NUCLEAR PARTON DISTRIBUTIONS FROM NEURAL NETWORKS*

Rabah Abdul Khalek, Jacob J. Ethier, Juan Rojo

the NNPDF Collaboration

Department of Physics and Astronomy, Vrije Universiteit Amsterdam
1081 HV Amsterdam, The Netherlands
and
Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands

(Received March 6, 2019)

In this contribution, we present a status report on the recent progress towards an analysis of nuclear parton distribution functions (nPDFs) using the NNPDF methodology. We discuss how the NNPDF fitting approach can be extended to account for the dependence on the atomic mass number A, and introduce novel algorithms to improve the training of the neural network parameters within the NNPDF framework. Finally, we present preliminary results of an nPDF fit to neutral current deep-inelastic lepton–nucleus scattering data, and demonstrate how one can validate the new fitting methodology by means of closure tests.

DOI:10.5506/APhysPolBSupp.12.927

1. Introduction

Parton distribution functions (PDFs) are universal, process-independent objects describing the longitudinal motion of quarks and gluons within hadrons [1, 2]. Since PDFs are difficult to compute from first principles, they are instead extracted from experimental data by means of a global analysis in the framework of QCD collinear factorization theorems. Currently, the PDFs of nucleons bound within heavy nuclei (nPDFs) [3] are less well-understood than their free-nucleon counterparts, due primarily to the limited experimental constraints available.

This state of affairs is unfortunate, since the determination of nPDFs is important to reveal the origin and properties of phenomena such as the Fermi motion, the EMC effect, nuclear shadowing, and possible non-linear

* Presented by R.A. Khalek at the Diffraction and Low-x 2018 Workshop, August 26–September 1, 2018, Reggio Calabria, Italy.

evolution effects in nuclei. In addition, nPDFs are key inputs for the interpretation of heavy-ion collisions and the characterization of the Quark–Gluon Plasma (QGP), as well as for high-energy astrophysics such as theoretical predictions of neutrino–nucleus interaction cross sections [4].

Several groups have presented nPDF determinations in recent years. Two of such analyses are EPPS16 [5], which fits a nuclear modification factor with respect to the CT14 [6] proton baseline, and nCTEQ15 [7], which fits directly the nPDF shape by mimicking the parameterization used in the CTEQ proton fits [8]. This recent activity in global nPDF studies has been largely prompted by the availability of proton–lead collision observables such as dijet, D meson, or W and Z gauge boson production. As in the case of proton PDFs, these measurements offer the potential of a greatly improved understanding of nPDFs and their uncertainties.

2. Towards nNNPDF1.0

Following the NNPDF methodology [9–11] (see [12] for a summary), we adopt here artificial neural networks (ANNs) as universal unbiased interpolants to parameterize the x and A dependence of the nPDFs. As an initial study, we consider only observables from neutral current (NC) deep-inelastic scattering (DIS) off heavy nuclei, which are assumed to be isoscalar.

The description of isoscalar nuclei observables in DIS below the Z-boson pole requires the parametrisation of three independent nPDFs, which are taken to be the quark singlet Σ, the gluon g, and the quark non-singlet octet T_8 distributions. In this basis, for example, the NC DIS structure function F_2^A is given by

$$\begin{aligned} &F_2^A\left(x,Q^2\right) = \Gamma^{\rm S}_{2,\Sigma}\left(x,Q_0^2,Q^2\right)\otimes\Sigma\left(x,A,Q_0^2\right) \\ &+\Gamma^{\rm S}_{2,g}\left(x,Q_0^2,Q^2\right)\otimes g\left(x,A,Q_0^2\right)+\Gamma^{\rm NS}_{2,T_8}\left(x,Q_0^2,Q^2\right)\otimes T_8\left(x,A,Q_0^2\right)\,, \end{aligned} \tag{1}$$

where the Γ factors encode both the hard-scattering coefficient functions and the DGLAP evolution kernels. The nPDFs are then parametrised at an initial scale denoted by Q_0, and depend both on the partonic momentum fraction x and the mass number A.

