CONSTRAINING PDFs AND nPDFs WITH RECENT DATA*

Petja Paakkinen

University of Jyväskylä, Department of Physics
P.O. Box 35, 40014 University of Jyväskylä, Finland
and
Helsinki Institute of Physics
P.O. Box 64, 00014 University of Helsinki, Finland

Received 28 July 2022, accepted 20 September 2022, published online 14 December 2022

The progress in constraining proton and nuclear parton distribution functions is briefly summarised. Some persistent uncertainties are pointed out and recent experimental advancements highlighted.

DOI:10.5506/APhysPolBSupp.16.1-A24

1. Introduction

In many applications of high-energy hadronic collisions, including those of heavy nuclei, the partonic structure of hadrons manifests itself through the collinear factorisation of QCD [1]. That is, the cross section for producing an inclusive final state $k+X$, where the X denotes "anything", can be described in terms of coefficient functions (CF) $\mathrm{d}\hat{\sigma}^{ij\to k+X'}$, which are calculable from perturbative QCD, and parton distribution functions (PDFs) f_i^A, f_j^B, which contain long-range physics and thus cannot be obtained by perturbative means, plus "higher twist" (HT) corrections which are suppressed at high enough momentum-transfer scale $Q \gg \Lambda_{\mathrm{QCD}}$. Formally,

$$\begin{aligned}&\mathrm{d}\sigma^{AB\to k+X}\left(Q^2\right)\\ &\overset{Q\gg\Lambda_{\mathrm{QCD}}}{=}\sum_{i,j,X'} f_i^A\left(Q^2\right)\otimes\mathrm{d}\hat{\sigma}^{ij\to k+X'}\left(Q^2\right)\otimes f_j^B\left(Q^2\right)+\mathcal{O}\left(\frac{1}{Q^2}\right),\end{aligned}\quad(1)$$

where "$\otimes$" indicates a multiplicative (Mellin) convolution over x, the fraction of the parents momentum carried by the parton.

* Presented at the 29th International Conference on Ultrarelativistic Nucleus–Nucleus Collisions: Quark Matter 2022, Kraków, Poland, 4–10 April, 2022.

The PDFs $f_i^A(x, Q^2)$, with i denoting the parton flavour and A the parent hadron or nucleus, exhibit two important properties: First, they are universal, independent of the partonic process (*i.e.* the CFs). Second, their scale dependence is governed by the DGLAP equations $\partial f_i^A / \partial \ln Q^2 = \sum_j P_{ij} \otimes f_j^A$ with perturbatively calculable splitting functions P_{ij}. For these properties, it is possible to extract the PDFs from global analyses of experimental data of multiple observables at different scales.

Figure 1 shows results from such analyses with the CT18A next-to-leading order (NLO) proton PDFs [2] and EPPS21 NLO nuclear PDFs (nPDFs) [3], where the latter uses the former as its free-proton baseline, taken as representative examples. As it is common, both analyses fit six independent parton flavours, $u, d, \bar{u}, \bar{d}, s = \bar{s}$, and glue, taking heavy-quark PDFs fully perturbatively generated[1]. The general features are evident from this plot: For protons, most flavours are already quite well constrained in a wide range of x. The exception to this rule is the strange quark, which is currently the least constrained flavour in both free-proton and nuclear PDF analyses. For nPDFs, the uncertainties persist to be larger compared to the free-proton ones, particularly in the small-x region, where direct data constraints are scarce.

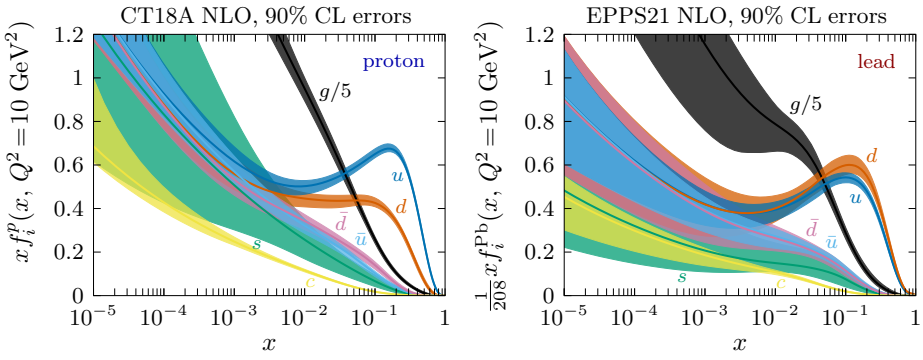


Fig. 1. (Colour on-line) A representative example of proton and nuclear PDFs taken from the CT18A NLO proton PDFs [2] and EPPS21 NLO nPDFs [3]. The coloured bands show the estimated PDF uncertainties at 90% confidence level.

