

EPS09 - Global NLO analysis of nuclear PDFs and their uncertainties

K. J. Eskola, H. Paukkunen*

Department of Physics, University of Jyväskylä and Helsinki Institute of Physics, Finland

E-mail: kari.eskola,hannu.paukkunen@phys.jyu.fi

C. A. Salgado

Departamento de Física de Partículas and IGFAE, Universidade de Santiago de Compostela, Spain

E-mail: carlos.salgado@usc.es

In this talk, we introduce our recently completed next-to-leading order (NLO) global analysis of the nuclear parton distribution functions (nPDFs) called EPS09 — a higher order successor to the well-known leading-order (LO) analysis EKS98 and also to our previous LO work EPS08. As an extension to similar global analyses carried out by other groups, we complement the data from deep inelastic $l + A$ scattering and Drell-Yan dilepton measurements in p+A collisions by inclusive midrapidity pion production data from d+Au collisions at RHIC, which results in better constrained gluon distributions than before. The most important new ingredient, however, is the detailed error analysis, which employs the Hessian method and which allows us to map out the parameter-space vicinity of the best-fit to a collection of nPDF error sets. These error sets provide the end-user a way to compute how the PDF-uncertainties will propagate into the cross sections of his/her interest. The EPS09 package to be released soon, will contain both the NLO and LO results for the best fits and the uncertainty sets.

*4th international workshop High-pT physics at LHC 09
February 4-7, 2009
Prague, Czech Republic*

*Speaker.

1. Introduction

The global analyses of the free-nucleon parton distribution functions (PDFs) are based on three fundamental aspects of QCD: asymptotic freedom, collinear factorization and scale evolution of the PDFs. These properties allow to compute the inclusive hard process cross-sections, schematically

$$\sigma_{AB \rightarrow h+X} = \sum_{i,j} f_i^A(Q^2) \otimes \hat{\sigma}_{ij \rightarrow h+X} \otimes f_j^B(Q^2) + \mathcal{O}(Q^2)^{-n},$$

where f_i s denote the universal, process-independent PDFs obeying the DGLAP [1] evolution equations, and $\hat{\sigma}_{ij \rightarrow h+X}$ are perturbatively computable coefficients. This approach has proven to work extremely well with increasingly more different types of data included in the analyses. In the case of bound nucleons the validity of factorization is not as well established (see e.g. Ref. [2]), but it has nevertheless turned out to provide a very good description of the world data from deep inelastic scattering (DIS) and Drell-Yan (DY) dilepton measurements involving nuclear targets [3, 4, 5, 6, 7, 8]: in the nuclear environment the shape of the PDFs, however, is different from the free-nucleon PDFs. Here, we give an overview of our recent NLO global analysis of the nuclear PDFs (nPDFs) and their uncertainties [9].

2. Analysis method and framework

Our analysis follows the usual global DGLAP procedure:

- **The PDFs are parametrized at a chosen initial scale Q_0^2 imposing the sum rules.** In this work we do not parametrize the absolute nPDFs, but the nuclear modification factors $R_i^A(x, Q_0^2)$ encoding the relative difference to the free-proton PDFs through

$$f_i^A(x, Q^2) \equiv R_i^A(x, Q_0^2) f_i^{\text{CTEQ6.1M}}(x, Q^2). \quad (2.1)$$

Above, $f_i^{\text{CTEQ6.1M}}(x, Q^2)$ refers to the CTEQ6.1M set of free proton PDFs [10] in the \overline{MS} scheme and zero-mass variable flavour-number scheme. We consider three different modification factors: $R_V^A(x, Q_0^2)$ for both u and d valence quarks, $R_S^A(x, Q_0^2)$ for all sea quarks, and $R_G^A(x, Q_0^2)$ for gluons.

- **The absolute PDFs at the parametrization scale Q_0^2 are connected to other perturbative scales $Q^2 > Q_0^2$ through the DGLAP evolution.** In this way, also the nuclear modification factors $R_i^A(x, Q^2)$ become scale-dependent and the initial flavour-independence may also disappear. An efficient numerical solver for the parton evolution is an indispensable tool for any global analysis, but in the case of nPDFs this is even more critical as we always need to perform the evolution separately for 13 different nuclei — an order of magnitude more computation time than what is needed in a free-proton analysis.
- **The cross-sections are computed using the factorization theorem.** In the present analysis the DIS and DY data constitute the main part of the experimental input, but we also employ the midrapidity π^0 -production data measured in d+Au collisions at RHIC. Inclusion of the π^0 -data provides important further constraints for the gluon modification that are partly complementary to the DIS and DY data.