Following Ref. [13], the convolutions in Eq. (1) can be reduced to a scalar product by means of an expansion over a set of interpolating polynomials, allowing us to write

$$F_2^A\left(x,Q^2\right) = \sum_i^{n_f}\sum_\alpha^{n_x}\widetilde{\Gamma}_{i,\alpha}\left(x,x_\alpha,Q^2,Q_0^2\right)\,q_i\left(x_\alpha,A,Q_0^2\right)\,, \tag{2}$$

where $\widetilde{\Gamma}$ stand for the precomputed FastKernel grids that contain all the perturbative information relevant for the calculation of F_2^A, and $q_i(x,A,Q_0^2)$

represent the initial scale nPDF for the flavour i in a given basis. Note that in Eq. (2) only the values of the input PDFs at a finite n_x-sized grid are required to compute the structure functions, leading to a significant improvement of the numerical computation with respect to the convolutions in Eq. (1).

The three independent nPDFs that enter Eq. (2) are parametrised as:

— $\Sigma(x, A, Q_0) = (1-x)^{\alpha_\Sigma} x^{-\beta_\Sigma} \mathrm{NN}_\Sigma(x, A)$,
— $g(x, A, Q_0) = A_g (1-x)^{\alpha_G} x^{-\beta_g} \mathrm{NN}_g(x, A)$,
— $T_8(x, A, Q_0) = (1-x)^{\alpha_{T_8}} x^{-\beta_{T_8}} \mathrm{NN}_{T_8}(x, A)$,

where NN_i corresponds to the output of the ANN for a given flavor i. The preprocessing exponents [14] α_i and β_i facilitate the training procedure and can either be fitted or drawn at random from a range determined iteratively. Furthermore, we fix the overall normalisation of the gluon nPDF

$$A_g \equiv \left(1 - \int_0^1 \Sigma(x, A, Q_0)\mathrm{d}x\right) \Bigg/ \int_0^1 g(x, A, Q_0)\mathrm{d}x \tag{3}$$

so that the momentum sum rule is satisfied. In general, this normalisation is different for every value of A.

Concerning the input dataset, we consider here a similar set of nuclear NC DIS measurements that were used by EPPS16 and nCTEQ15. In Fig. 1, the kinematic coverage of the (x, Q^2) plane of the nuclear DIS data are

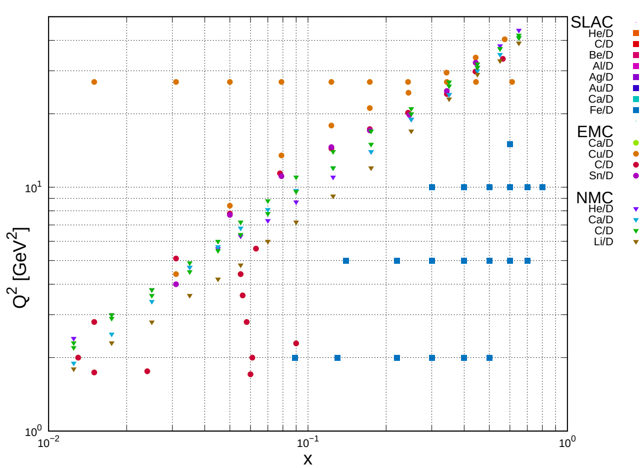

Fig. 1. The kinematic coverage in x and Q^2 of the NC DIS data included in this work.

shown. Here, the coverage in x is significantly reduced compared to the proton case ($x > 10^{-2}$ *versus* $x > 10^{-5}$ respectively). Enlarging this kinematic range to smaller values of x and higher values of Q^2 is possible by means of the RHIC and LHC data on nucleon–nucleus collisions.

3. Neural network training

The underlying procedure for any optimisation problem, such as the present one, can be summarized by

$$\min_{\boldsymbol{\omega}} C(f(\boldsymbol{\omega}))\,, \tag{4}$$

where C is a cost function to minimize and f is the target function that depends on a vector of parameters $\boldsymbol{\omega}$. In our case, the target functions that need to be determined from the data are the nPDFs parametrised by the neural networks, while $\boldsymbol{\omega}$ represents the neural network weights and thresholds. The cost function is defined here to be the χ^2, which measures the agreement between the experimental data points D_i and the corresponding theoretical predictions T_i of nuclear DIS observables

$$\chi^2 = \sum_{i=1}^{n_{\rm dat}} \frac{(T_i[f(\boldsymbol{\omega})] - D_i)^2}{\sigma_i^2}\,, \tag{5}$$

where σ_i is the total experimental statistical and systematic uncertainties added in quadrature.

