Nuclear PDFs Today

Hannu Paukkunen

University of Jyvaskyla, Department of Physics, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland

Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland

E-mail: `hannu.paukkunen@jyu.fi`

I review the current status of global analysis of nuclear parton distributions and discuss some near-future prospects. I also advocate proton-nucleus (p-A) measurements at the LHC with $A \ll 208$ from the viewpoint of nuclear parton distributions.

1. State of the art 2018

A proper description of hard processes in high-energy collisions involving nuclei — whether they take place in laboratory circumstances or in nature — requires knowledge of nuclear parton distributions (PDFs). Table 1 summarizes the latest available global analyses for these objects. The situation has been quite static for almost two years now, the most recent parametrizations being nCTEQ15 [1] and EPPS16 [2]. Datawise, EPPS16 is the most comprehensive including e.g. LHC Run-I p-Pb data on dijets and electroweak bosons, as well as data on neutrino-nucleus deeply-inelastic scattering. On the theory side, the analyses at an NNLO precision are emerging and the heavy-quark effects are being taken into account, i.e. the PDFs are defined in general-mass variable flavour number schemes (GM-VFNS). For a long time, the light-quark flavour dependence was essentially neglected in the nuclear-PDF analysis, but now we are also making progress there and beginning to fold out the nuclear PDFs truly flavour by flavour.

Table 1: Key specifications of contemporary nuclear-PDF analyses

	EPS09 [3]	DSSZ12 [4]	KA15 [5]	NCTEQ15 [1]	EPPS16 [2]
DIS in ℓ^-+A	✓	✓	✓	✓	✓
Drell-Yan in p+A	✓	✓	✓	✓	✓
RHIC pions d+Au	✓	✓		✓	✓
ν-nucleus DIS		✓			✓
Drell-Yan in π+A					✓
LHC p+Pb dijets					✓
LHC p+Pb W, Z					✓
Order in α_s	NLO	NLO	NNLO	NLO	NLO
Q cut in DIS	1.3 GeV	1 GeV	1 GeV	2 GeV	1.3 GeV
datapoints	929	1579	1479	708	1811
free parameters	15	25	16	16	20
error tolerance	50	30	N.N	35	52
proton baseline	CTEQ6.1	MSTW2008	JR09	CTEQ6M-like	CT14NLO
GM-VFNS		✓		✓	✓
flavour separation				valence	valence + sea

Let us have a look on the most recent NLO extractions. To this end, Figure 1 presents a comparison of the PDF nuclear modifications $R_i^A(x, Q^2)$ defined as ratios between the free- and bound-proton PDFs,

$$R_i^A(x, Q^2) \equiv f_i^{\text{proton in nucleus } A}(x, Q^2) / f_i^{\text{free proton}}(x, Q^2), \qquad (1.1)$$

from EPPS16, nCTEQ15 and DSSZ12. **Valence Quarks:** In the EPPS16 and nCTEQ15 analyses the up- and down-valence quarks were independently parametrized. Whereas in EPPS16 the central values for u and d come out mutually similar — thanks e.g. to the neutrino-Pb DIS data — in nCTEQ15 the up valence has a very strong high-x EMC effect, while the down-valence is enhanced at large *x*. Within the uncertainty bands, however, EPPS16 and nCTE15 are still consistent. In the