- **The computed cross-sections are compared with the experimental data, and the initial parametrization is varied until the best agreement with the data is found.** Our definition for the “best agreement” is based on minimizing the following χ^2 -function

$$\chi^2(\{a\}) \equiv \sum_N w_N \chi_N^2(\{a\}) \quad (2.2)$$

$$\chi_N^2(\{a\}) \equiv \left(\frac{1 - f_N}{\sigma_N^{\text{norm}}} \right)^2 + \sum_{i \in N} \left[\frac{f_N D_i - T_i(\{a\})}{\sigma_i} \right]^2. \quad (2.3)$$

Within each data set N , D_i denotes the experimental data value with σ_i point-to-point uncertainty, and T_i is the theory prediction corresponding to a parameter set $\{a\}$. If an overall normalization uncertainty σ_N^{norm} is specified by the experiment, the normalization factor $f_N \in [1 - \sigma_N^{\text{norm}}, 1 + \sigma_N^{\text{norm}}]$ is introduced. Its value is determined by minimizing χ_N^2 and the final f_N is thus an output of the analysis. The weight factors w_N are used to amplify the importance of those data sets whose content is physically relevant, but contribution to χ^2 would otherwise be too small to have an effect on the automated χ^2 minimization.

In addition to finding the parameter set $\{a^0\}$ that optimally fits the experimental data, quantifying the uncertainties stemming from the experimental errors has become an increasingly important topic in the context of global PDF analyses. The Hessian method [11] provides a practical way of treating this issue. It is based on a quadratic approximation for the χ^2 around its minimum χ_0^2 ,

$$\chi^2 \approx \chi_0^2 + \sum_{ij} \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_i \partial a_j} (a_i - a_i^0)(a_j - a_j^0) \equiv \chi_0^2 + \sum_{ij} H_{ij} (a_i - a_i^0)(a_j - a_j^0), \quad (2.4)$$

which defines the Hessian matrix H . Non-zero off-diagonal elements in the Hessian matrix are a sign of correlations between the original fit parameters, invalidating the standard (diagonal) error propagation formula

$$(\Delta X)^2 = \left(\frac{\partial X}{\partial a_1} \cdot \delta a_1 \right)^2 + \left(\frac{\partial X}{\partial a_2} \cdot \delta a_2 \right)^2 + \dots \quad (2.5)$$

for a PDF-dependent quantity (cross section) X . Therefore, it is useful to diagonalize the Hessian matrix, such that

$$\chi^2 \approx \chi_0^2 + \sum_i z_i^2, \quad (2.6)$$

where each z_i is a certain linear combination of the original parameters around $\{a^0\}$. In these variables, the usual form of the error propagation

$$(\Delta X)^2 = \left(\frac{\partial X}{\partial z_1} \cdot \delta z_1 \right)^2 + \left(\frac{\partial X}{\partial z_2} \cdot \delta z_2 \right)^2 + \dots \quad (2.7)$$

stands on a much more solid ground. How to determine the size of the deviations δz_k is, however, a difficult issue where no universally agreed procedure exists. We consider 15 parameters in the expansion (2.6) and the prescription which we adopt here is to choose each δz_k such that χ^2 grows by a certain fixed amount $\Delta \chi^2$. Requiring each data set to remain close to its 90%-confidence range, we end up with a choice $\Delta \chi^2 = 50$ (see Appendix A of Ref. [9] for details).

Much of the practicality of the Hessian method resides in constructing PDF error sets, which we denote by S_k^\pm . Each S_k^\pm is obtained by displacing the fit parameters to the positive/negative direction along z_k such that χ^2 grows by the chosen $\Delta\chi^2 = 50$. Approximating the derivatives in Eq. (2.7) by a finite difference, we may then re-write the error formula as

$$(\Delta X)^2 = \frac{1}{4} \sum_k [X(S_k^+) - X(S_k^-)]^2, \quad (2.8)$$

where $X(S_k^\pm)$ denotes the value of the quantity X computed by the set S_k^\pm . If the lower and upper uncertainties ΔX^\pm differ, they should be computed separately, using the prescription [12]

$$\begin{aligned} (\Delta X^+)^2 &\approx \sum_k [\max\{X(S_k^+) - X(S^0), X(S_k^-) - X(S^0), 0\}]^2 \\ (\Delta X^-)^2 &\approx \sum_k [\max\{X(S^0) - X(S_k^+), X(S^0) - X(S_k^-), 0\}]^2, \end{aligned} \quad (2.9)$$

where S^0 denotes the best fit.