There are different options that can be used to solve Eq. (4). Previous NNPDF global fits have been based either in Genetic Algorithms (GAs) or the Covariance Matrix Adaptation — Evolutionary Strategy (CMA-ES) algorithms. Both methods require knowledge only on the local values of the χ^2 and not of its derivatives. Here, for the first time in the context of NNPDF studies, we have implemented the method of gradient descent, one of the most widely used minimization techniques in machine learning applications. In this procedure, the parameters are shifted by an amount proportional to the negative of the gradient of the cost function evaluated at the current position in parameter space

$$\omega_i \to \omega_i - \frac{\eta}{n_{\rm par}} \frac{\partial \chi^2}{\partial \omega_i}\,, \tag{6}$$

where ω_i is one of the $n_{\rm par}$ free parameters of the ANN and η is a hyperparameter of the algorithm known as the learning rate. In this work, the gradients are computed numerically by means of automatic differentiation using the TensorFlow library [15]. The updating process in Eq. (6) is then iterated until a suitable set of convergence criteria is satisfied, for instance, using look-back or early stopping with cross-validation.

4. Preliminary results

The application of the NNPDF methodology to a QCD analysis of nuclear parton distributions can be validated by means of closure tests, as was done in previous global fits of proton PDFs [11] and fragmentation functions [16]. In these closure tests, pseudo-data is generated based on an established theoretical input. In this work, we construct pseudo-data with Eq. (1) up to NLO in perturbative QCD using the EPPS16 nPDF set. A fit is then performed to this pseudo-data, and by comparing the fit output to the known input, we can assess if a given fitting methodology is working successfully. Since this pseudo-data is free from possible data inconsistencies or limitations in the theory calculations, it provides a clean testing group to validate the fitting strategy employed.

We have performed the simplest type of closure test, known as Level 0 (L0), where the pseudo-data coincides with the EPPS16 prediction without any additional set of statistical noise added. In this case, a successful fit should be able to reach asymptotically $\chi^2 \to 0$, so that the fit predictions at the structure function level are identical to those obtained with the input EPPS16 theory. For this L0 closure test, we have used a single neural network with three input nodes $(x, \ln x, A)$, three output nodes NN_Σ, NN_g, NN_{T_8}, and a single hidden layer with 20 neurons. This ANN architecture contains then 143 free parameters. Our final result consists of 200 Monte Carlo "replicas", which in L0 tests correspond to different initial parameter values that are randomly chosen for each fit.

In Fig. 2, we display the x dependence of the preliminary nNNPDF1.0 L0 closure test fit results for iron nuclei ($A = 56$). Here, the gluon, quark singlet Σ, and non-singlet octet T_8 distributions are shown at the input scale $Q_0 = 1.3$ GeV and are compared with the EPPS16 result. While the EPPS16 uncertainty band is computed using the asymmetric Hessian method, the nNNPDF1.0 band instead represents the variance evaluated over the 200 fitted replicas. The fact that both central values agree reasonably well, especially for the gluon, is a first indication that the closure test is successful. Note that Σ and T_8 are strongly anti-correlated in the data region, since the actual quantity which is being constrained from data is $F_2^A \propto \Sigma + T_8/4$. Also, for L0 closure tests, the two error bands have different statistical interpretations and cannot be compared directly.

In Fig. 3, we present a similar comparison as in Fig. 2, but now show the mass number A dependence for a fixed value of x. Here, we see that the closure test results agree with the input EPPS16 theory within uncertainties, and that they reproduce the same qualitative behaviour as A is varied.

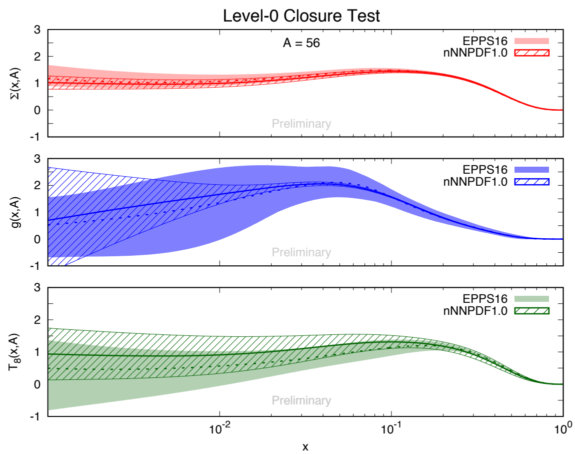

Fig. 2. (Colour on-line) The results of the nNNPDF1.0 Level 0 closure test. The nNNPDF1.0 result (solid line with hashed bands) is compared to the EPPS16 nPDF result (hashed line with shaded bands) for the iron nucleus ($A = 56$) at $Q_0 = 1.3$ GeV for the quark singlet Σ (red/top panel), the gluon g (blue/medium panel), and the quark non-singlet octet distribution T_8 (green/bottom panel).