DSSZ12 analysis there was no flavour freedom, and thus the uncertainties are artificially much smaller. **Sea Quarks:** In the EPPS16 analysis the three light sea-quark flavours (u,d,s) were independently parametrized while in nCTEQ15 and DSSZ there is no such freedom. Thus the EPPS16 uncertainties are larger, but on the other hand there is less bias. The largest uncertainties are there for the strange quarks, for which the constraints are rather scarce. Currently the best constraints for the strange come from the Z-boson production at the LHC where there is around 20% contribution from $s\bar{s}$ scattering at midrapidity, as illustrated in the leftmost panel in Figure 2. **Gluons:** At large x, the nCTEQ15 uncertainty bands are wider than those of EPPS16. This is principally due to the LHC dijet data included in the EPPS16 analysis, leading to a better constrained large-x gluon. At low x, the nCTEQ15 uncertainties are smaller, which is probably just related to the form of the fit functions — I will elaborate more on this below. In DSSZ12 there are practically no nuclear effects in gluons, as the authors assumed nuclear modifications in parton-to-pion fragmentation functions when analyzing the RHIC pion data.
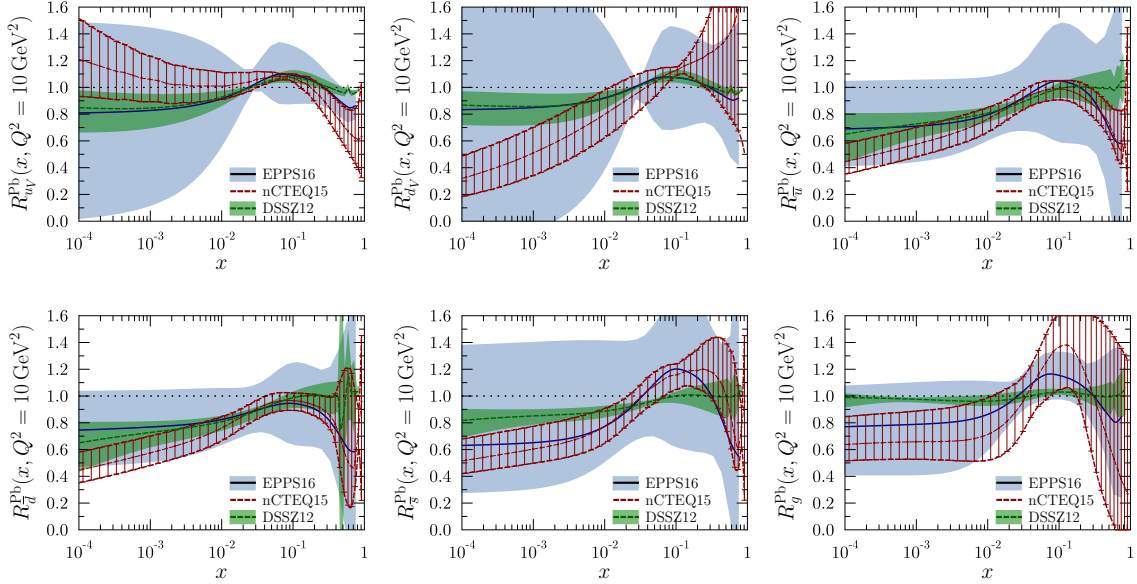

Figure 1: Nuclear modifications from the EPPS16, nCTEQ15 and DSSZ global NLO fits at $Q^2 = 10\,\text{GeV}^2$.