2. Progress in proton-PDF analyses

The most recent proton-PDF analyses use several thousand data points from various experiments. Such a global approach is a necessity: a multi-observable fit is needed in order to constrain individual flavours and a

[1] Some analyses, like the recent NNPDF4.0 extraction [4], extend this set to fit also intrinsic charm and strange asymmetry.

multi-experiment fit in order to do so in a wide range of x. For example, fixed target deep-inelastic scattering (DIS) and Drell–Yan (DY) data are important in setting the large-x quark distributions (valence-to-sea and flavour separation). While much of these data are already rather old, new data are still coming from Fermilab and JLab, where *e.g.* the new SeaQuest Collaboration measurement of proton–deuteron to proton–proton ratio of DY dilepton production [5] improves our understanding of the antimatter asymmetry in the proton. To reach smaller values of x, collider DIS data are needed, and nowadays the HERA inclusive and heavy-quark data [6, 7] form the backbone of any modern proton-PDF analysis. With the large x, Q^2 lever arm, these data are able to constrain also the gluon distribution through the DGLAP equations. Recently, however, most of the new data have come from hadron colliders (in particular from the CERN LHC), granting access to a wide spectrum of new processes such as the production of electroweak (EW) bosons, jets, top-pairs, *etc.*, allowing for testing the factorisation and PDF flavour dependence in new kinematic windows.

One of the most important latest developments in this field has been the inclusion of next-to-next-to-leading-order (NNLO) corrections [8, 9] to the perturbative jet-production calculations in the most recent global analyses[2]. These jet observables, together with the $t\bar{t}$ production, are important in setting the large-x gluon distributions [4]. The NNLO precision improves fit quality particularly for the LHC (EW boson and jet) data and, for example, the recent MSHT20 analysis [12] reports on an improvement of 700 units in χ^2_{tot} for the 4363 global data points. Therefore, with the present LHC data precision, NNLO proton-PDFs are no longer a mere formal improvement, reducing theoretical uncertainties from missing-higher-order corrections, but a necessity for an accurate fit.

LHC data can also challenge some of the earlier approaches taken in global analyses. Traditionally, proton-PDF fits have included neutrino–nucleus (νA) DIS data for improved strange-quark constraints, which have led to suppressed strangeness *e.g.* in the ABMP16 fit [13], as is quantified by the fraction

$$K_s = \frac{\int_0^1 \mathrm{d}x\, x\,[s(x) + \bar{s}(x)]}{\int_0^1 \mathrm{d}x\, x\,\left[\bar{u}(x) + \bar{d}(x)\right]} \tag{2}$$

shown in Fig. 2. Complementary data from ATLAS EW-boson production have then confronted this view with a preference for unsuppressed strange (see ATLAS-epWZ16 in Fig. 2) [14, 15]. Finding some tension between these two data types, the CT18 [2] analysis provides two different fits, where the nominal CT18 set does not include the ATLAS data, but the CT18A

[2] At leading colour, that is. Calculations at full colour have been performed only very recently [10, 11].

version does. Later, a simultaneous fit was found to be feasible when NNLO charm-quark mass corrections in the νA cross sections were accounted for, as in the nNNPDF4.0 and MSHT20 analyses [4, 12].

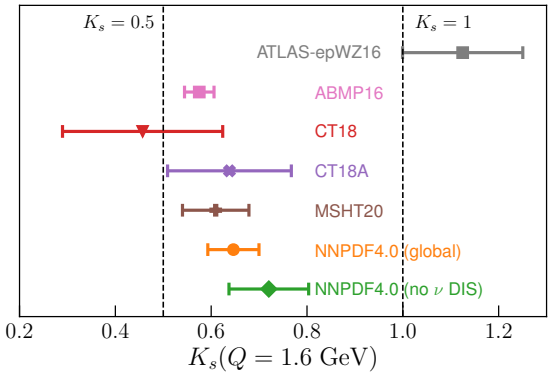


Fig. 2. Strangeness suppression in different proton-PDF analyses. Figure from Ref. [16].