Along with the grids and interpolation routine for the best NLO and LO fits, our new release — a computer routine called EPS09 [13] — contains also 30 error sets required for computation of the uncertainties to nuclear cross-section ratios similar to those we present in the following section. However, as can be understood from the definition (2.1), the total uncertainty in the absolute nPDFs is a combination of uncertainties from the baseline set CTEQ6.1M and those from the nuclear modifications R_i^A . Therefore, if an absolute cross-section is computed, also the CTEQ6.1M error-sets (with the EPS09 central set) should be included when calculating its uncertainty.

3. Results

In this section we present a selection of results from the NLO analysis. First, in Fig. 1 we plot the obtained modifications at two scales, at $Q_0^2 = 1.69 \text{ GeV}^2$ and at $Q^2 = 100 \text{ GeV}^2$, which is to demonstrate the scale-dependence of the modifications and their uncertainties. One prominent feature to be noticed is that even if there is a rather large uncertainty band for the small- x gluon modification R_G^A at Q_0^2 , the scale evolution tends to bring even a very strong gluon shadowing close to no shadowing at $Q^2 \sim 100 \text{ GeV}^2$ — a very clear prediction of the DGLAP approach.

The DIS data constitutes the bulk of the experimental data available for the global analysis of nPDFs. In Fig. 2 we show some of the measured nuclear modifications for $(l + A)$ DIS structure functions with respect to Deuterium,

$$R_{F_2}^A(x, Q^2) \equiv \frac{F_2^A(x, Q^2)}{F_2^d(x, Q^2)}, \quad (3.1)$$

and the comparison with our EPS09 NLO results. The shaded bands denote the uncertainty derived from the 30 EPS09NLO error sets and, as should be noticed, their width is comparable to the error bars in the experimental data. This *a posteriori* supports the validity of our procedure for determining the $\Delta\chi^2$. Similar conclusion can be drawn upon inspecting the nuclear effects in the Drell-Yan data

$$R_{\text{DY}}^A(x_{1,2}, M^2) \equiv \frac{\frac{1}{A} d\sigma_{\text{DY}}^{\text{pA}}/dM^2 dx_{1,2}}{\frac{1}{2} d\sigma_{\text{DY}}^{\text{pd}}/dM^2 dx_{1,2}}, \quad (3.2)$$

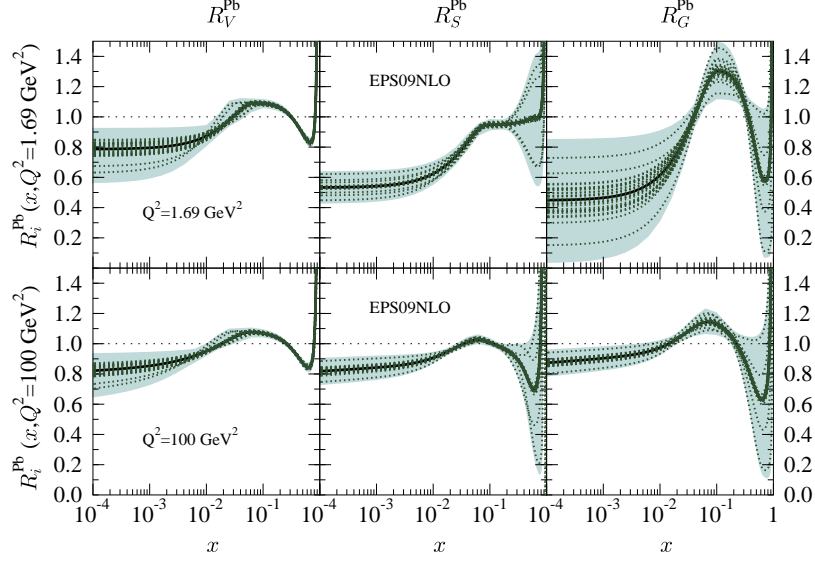


Figure 1: The nuclear modifications R_V , R_S , R_G for Lead at our initial scale $Q_0^2 = 1.69 \text{ GeV}^2$ and at $Q^2 = 100 \text{ GeV}^2$. The thick black lines indicate the best-fit results, whereas the dotted green curves denote the individual error sets. The shaded bands are computed using Eq. (2.9).