Concerning the functional forms used to fit nuclear PDFs. What we need are certain assumptions for the x and A dependence for $R_i^A(x, Q^2)$ at the parametrization scale Q_0^2 — the behaviour at higher Q^2 follows then from the evolution equations. The assumed functional forms are rather simple and — as has been shown in Refs. [6, 7] — the bias in the current parametrizations is huge, particularly at small x. As can be understood from the right-hand panels of Figure 2, not much variation in small-x behaviour of $R_i^A(x, Q^2)$ is allowed by the current parametrizations. The difficulty is not so much in inventing a flexible ansatz for the x dependence, but to do it in such a way that also the A dependence can be made physically sound ("larger effects for a larger nucleus"). In this sense nuclear PDFs are more difficult to fit than the free proton PDFs. There is also an interesting ongoing effort to fit the nuclear PDFs in a neural-network framework, which should be superior when it comes to reducing the parametrization bias [8].
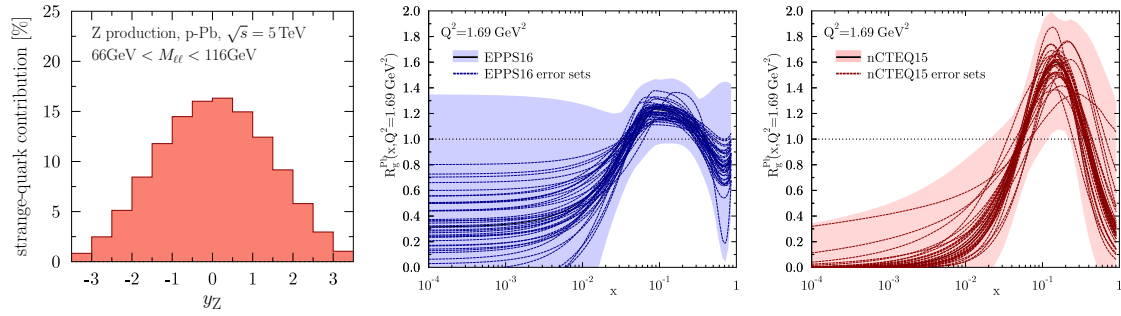
Figure 2: Left: Percentual contribution of $s\bar{s}$ scattering to Z production at $\sqrt{s} = 5\,\text{TeV}$. **Middle:** The EPPS16 gluon nuclear modifications with individual error sets at $Q^2 = 1.69\,\text{GeV}^2$. **Right:** The nCTEQ15 gluon nuclear modifications with individual error sets at $Q^2 = 1.69\,\text{GeV}^2$.
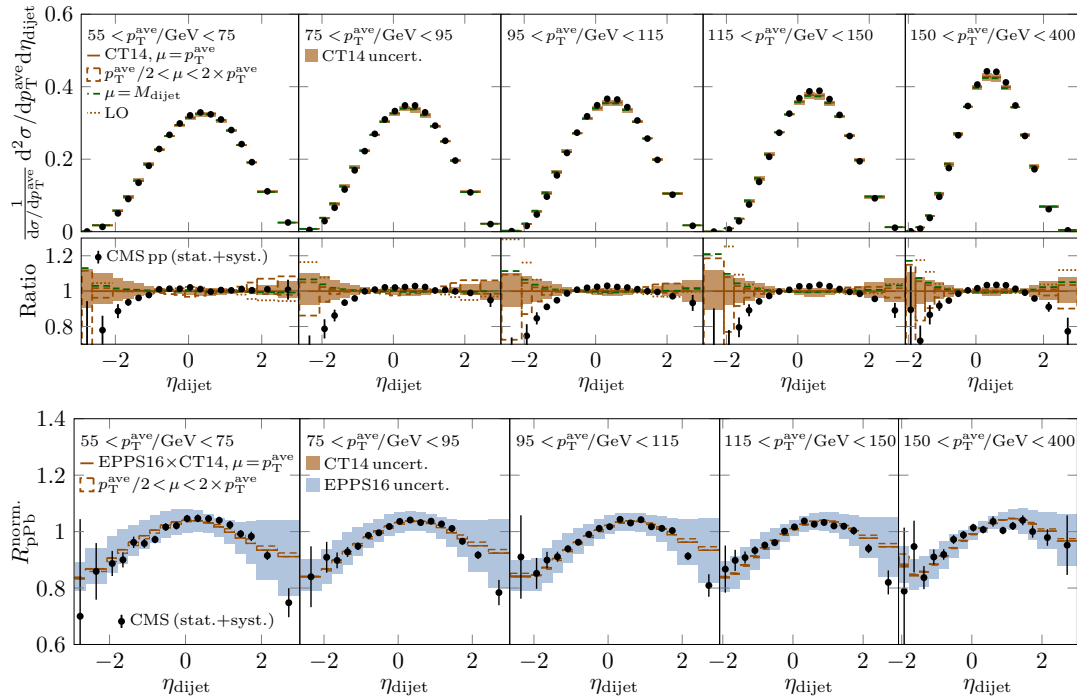

Figure 3: Up: The CMS dijet spectrum in p-p compared with the NLO calculations using CT14NLO proton PDFs [11]. **Down:** Measured dijet nuclear modification compared with EPPS16 predictions [11].