All of this points to the direction that understanding the role of nuclear corrections in fitting proton PDFs with nuclear data is becoming increasingly important. Indeed, the recent NNPDF4.0 analysis finds very different large-x sea-quark behaviour depending on whether the uncertainties from nNNPDF2.0 nPDFs were included or not [4], in contrast to some earlier evaluations [17]. How to take these nuclear corrections into account varies from analysis to analysis, the MSHT20 taking them from the DSSZ nPDFs [18] plus an additional three-parameter fit [12], whereas CT18 uses a scale-independent phenomenological parametrization as described in Ref. [19].

3. Progress in nPDF analyses

The nPDFs are fitted with similar global analyses as their free-proton counterparts, relying only on the QCD collinear factorisation and thus giving a model-agnostic way to study the nuclear effects. Unlike in proton-PDF analyses, where NNLO is the standard, most of the nPDF analyses are still performed at NLO and those recent analyses which also provide NNLO fits [20, 21] are performed with a rather limited number of data types and thus cannot parametrise the flavour dependence very freely. Those (NLO) analyses standing on a more global footing instead include a variety of data from the LHC proton–lead (p+Pb) runs, with EPPS21 [3] and nNNPDF3.0 [22] also propagating the uncertainties from their free-proton baselines.

A very important advancement in this field has been the inclusion of the double-differential LHC p+Pb dijet and D^0-production data in the recent analyses, leading to strong new constraints on the nuclear gluon PDFs [3, 22] which have been previously poorly known in lack of collider DIS data on nuclei[3]. In accordance with the earlier reweighting study [24], these analyses obtain a good fit to the dijet data, except in the very forward bins where the predictions overshoot the data. This could be due to the missing data correlations, but also larger than expected NNLO or non-perturbative corrections could play a role. For the D^0-production, nNNPDF3.0, with its POWHEG+PYTHIA approach, finds a large scale uncertainty in the nuclear ratio and therefore fit only to the more constraining forward data, whereas EPPS21, with the S-ACOT-m_{T} mass scheme [25], finds no large theory uncertainty in this ratio in the fitted $p_{\mathrm{T}} > 3$ GeV region [26]. Complementary gluon constraints can be derived from single-inclusive production of π^0, $\pi^{\pm}$, and $K^{\pm}$ as in the nCTEQ15WZSIH analysis [27] fitting to a large set of data from PHENIX, STAR, and ALICE measurements. In such fits, like with the D^0-production, one is sensitive to the fragmentation functions and their associated uncertainties, which however partially cancel in the nuclear ratios.

In addition, $W^{\pm}$ production and the most precise CMS Collaboration measurement from Run 2 [28] in particular, have been already included in almost every recent nPDF fit. The approaches taken in doing so, however, vary, TUJU21 [21], nNNPDF3.0 [22], and nCTEQ15WZSIH [27] fitting to absolute cross sections, whereas EPPS21 [3] using nuclear-modification ratios to cancel proton-PDF uncertainties [29]. Also new Z-boson data from Run 2 recently became available [30] and have been studied in several nPDF analyses, all finding some difficulties in describing the data. While the discrepancy in the small-invariant-mass bin can be explained by NNLO corrections [21, 22], the on-peak deviations appear to be due to a rather abrupt change in the shape of the data at midrapidity, possibly due to untypically large data fluctuations [3].

The resulting gluon distributions from the EPPS21, nNNPDF3.0, and nCTEQ15WZSIH analyses [3, 22, 27] are compared in Fig. 3. All of these nPDFs exhibit small-x gluon shadowing and mid-x antishadowing in lead, which are now established from multiple experimental observables. In finer details, the results however differ. At small x, EPPS21 and nNNPDF3.0 agree extremely well due to the use of D^0 and dijet data, though with rather different uncertainty estimations, whereas nCTEQ15WZSIH with the main small-x constraints coming from EW-boson production favouring smaller shadowing. At large x, nNNPDF3.0 shows a different behaviour from EPPS21 which now agrees better with nCTEQ15WZSIH instead. This is possibly

[3] The heavy-quark data were recently considered also in Ref. [23].

associated with the PHENIX π^0 data included in EPPS21 and nCTEQ15-WZSIH and also with the exclusion of backward D^0 data from nNNPDF3.0. It should be also noted that the discussed direct gluon constraints come only from data on heavy nuclei and therefore the gluon distributions in lighter nuclei, such as oxygen, remain poorly constrained [31].