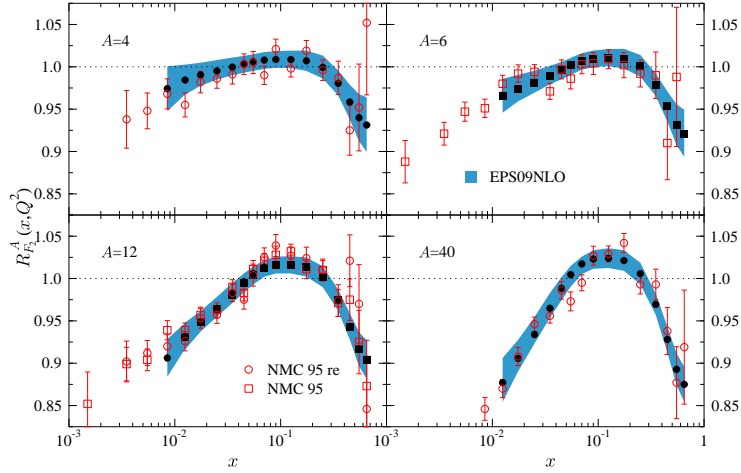


Figure 2: The calculated NLO $R_{F_2}^A(x, Q^2)$ (filled symbols and error bands) compared with the NMC 95 [14] and the reanalysed NMC 95 [15] data (open symbols).

where M^2 is the invariant mass of the lepton pair and $x_{1,2} \equiv \sqrt{M^2/s} e^{\pm y}$ (y is the pair rapidity). Comparison to the E772 and E866 data is shown in Fig. 3. Let us mention that the E772 data at $x_2 > 0.1$ carry some residual sensitivity also to the gluons and sea quarks, which has not been noticed before.

The nuclear modification for inclusive pion production in d+Au collisions relative to p+p is

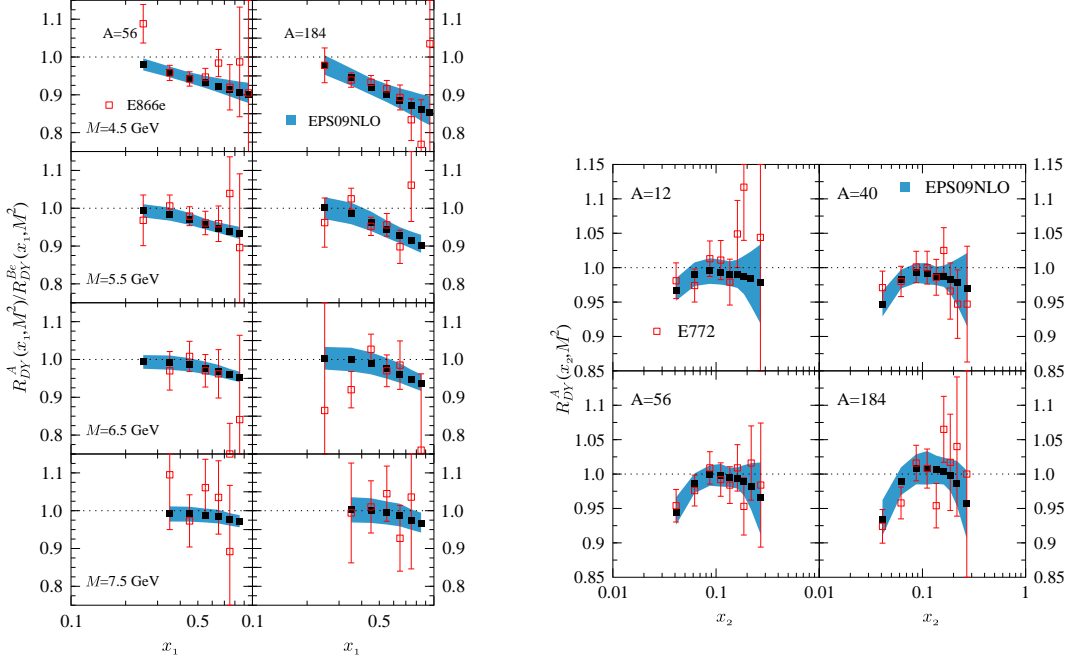


Figure 3: The computed NLO $R_{DY}^A(x, M^2)$ (filled squares and error bands) as a function of x_1 and x_2 compared with the E866 [17] and the E772 [16] data (open squares).