2. New constraints from the LHC

Dijets: One of the most stringent new PDF constraints in the horizon are the CMS 5-TeV dijet data [9]. The upper panels in Figure 3 show the p-p baseline data for the normalized yields

$$\frac{d\sigma^{\text{pp}}(\eta_{\text{dijet}}, p_T^{\text{ave}})}{\int d\sigma^{\text{pp}}(\eta_{\text{dijet}}, p_T^{\text{ave}}) d\eta_{\text{dijet}}} \,,\; \eta_{\text{dijet}} \equiv \frac{\eta_{\text{leading}} + \eta_{\text{subleading}}}{2} \,,\; p_T^{\text{ave}} \equiv \frac{p_T^{\text{leading}} + p_T^{\text{subleading}}}{2} \,,$$

where η_i refers to the jet pseudorapidity. A remarkable thing is that the data cannot be well described within the uncertainties of the current NLO proton PDFs — the theory predictions are way

too broad. This observable, apart from the large $|\eta_{\text{dijet}}|$ tails, should not be too sensitive to the NNLO corrections either, so there seems to be a problem. However, the problem appears to be the same in the p-Pb case, and upon taking the p-Pb/p-p ratio, the outcome is rather well described by EPPS16, as shown in the lower panels of Figure 3. The large-η_{dijet} data appear to prefer a stronger shadowing that is possible within EPPS16. Nevertheless, the data uncertainties are way smaller than the EPPS16 error bands, and in this sense these data promise to have a major impact on the gluon PDFs. Preliminary dijet measurements by ATLAS [10] have also recently appeared.

Z, γ and W$^{\pm}$: On-shell electroweak bosons are clean probes of PDFs at high factorization scales [12, 13, 14]. However, the Run-I data included in the EPPS16 analysis still had a relatively small statistical weight as only a handful of data points was available. For the higher luminosity and higher $\sqrt{s}$, the new 8-TeV data are now significantly more precise and begin to discriminate between the available sets of nuclear PDFs [15]. Also, the data uncertainties are of the same order as the NNLO QCD corrections, motivating the NNLO-level analyses of nuclear PDFs. Now that the yields are higher, it may also be possible to start realistically looking into e.g. the p_{T} distributions of the electroweak bosons which can provide complementary information with respect to the p_{T}-integrated case [16, 17]. The direct photons are also good probes of the gluon density [18]. As the cross sections for lowish p_{T} photons are much much larger than e.g. for Z and W$^{\pm}$, the statistics are high. The limiting factor seems to be currently the systematic experimental uncertainty [19].

Open heavy flavour: The open heavy flavours, D mesons in particular, have recently attracted a growing interest, and we have already rather precise data from various LHC experiments [20, 21, 22]. In particular, the LHCb data [22] at forward direction show a compelling evidence of shadowing at clearly perturbative scales. From the PDF viewpoint the challenge is, maybe a bit surprisingly, in the theoretical treatment for which there are several options, including the fixed-order perturbative QCD [24], FONLL [23], and GM-VFNS frameworks [25, 26]. As most of the general-purpose nuclear PDFs are defined within GM-VFNS, where the heavy quarks treated as partons above the mass thresholds, this framework would sound like the most adequate one. And it is not only that. It is rather well known that the other frameworks, e.g. FONLL and even so-phisticated fixed-order calculations matched to parton showers ("POWHEG+PYTHIA") [27], tend to undershoot the LHC (and other) data by a factor of two or so, though within the large scale uncertainties an agreement can be reached. The GM-VFNS framework, on the other hand, reproduces the data very well and at high p_{T} has a much smaller scale uncertainty than e.g. POWHEG+PYTHIA setup. An example is shown in the left-hand panel of Figure 4.