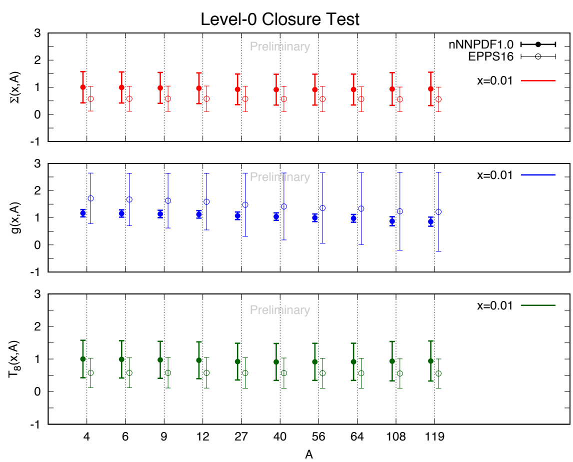

Fig. 3. Similar to Fig. 2, but now showing the comparison for discrete values of mass number A at fixed $x = 0.01$.

5. Next steps

In this contribution, we have presented the initial steps towards the first determination of the nuclear parton distributions in the framework of the NNPDF methodology. We have validated the effectiveness of improved neural network training algorithms, in particular the gradient descent minimization with TensorFlow. At the closure test level, we have shown that we are able to reproduce the results of the chosen input theory, in this case EPPS16 nPDFs. Work in progress is now focused on extending our approach to Level 1 and 2 closure tests, as well as to a full QCD analysis of experimental data. Following an initial study on DIS measurements, we aim to deliver a full-fledged global nPDF fit that accounts for all available experimental constraints and is based on state-of-the-art theoretical calculations.

This research has been supported by a European Research Council Starting grant "PDF4BSM", and by the Netherlands Organization for Scientific Research (NWO).

REFERENCES

[1] J. Butterworth *et al.*, *J. Phys. G* **43**, 023001 (2016) [`arXiv:1510.03865 [hep-ph]`].

[2] J. Gao, L. Harland-Lang, J. Rojo, *Phys. Rep.* **742**, 1 (2018) [`arXiv:1709.04922 [hep-ph]`].

[3] H. Paukkunen, `arXiv:1811.01976 [hep-ph]`.

[4] V. Bertone, R. Gauld, J. Rojo, *J. High Energy Phys.* **1901**, 217 (2019) [`arXiv:1808.02034 [hep-ph]`].

[5] K.J. Eskola, P. Paakkinen, H. Paukkunen, C.A. Salgado, *Eur. Phys. J. C* **77**, 163 (2017) [`arXiv:1612.05741 [hep-ph]`].

[6] S. Dulat *et al.*, *Phys. Rev. D* **93**, 033006 (2016) [`arXiv:1506.07443 [hep-ph]`].

[7] K. Kovarik *et al.*, *Phys. Rev. D* **93**, 085037 (2016) [`arXiv:1509.00792 [hep-ph]`].

[8] J. Pumplin *et al.*, *J. High Energy Phys.* **0207**, 012 (2002) [`arXiv:hep-ph/0201195`hep-ph/].

[9] R.D. Ball *et al.* [NNPDF Collaboration], *Nucl. Phys. B* **809**, 1 (2009) [*Erratum ibid.* **816**, 293 (2009)] [`arXiv:0808.1231 [hep-ph]`].

[10] R.D. Ball *et al.*, *Nucl. Phys. B* **867**, 244 (2013) [`arXiv:1207.1303 [hep-ph]`].

[11] R.D. Ball *et al.* [NNPDF Collaboration], *J. High Energy Phys.* **1504**, 040 (2015) [`arXiv:1410.8849 [hep-ph]`].

[12] J. Rojo, `arXiv:1809.04392 [hep-ph]`.

[13] V. Bertone, S. Carrazza, N.P. Hartland, *Comput. Phys. Commun.* **212**, 205 (2017) [`arXiv:1605.02070 [hep-ph]`].

[14] R.D. Ball, E.R. Nocera, J. Rojo, *Eur. Phys. J. C* **76**, 383 (2016) [`arXiv:1604.00024 [hep-ph]`].

[15] M. Abadi *et al.*, 2015, software available from `tensorflow.org`

[16] V. Bertone [NNPDF Collaboration], *Eur. Phys. J. C* **77**, 516 (2017) [`arXiv:1706.07049 [hep-ph]`].