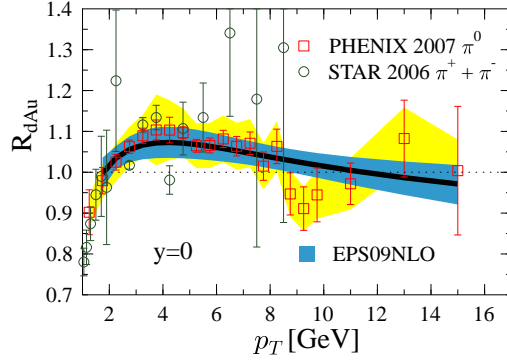


Figure 4: The computed R_{dAu} (thick black line and blue error band) at $y = 0$ for inclusive pion production compared with the PHENIX [18] data (open squares). The error bars are the statistical uncertainties, and the yellow band indicates the point-to-point systematic errors. The PHENIX data have been multiplied by the optimized normalization factor $f_N = 1.03$, which is an output of our analysis. The STAR data [19] (open circles), multiplied by $f_N = 0.90$, are also shown although they were not included in the EPS09 analysis.

defined as

$$R_{dAu}^{\pi} \equiv \frac{1}{\langle N_{\text{coll}} \rangle} \frac{d^2 N_{\pi}^{\text{dAu}} / dp_T dy}{d^2 N_{\pi}^{\text{pp}} / dp_T dy} \stackrel{\text{min.bias}}{=} \frac{\frac{1}{2A} d^2 \sigma_{\pi}^{\text{dAu}} / dp_T dy}{d^2 \sigma_{\pi}^{\text{pp}} / dp_T dy}, \quad (3.3)$$

where p_T , y are the transverse momentum and rapidity of the pion, and $\langle N_{\text{coll}} \rangle$ denotes the number of binary nucleon-nucleon collisions. A comparison with the PHENIX and STAR data is shown in

Fig. 4, and evidently, the shape of the spectrum gets very well reproduced by our parametrization. The downward trend toward large p_T provides direct evidence for an EMC-effect in the large- x gluons, while the suppression toward small p_T is in line with the gluon shadowing. No additional effects are needed to reproduce the observed spectra. It is also reassuring that this shape is practically independent of the fragmentation functions used in the calculation — modern sets like [20, 21, 22] all give equal results.

Special attention should be paid to the scale-breaking effects in the data and to their good description by the DGLAP evolution. These effects are clearly visible e.g. in the E886 Drell-Yan data in Fig. 3 where the trend of diminishing nuclear effects toward larger invariant mass M^2 is observed. Also, from the DIS data as a function of Q^2 , shown in Fig. 5, the general features can be filtered out: At small x the Q^2 -slopes tend to be positive, while toward larger x the slopes gradually die out and become even slightly negative.