A very different strategy was adopted in a recent article by Lansberg and Shao [29], where the idea was not to calculate the heavy-flavour cross sections from the first principles but rather to write the cross sections as $d\sigma^{\text{D}^0} = f_g(x_1, Q_f^2) \otimes d\hat{\sigma}_{gg}^{\text{D}^0} \otimes f_g(x_2, Q_f^2)$, parametrize $d\hat{\sigma}_{gg}^{\text{D}^0}$, and fit it to available p-p data. The p-Pb predictions are then obtained by simply replacing the proton PDF by a nuclear PDF. I call this here "matrix-element fitting". As was shown in a follow-up paper [30], one can obtain a very good description of e.g. the D-meson data by this construction. It works actually astonishingly well considering that the x dependence of the assumed squared matrix elements was rather trivial, $\overline{|\mathcal{M}_{gg \to D^0 + X}|^2} = \text{constant} \times x_1 x_2$, and that only the gluon-gluon scattering with leading-order kinematics was considered. Here, I should also point out that it was not just D mesons that were studied, but using the same framework it was possible to describe e.g. the inclusive J/ψ production without invoking any energy loss, multiple scattering or equivalent. However, return-
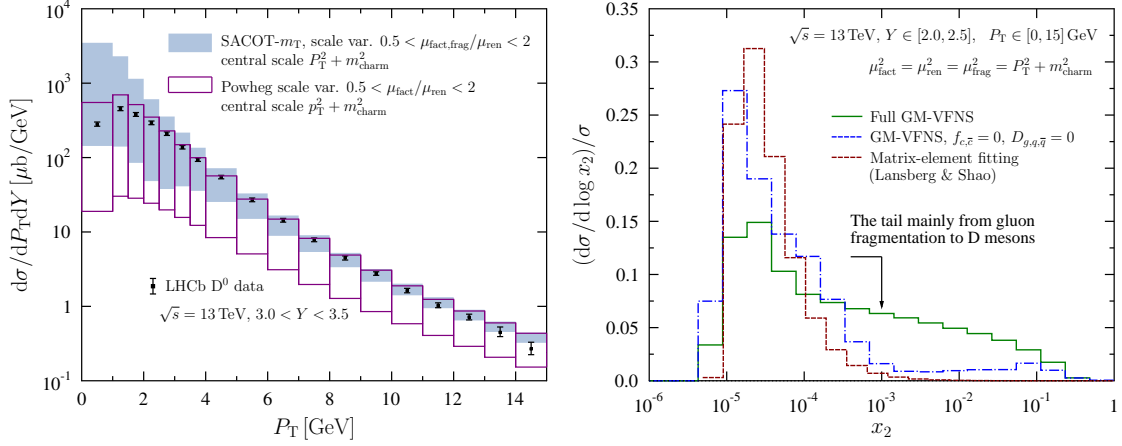
4

Figure 4: Left: Comparison of the LHCb 13-TeV D^0 data [28] with theoretical GM-VFNS and Powheg-method calculations. The error bands represent the scale uncertainties. Figure is from Ref. [26]. **Right:** The x_2 distributions corresponding to LHCb kinematics from GM-VFNS and matrix-element-fitting approaches.
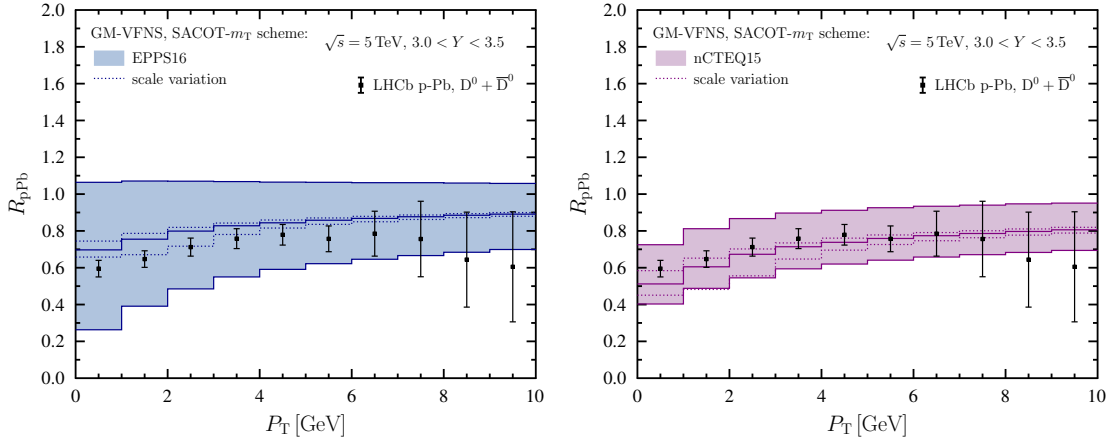

Figure 5: LHCb measurements on the nuclear modification R_{pPb} for D^0 mesons compared with GM-VFNS calculations using EPPS16 (left) and nCTEQ15 (right) [31].

