1801.07682 [hep-ph]].

[12] M. Stratmann, Z. Phys. C **60** (1993), 763-772 doi:10.1007/BF01558408

[13] B. Pasquini *et al.* [MAP (Multi-dimensional Analyses of Partonic distributions)], Phys. Rev. D **107** (2023) no.11, 114023 doi:10.1103/PhysRevD.107.114023 [arXiv:2303.01789 [hep-ph]].

[14] M. M. Block, Eur. Phys. J. C **65** (2010), 1-7 doi:10.1140/epjc/s10052-009-1195-8 [arXiv:0907.4790 [hep-ph]].

[15] K. J. Golec-Biernat and A. D. Martin, Phys. Rev. D **59** (1999), 014029 doi:10.1103/PhysRevD.59.014029 [arXiv:hep-ph/9807497 [hep-ph]].