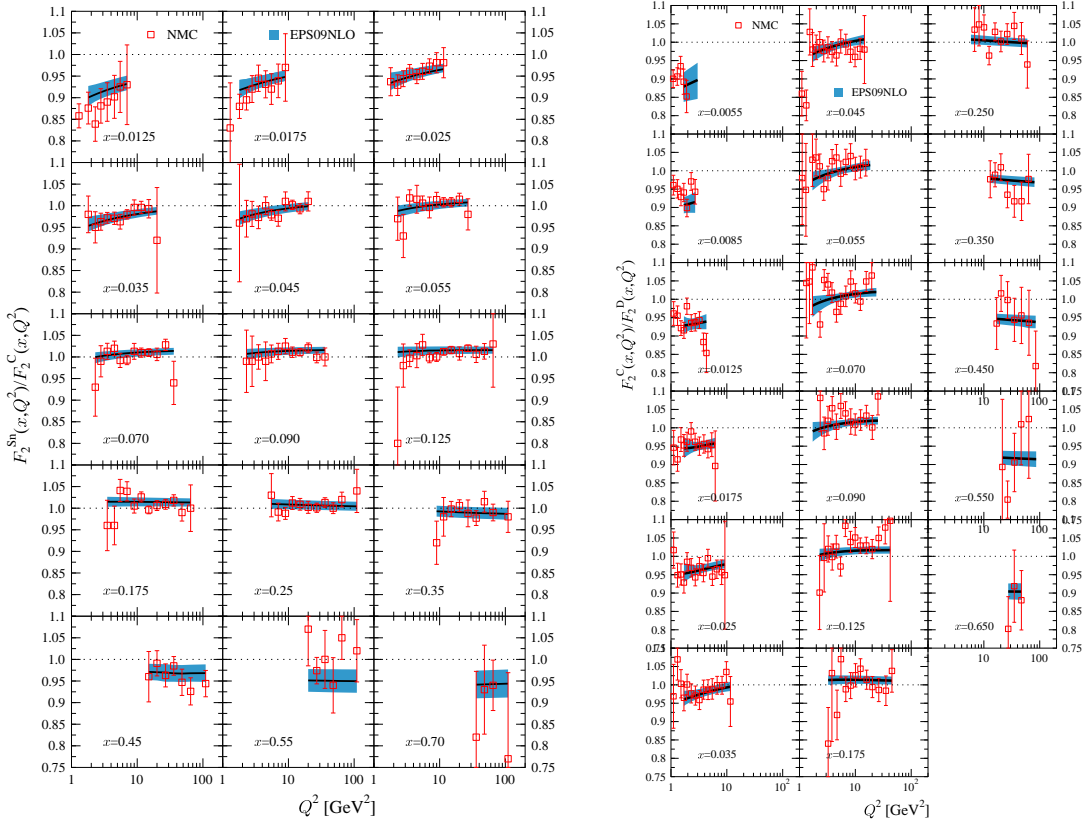


Figure 5: The calculated NLO scale evolution (solid black lines and error bands) of the ratio F_2^{Sn}/F_2^C and F_2^C/F_2^D , compared with the NMC data [23] for several fixed values of x .

To close this section we show, in Fig. 6, a comparison of the nuclear modifications for Lead from all available global NLO analyses. The comparison is shown again at $Q_0^2 = 1.69 \text{ GeV}^2$ and at $Q^2 = 100 \text{ GeV}^2$. Most significant differences — that is, curves being outside our error bands — are found from the sea quarks and gluons. At low x , the differences shrink when the scale Q^2 is increased, but at the high- x region notable discrepancies persist at all scales. Most of the differences are presumably explainable by the assumed behaviours of the fit functions, but also the choices for

the considered data sets (e.g. HKN and nDS do not implement the pion data), and differences in the definition of χ^2 carry some importance.

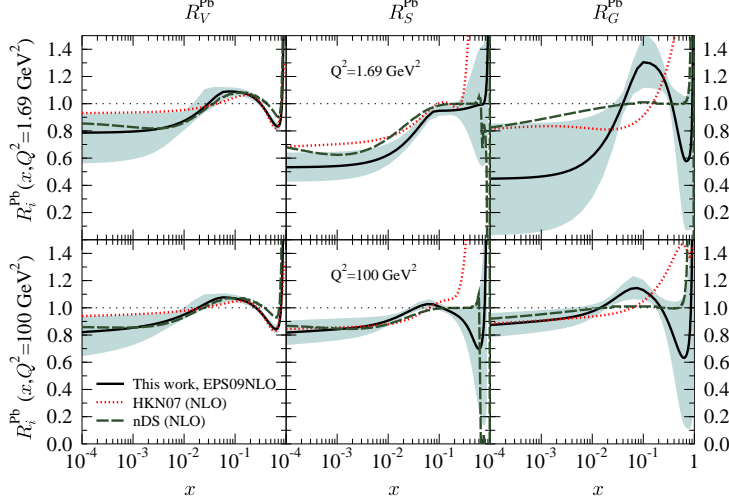


Figure 6: Comparison of the nuclear modifications for Lead at $Q^2 = 1.69 \text{ GeV}^2$ and at $Q^2 = 100 \text{ GeV}^2$ from the NLO global DGLAP analyses HKN07 [7], nDS [8] and this work, EPS09.

4. Hadrons in the forward direction in d+Au at RHIC?

In our previous article [6], we studied the case of a very strong gluon shadowing, motivated by the suppression in the nuclear modification R_{dAu} for the negatively-charged hadron yield in the forward rapidities ($\eta = 2.2, 3.2$) measured by the BRAHMS collaboration [24] in d+Au collisions at RHIC. We now return to this issue by applying the EPS09NLO parametrization to this specific process. The EPS09 predictions (with fDSS fragmentation functions [25]) compared with the BRAHMS data for the absolute h^- spectra,

$$\frac{d^3 N^{\text{pp}}}{d^2 p_T dy} \stackrel{\text{min.bias}}{=} \frac{1}{\sigma_{NN}^{\text{inelastic}}} \frac{d^3 \sigma^{\text{pp}}}{d^2 p_T dy} ; \quad \frac{d^3 N^{\text{dAu}}}{d^2 p_T dy} \stackrel{\text{min.bias}}{=} \frac{\langle N_{\text{coll}} \rangle}{\sigma_{NN}^{\text{inelastic}}} \frac{1}{2A} \frac{d^3 \sigma^{\text{dAu}}}{d^2 p_T dy}, \quad (4.1)$$

are shown in Fig. 7. In the $\eta = 2.2$ bin, the measured p+p and d+Au spectra are both in good agreement with the NLO pQCD. However, in the most forward $\eta = 3.2$ bin in the p+p case, there is a systematic and significant discrepancy between the measured and computed p_T spectra. This observation casts doubts on any conclusion made from the nuclear modification R_{dAu} alone — clearly, one should first account for the absolute baseline spectrum in p+p collisions. This is why we have not included this BRAHMS data set in our global analysis, either.