ing to the D-meson case, given the simple assumed form for the partonic cross sections, one may wonder whether it is really able to well enough capture the main features of the full-fledged NLO calculations. In the right-hand panel of Figure 4, we see typical x_2 distributions corresponding to D mesons at forward LHCb kinematics from NLO GM-VFNS, and matrix-element-fitting method. The two distributions are quite different: GM-VFNS predicts a long tail towards large x whereas the ansatz made in the matrix-element fitting yields a distribution which is more small-x concentrated. Actually, the long tail in GM-VFNS comes mainly from the gluon fragmentation contributions [26], as suggested by the curves in Figure 4 with charm-quark PDFs and gluon/light-quark fragmentation contributions turned off. Thus, even for R_{pA}, in which many theory uncertainties cancel, it is not indifferent how the cross section are computed. Of course, the big picture is qualitatively similar in these two frameworks, and also GM-VFNS predicts a significant shadowing at forward direction and the data are consistent with both, EPPS16 and nCTEQ15 as shown in Figure 5.

Pions, h^+h^-: There are also LHC p-Pb data on production of other species like neutral pions [33] or just unidentified charged hadrons [34]. All these are also sensitive to nuclear PDFs. For the

presence of baryons in any h^+h^- sample, the charged-hadrons can be used only at sufficiently high p_T, $p_T \gtrsim 10\,\text{GeV}$ near midrapidity [32] (could be lower at forward/backward directions), and there e.g. the data on foward-backward asymmetry are consistent with NLO predictions [34]. In the case of pions one should be able to go to lower values of p_T. There, e.g. the ALICE data [33] agree with the EPPS16 predictions but disagree with nCTEQ15, thus promising some discrimination power.

Top quarks: The top-quarks have also finally been observed in p-Pb collisions, by CMS [35], and the theory-vs-data pattern for the total cross section looks similar as in p-p. The current statistics are, however, not enough to set significant constraints on nuclear PDFs [36]. Thus, for the moment, the top production remains more a "control" measurement rather than a precision constraint.

Ultraperipheral collisions: The ultraperipheral collisions (UPC) in p-p/p-Pb/Pb-Pb interactions constitute a new avenue of LHC measurements. An interesting observable is the exclusive J/ψ production which at zero-momentum transfer can be related to the square of the nuclear gluon density [37]. The data (e.g. [38, 39]) also agree nicely with predictions based on such expectations. However, since the momentum transfer is not exactly zero in the measurements, the relation to the gluon comes with an uncertainty and for the moment none of the nuclear-PDF fitting groups have dared to consider this observable in their fits. A theoretically more robust UPC observable from the nuclear-PDF view point is the inclusive dijet photoproduction, measured by ATLAS [40]. In principle, the yields should be high enough to get quantitative constraints for the nuclear PDFs. This is illustrated in Figure 6 where a prediction from PYTHIA 8 [41] using EPPS16 PDFs and the ATLAS kinematics is shown. The EPPS16 errorbands are clearly wider than the expected statistical precision with 1nb^{-1} integrated luminosity and one should be able to resolve shadowing, antishadowing and EMC effects. There are several other inclusive UPC processes which could be measured [42] and are sensitive the PDFs, including heavy-flavour and isolated photons.

3. On the case for lighter-than-lead ions

The Run-II LHC p-Pb data — with almost 10 times higher statistics than in Run I — have gradually started to become available. Thus, the global fits of nuclear PDFs will soon be dominated by the measurements involving the lead nucleus (EPPS16 is already). The data on lighter ions come almost exclusively from fixed-target experiments and there is not much variety processwise. In this situation, p-A data with $A \ll 208$ would probably have much more significant impact than, say, running more p-Pb. With lighter ions one can also reach higher luminosities [43] and thus measure rarer processes, or reach the same precision as we have now for p-Pb but in less time. In addition, from the astrophysics view point p-A data with a lighter ion would be useful as e.g. in air the cosmic rays and neutrinos predominantly scatter off from oxygen and nitrogen nuclei. To illustrate a bit how badly the A dependence of PDFs is known, I note that e.g. in the EPPS16 analysis there were only 3 free fit parameters altogether to control the A systematics. Consequently, the bias is presumably significant. The right-hand panels of Figure 6 compare the gluon nuclear modifications for argon and lead as predicted by EPPS16 and nCTEQ15. While in EPPS16 both nuclei look similar, in nCTEQ15 the gluons are even qualitatively very different for these two nuclei. From the nuclear-PDF viewpoint, to constrain the A dependence, it should not matter so much whether it is data on oxygen, argon or some other nucleus, as long as it is clearly smaller than Pb.