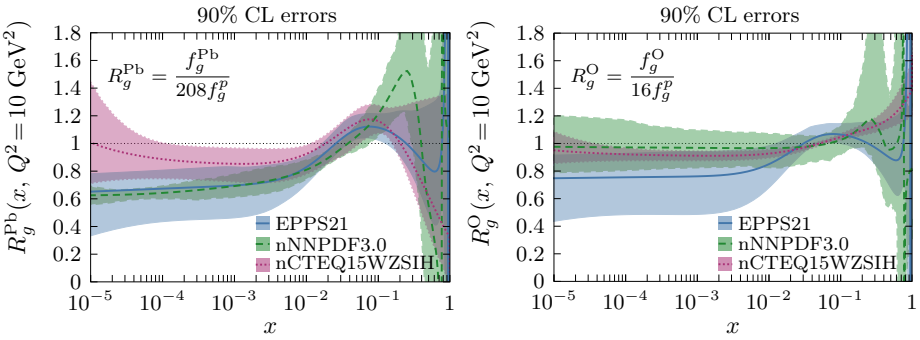


Fig. 3. Nuclear modifications of the gluon PDFs in lead (left) and oxygen (right) nuclei. Results from Refs. [3, 22, 27].

The progress does of course not end here. In Ref. [32], the advantages of measuring triple-differential dijet production in p+Pb were discussed. Since the triple-differential measurement fixes partonic kinematics at leading order, this would give a powerful test of factorisation and PDFs. Also the new π^0-production small-system scan from PHENIX [33] with improved precision and access to higher p_{T} can help in deciphering the nuclear modifications at large x. Contrary to nPDF expectations, the measured "Cronin peak" size follows an ordering $^3\mathrm{He} + \mathrm{Au} < d + \mathrm{Au} < p + \mathrm{Au}$ and Ref. [33] discusses the cause of these effects in terms of multiple-scattering (*i.e.* HT) and flow-like effects. In this conference, also new LHCb measurements of D^0s [34] and π^0s [35] at 8.16 TeV and charged hadrons at 5.02 TeV [36] in p+Pb were presented, probing smaller values of x. The hopes for using exclusive J/ψ photoproduction in constraining nPDFs were also elevated with the first phenomenological implementation of the NLO corrections in ultrapheripheral Pb+Pb [37]. A more direct probe of the collinearly factorised nPDFs in ultrapheripheral events is the inclusive dijet photoproduction, for which the ATLAS measurement at 5.02 TeV has now been finally fully unfolded [38].

4. Conclusion

The new data constraints from the LHC have brought the proton PDFs (and increasingly also the nPDFs) to the point where taking NNLO corrections into account is a necessity for a good fit, and the p+Pb data have also put unprecedented constraints on the gluon nPDFs. There is also ongoing work to

understand the (cross)correlations between proton and nuclear PDF analyses, and many new data sets (with the LHC Run 3 also underway) to be exploited in the future.

I thank T.J. Hobbs and P.M. Nadolsky for clarifying the use of nuclear corrections in CT18. Financial support from the Academy of Finland, project 330448, is acknowledged. This research was funded as part of the Center of Excellence in Quark Matter of the Academy of Finland and as part of the European Research Council project ERC-2018-ADG-835105 YoctoLHC. Figures 1 and 3 use protanopia- and deuteranopia-safe colour schemes from Refs. [39] and [40].

REFERENCES

[1] J.C. Collins, D.E. Soper, G.F. Sterman, *Adv. Ser. Direct. High Energy Phys.* **5**, 1 (1989).

[2] T.-J. Hou *et al.*, *Phys. Rev. D* **103**, 014013 (2021).

[3] K.J. Eskola, P. Paakkinen, H. Paukkunen, C.A. Salgado, *Eur. Phys. J. C* **82**, 413 (2022).

[4] R.D. Ball *et al.*, *Eur. Phys. J. C* **82**, 428 (2022).

[5] J. Dove *et al.*, *Nature* **590**, 561 (2021); Publisher Correction, *ibid.* **604**, E26 (2022).