5. Summary

We have here outlined our NLO analysis of nPDFs. The very good agreement with the experimental data $\chi^2/N \approx 0.79$ — especially the correct description of the scaling-violation effects —

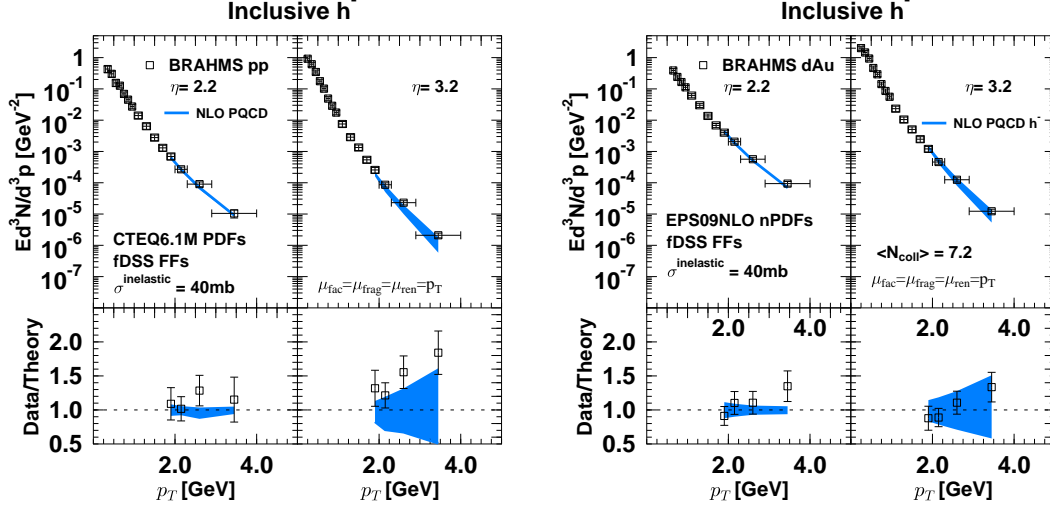


Figure 7: Inclusive h^- yield in p+p and d+Au collisions. The experimental p_T -binned data [24] are shown by open squares with statistical and systematical errors added in quadrature. The blue band indicates the 90% confidence range derived from the CTEQ6.1M and EPS09 uncertainties. The calculated cross-sections have been averaged over the p_T -bin width.

lends support to the validity of collinear factorization in nuclear collisions. In addition to the best fit, we have distilled the experimental errors into the 30 nPDF error-sets, which encode the relevant parameter-space neighbourhood of the χ^2 minimum. All these sets will be available as a computer routine at [13] for general use.

Although not discussed here, we have also performed the leading-order (LO) counterpart of the NLO analysis as we want to provide the uncertainty tools also for the widely-used LO framework. The best-fit quality is very similar both in LO and NLO, but the uncertainty bands become somewhat smaller when going to higher order. In the LO case, our error analysis indeed accommodates also the strong gluon shadowing suggested in our previous work EPS08 [6].

In the near future, more RHIC data will become available and published, and factorization will be tested further. Also p+A runs at the LHC would be welcome for this purpose. Even better possibilities for the nPDF studies would, however, be provided by the lepton-ion colliders like the planned eRHIC or LHeC.

During the recent years, the free-proton PDF analyses have gradually shifted to a better organized prescription for treating the heavy quarks than the zero-mass scheme employed in this analysis and in CTEQ6.1M. Such general-mass scheme should be especially important e.g. in the case of charged-current neutrino interactions. Interestingly, the CTEQ collaboration has lately looked also at this type of data among other neutrino-Iron measurements [26], and noticed that the best fit tends to point to somewhat different nuclear modifications than what have been obtained in the global nPDF analyses. Both the extension of the nPDF analysis to the general-mass scheme and a systematic investigation of the possible discrepancies between neutrino and other data, remains as future work.