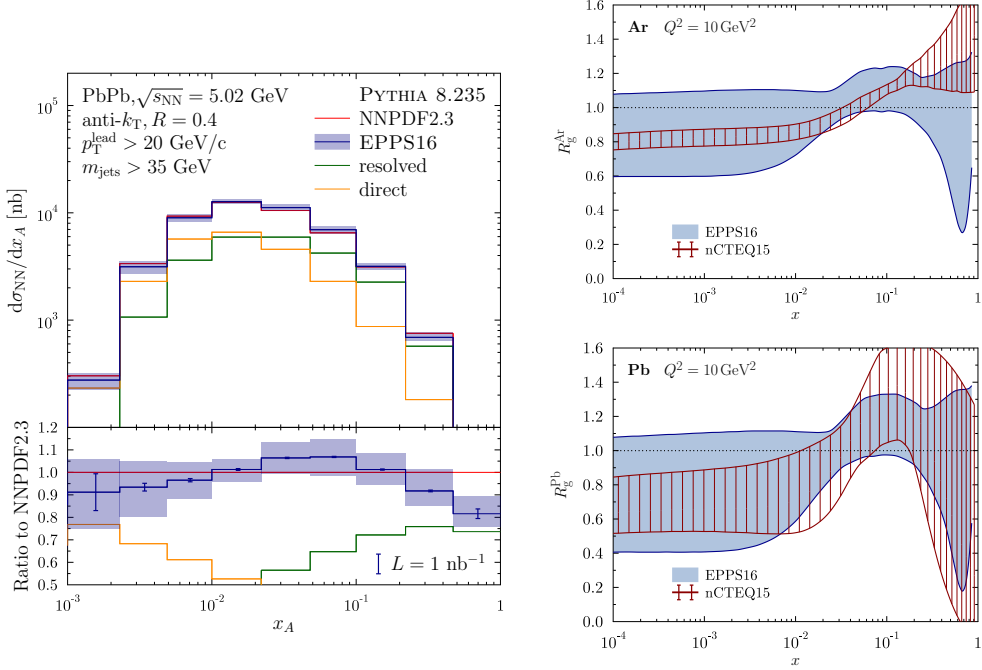**Figure 6: Left:** A PYTHIA simulation of dijet photoproduction in ultraperiheral Pb-Pb collisions (from Ref. [41]). **Right:** Gluonic nuclear effects for argon and lead nuclei from EPPS16 (upper panel) and nCTEQ15 (lower panel) at $Q^2 = 10\,\text{GeV}^2$.

Actually, the large-*x* regime is something one may be able to study also within the fixed-target program of LHCb [44]. For example, there are some D-meson data becoming available [45], which may shed light on the *A* dependence of the gluon at large *x*, though there may be other theoretical issues like a possible intrinsic charm component, for instance. Also, the small $\sqrt{s}$ in these measurements may pose theoretical difficulties.

4. Summary & Outlook

I have shortly described the present status of the global extractions of nuclear PDFs, and discussed some near-future prospects. Thanks to the increasing variety in data input, the earlier restrictions of imposing flavour-independent nuclear effects for the light quarks have been overcome. Also, as the LHC data are becoming increasingly precise, there are now more pressing motivations for the already ongoing efforts to improve the analyses to an NNLO level. The new LHC Run-II data will bring significant new constraints into the game and in many cases the systematic uncertainties can be expected to dominate over the statistical ones. In this situation the availability of the experimental correlations from one bin to another would be quite important. When planning the future LHC runs, to better understand the *A* dependence, p-*A* collisions with some other nucleus than lead would have a significant impact and would be also of great relevance for astrophysics.

Acknowledgments

References

[1] K. Kovarik *et al.*, Phys. Rev. D **93** (2016) 085037.