[6] H1 and ZEUS collaborations (H. Abramowicz *et al.*), *Eur. Phys. J. C* **75**, 580 (2015).

[7] H. Abramowicz *et al.*, *Eur. Phys. J. C* **78**, 473 (2018).

[8] J. Currie, E.W.N. Glover, J. Pires, *Phys. Rev. Lett.* **118**, 072002 (2017).

[9] J. Currie *et al.*, *Phys. Rev. Lett.* **119**, 152001 (2017).

[10] M. Czakon, A. van Hameren, A. Mitov, R. Poncelet, *J. High Energy Phys.* **2019**, 262 (2019).

[11] X. Chen *et al.*, *J. High Energy Phys.* **2022**, 25 (2022).

[12] S. Bailey *et al.*, *Eur. Phys. J. C* **81**, 341 (2021).

[13] S. Alekhin, J. Blümlein, S. Moch, R. Placakyte, *Phys. Rev. D* **96**, 014011 (2017).

[14] ATLAS Collaboration (G. Aad *et al.*), *Phys. Rev. Lett.* **109**, 012001 (2012).

[15] ATLAS Collaboration (M. Aaboud *et al.*), *Eur. Phys. J. C* **77**, 367 (2017).

[16] J.L. Feng *et al.*, arXiv:2203.05090 [hep-ex].

[17] R.D. Ball, E.R. Nocera, R.L. Pearson, *Eur. Phys. J. C* **79**, 282 (2019).

[18] D. de Florian, R. Sassot, P. Zurita, M. Stratmann, *Phys. Rev. D* **85**, 074028 (2012).

[19] A. Accardi, T.J. Hobbs, X. Jing, P.M. Nadolsky, *Eur. Phys. J. C* **81**, 603 (2021).

[20] H. Khanpour *et al.*, *Phys. Rev. D* **104**, 034010 (2021).

[21] I. Helenius, M. Walt, W. Vogelsang, *Phys. Rev. D* **105**, 094031 (2022).

[22] R. Abdul Khalek *et al.*, *Eur. Phys. J. C* **82**, 507 (2022).

[23] P. Duwentäster *et al.*, *Phys. Rev. D* **105**, 114043 (2022).

[24] K.J. Eskola, P. Paakkinen, H. Paukkunen, *Eur. Phys. J. C* **79**, 511 (2019).

[25] I. Helenius, H. Paukkunen, *J. High Energy Phys.* **1805**, 196 (2018).

[26] K.J. Eskola, I. Helenius, P. Paakkinen, H. Paukkunen, *J. High Energy Phys.* **2005**, 037 (2020).

[27] P. Duwentäster *et al.*, *Phys. Rev. D* **104**, 094005 (2021).

[28] CMS Collaboration (A.M. Sirunyan *et al.*), *Phys. Lett. B* **800**, 135048 (2020).

[29] K.J. Eskola, P. Paakkinen, H. Paukkunen, C.A. Salgado, *Eur. Phys. J. C* **82**, 271 (2022).

[30] CMS Collaboration (A.M. Sirunyan *et al.*), *J. High Energy Phys.* **2105**, 182 (2021).

[31] P. Paakkinen, *Phys. Rev. D* **105**, L031504 (2022).

[32] S. Shen, P. Ru, B.-W. Zhang, *Phys. Rev. D* **105**, 096025 (2022).

[33] PHENIX Collaboration (U.A. Acharya *et al.*), *Phys. Rev. C* **105**, 064902 (2022).

[34] LHCb Collaboration, arXiv:2205.03936 [nucl-ex].

[35] LHCb Collaboration, arXiv:2204.10608 [nucl-ex].

[36] LHCb Collaboration (R. Aaij *et al.*), *Phys. Rev. Lett.* **128**, 142004 (2022).

[37] K.J. Eskola *et al.*, *Phys. Rev. C* **106**, 035202 (2022).

[38] ATLAS Collaboration, «Photo-nuclear jet production in ultra-peripheral Pb+Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV with the ATLAS detector», Conference note atlas-conf-2022-021, CERN, 2022. Accessed July 28, 2022.

[39] B. Wong, *Nat. Methods* **8**, 441 (2011).

[40] P. Tol, «Colour Schemes», Technical Note SRON/EPS/TN/09-002 issue 3.2, Netherlands Institute for Space Research, 2021. Accessed July 28, 2022.