References

- [1] Y. L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 641 [*Zh. Eksp. Teor. Fiz.* **73** (1977) 1216]; V. N. Gribov and L. N. Lipatov, *Yad. Fiz.* **15** (1972) 781 [*Sov. J. Nucl. Phys.* **15** (1972) 438]; V. N. Gribov and L. N. Lipatov, *Yad. Fiz.* **15** (1972) 1218 [*Sov. J. Nucl. Phys.* **15** (1972) 675]; G. Altarelli and G. Parisi, *Nucl. Phys. B* **126** (1977) 298.
- [2] N. Armesto, *J. Phys. G* **32** (2006) R367 [arXiv:hep-ph/0604108].
- [3] K. J. Eskola, V. J. Kolhinen and P. V. Ruuskanen, *Nucl. Phys. B* **535** (1998) 351 [arXiv:hep-ph/9802350].
- [4] K. J. Eskola, V. J. Kolhinen and C. A. Salgado, *Eur. Phys. J. C* **9** (1999) 61 [arXiv:hep-ph/9807297].
- [5] K. J. Eskola, V. J. Kolhinen, H. Paukkunen and C. A. Salgado, *JHEP* **0705** (2007) 002 [arXiv:hep-ph/0703104].
- [6] K. J. Eskola, H. Paukkunen and C. A. Salgado, *JHEP* **0807** (2008) 102 [arXiv:0802.0139 [hep-ph]].
- [7] M. Hirai, S. Kumano and T. H. Nagai, arXiv:0709.3038 [hep-ph].
- [8] D. de Florian and R. Sassot, *Phys. Rev. D* **69** (2004) 074028 [arXiv:hep-ph/0311227].
- [9] K. J. Eskola, H. Paukkunen and C. A. Salgado, arXiv:0902.4154 [hep-ph], submitted to JHEP.
- [10] D. Stump, J. Huston, J. Pumplin, W. K. Tung, H. L. Lai, S. Kuhlmann and J. F. Owens, *JHEP* **0310** (2003) 046 [arXiv:hep-ph/0303013].
- [11] J. Pumplin *et al.*, *Phys. Rev. D* **65** (2001) 014013 [arXiv:hep-ph/0101032].
- [12] P. M. Nadolsky and Z. Sullivan, in *Proc. of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001)* ed. N. Graf, *In the Proceedings of APS / DPF / DPB Summer Study on the Future of Particle Physics (Snowmass 2001), Snowmass, Colorado, 30 Jun - 21 Jul 2001, pp P510* [arXiv:hep-ph/0110378].
- [13] <https://www.jyu.fi/fysiikka/en/research/highenergy/urhic/nPDFs>
- [14] M. Arneodo *et al.* [New Muon Collaboration.], *Nucl. Phys. B* **441** (1995) 12 [arXiv:hep-ex/9504002].
- [15] P. Amaudruz *et al.* [New Muon Collaboration.], *Nucl. Phys. B* **441** (1995) 3 [arXiv:hep-ph/9503291].
- [16] D. M. Alde *et al.*, *Phys. Rev. Lett.* **64** (1990) 2479.
- [17] M. A. Vasilev *et al.* [FNAL E866 Collab.], *Phys. Rev. Lett.* **83** (1999) 2304 [arXiv:hep-ex/9906010].
- [18] S. S. Adler *et al.* [PHENIX Collab.], *Phys. Rev. Lett.* **98** (2007) 172302 [arXiv:nucl-ex/0610036].
- [19] J. Adams *et al.* [STAR Collaboration], *Phys. Lett. B* **637** (2006) 161 [arXiv:nucl-ex/0601033].
- [20] B. A. Kniehl, G. Kramer and B. Potter, *Nucl. Phys. B* **582** (2000) 514 [arXiv:hep-ph/0010289].
- [21] S. Albino, B. A. Kniehl and G. Kramer, arXiv:0803.2768 [hep-ph].
- [22] D. de Florian, R. Sassot and M. Stratmann, *Phys. Rev. D* **75** (2007) 114010 [arXiv:hep-ph/0703242].
- [23] M. Arneodo *et al.* [New Muon Collaboration.], *Nucl. Phys. B* **481** (1996) 23.
- [24] I. Arsene *et al.* [BRAHMS Collab.], *Phys. Rev. Lett.* **93** (2004) 242303 [arXiv:nucl-ex/0403005].
- [25] D. de Florian, R. Sassot and M. Stratmann, *Phys. Rev. D* **76** (2007) 074033 [arXiv:0707.1506 [hep-ph]].
- [26] I. Schienbein, J. Y. Yu, C. Keppel, J. G. Morfin, F. Olness and J. F. Owens, arXiv:0710.4897 [hep-ph].