7

[2] K. J. Eskola, P. Paakkinen, H. Paukkunen and C. A. Salgado, Eur. Phys. J. C **77** (2017) 163.

[3] K. J. Eskola, H. Paukkunen and C. A. Salgado, JHEP **0904** (2009) 065.

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

[5] H. Khanpour and S. Atashbar Tehrani, Phys. Rev. D **93** (2016) 014026.

[6] E. C. Aschenauer *et al.*, Phys. Rev. D **96** (2017) 114005.

[7] H. Paukkunen, PoS DIS **2017** (2018) 109.

[8] R. A. Khalek, J. J. Ethier and J. Rojo, arXiv:1811.05858 [hep-ph].

[9] A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121** (2018) 062002.

[10] The ATLAS collaboration, ATLAS-CONF-2018-050.

[11] Eskola, Paakkinen, Paukkunen, in preparation.

[12] H. Paukkunen and C. A. Salgado, JHEP **1103** (2011) 071.

[13] A. Kusina *et al.*, Eur. Phys. J. C **77** (2017) 488.

[14] P. Ru, S. A. Kulagin, R. Petti and B. W. Zhang, Phys. Rev. D **94** (2016) 113013.

[15] CMS Collaboration, CMS-PAS-HIN-17-007.

[16] V. Guzey, M. Guzzi, P. M. Nadolsky, M. Strikman and B. Wang, Eur. Phys. J. A **49** (2013) 35.

[17] M. Brandt, M. Klasen and F. König, Nucl. Phys. A **927** (2014) 78.

[18] I. Helenius, K. J. Eskola and H. Paukkunen, JHEP **1409** (2014) 138.

[19] The ATLAS collaboration, ATLAS-CONF-2017-072.

[20] S. Acharya *et al.* [ALICE Collaboration], Phys. Lett. B **770** (2017) 459.

[21] The ATLAS collaboration, ATLAS-CONF-2017-073.

[22] R. Aaij *et al.* [LHCb Collaboration], JHEP **1710** (2017) 090.

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

[24] M. L. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. B **373** (1992) 295.

[25] B. A. Kniehl, G. Kramer, I. Schienbein and H. Spiesberger, Phys. Rev. D **71** (2005) 014018.

[26] I. Helenius and H. Paukkunen, JHEP **1805** (2018) 196.

[27] S. Frixione, P. Nason and G. Ridolfi, JHEP **0709** (2007) 126.

[28] R. Aaij *et al.* [LHCb Collaboration], JHEP **1603** (2016) 159.

[29] J. P. Lansberg and H. S. Shao, Eur. Phys. J. C **77** (2017) 1.

[30] A. Kusina, J. P. Lansberg, I. Schienbein and H. S. Shao, Phys. Rev. Lett. **121** (2018) 052004.

[31] Eskola, Helenius, Paakkinen, Paukkunen, in preparation.

[32] D. d'Enterria, K. J. Eskola, I. Helenius and H. Paukkunen, Nucl. Phys. B **883** (2014) 615.

[33] S. Acharya *et al.* [ALICE Collaboration], Eur. Phys. J. C **78** (2018) 624.

[34] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **75** (2015) 237.

[35] A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119** (2017) 242001.

[36] D. d'Enterria, K. Krajczár and H. Paukkunen, Phys. Lett. B **746** (2015) 64.

[37] M. G. Ryskin, Z. Phys. C **57** (1993) 89.

[38] V. Khachatryan *et al.* [CMS Collaboration], Phys. Lett. B **772** (2017) 489.

[39] LHCb Collaboration, LHCb-CONF-2018-003.

[40] The ATLAS collaboration, ATLAS-CONF-2017-011.

[41] I. Helenius, arXiv:1806.07326 [hep-ph].

[42] A. J. Baltz *et al.*, Phys. Rept. **458** (2008) 1.

[43] J. Jowett, Quark Matter 2018.

[44] L. Anderlini [LHCb Collaboration], PoS EPS **-HEP2017** (2017) 152.

[45] R. Aaij *et al.* [LHCb Collaboration], arXiv:1810.07907 [hep-ex].