

Progress on Neural Parton Distributions

J. Rojo¹, R. D. Ball², L. Del Debbio², S. Forte³, A. Guffanti²,
J. I. Latorre⁴, A. Piccione⁵ and M. Ubiali² (The NNPDF Collaboration)

- 1.- LPTHE, CNRS UMR 7589, Universités Paris VI-Paris VII, F-75252, Paris Cedex 05, France
- 2.- School of Physics, University of Edinburgh, Edinburgh EH9 3JZ, Scotland
- 3.- Dipartimento di Fisica, Università di Milano,
and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy
- 4.- Departament d'Estructura i Constituents de la Matèria,
Universitat de Barcelona, Diagonal 647, E-08028 Barcelona, Spain
- 5.- INFN, Sezione di Genova, via Dodecaneso 33, I-16146 Genova, Italy

We give a status report on the determination of a set of parton distributions based on neural networks. In particular, we summarize the determination of the nonsinglet quark distribution up to NNLO, we compare it with results obtained using other approaches, and we discuss its use for a determination of α_s .

1 Introduction

The LHC will require an approach to the search for new physics based on the precision techniques which are customary at lepton machines [2, 3]. This has recently led to significant progress in the determination of parton distribution functions (PDFs) of the nucleon. The main recent development has been the availability of sets of PDFs with an estimate of the associated uncertainty [4, 5, 6]. However, the standard approach to the determination of the uncertainty on parton distributions has several weaknesses, such as the lack of control on the bias due choice of a parametrization and, more in general, the difficulty in giving a consistent statistical interpretation to the quoted uncertainties.

These problems have stimulated various new approaches to the determination of PDFs [8], in particular the neural network approach, first proposed in Ref. [7]. The basic idea is to combine a Monte Carlo sampling of the probability measure on the space of functions that one is trying to determine [8] with the use of neural networks as universal unbiased interpolating functions. In Refs. [7, 9] this strategy was successfully applied to a somewhat simpler problem, namely, the construction of a parametrization of existing data on the DIS structure function $F_2(x, Q^2)$ of the proton and neutron. The method was proven to be fast and robust, to be amenable to detailed statistical studies, and to be in many respects superior to conventional parametrizations of structure functions based on a fixed functional form.

The determination of a parton set involves the significant complication of having to go from one or more physical observables to a set of parton distributions. Recently [10] most of the technical complications required for the construction of a neural parton set have been tackled and solved in the process of constructing a determination of the quark isotriplet parton distribution. This work will be reviewed here. Also, based on this work, we will present preliminary results on the determination of α_s and a determination of the variation in χ^2 which corresponds to a one-sigma variation of the underlying parton distributions. Work to apply the techniques of [10] to the singlet sector is at an advanced stage [11].

2 Determination of the nonsinglet quark distribution

The first application of the neural network approach to parton distributions, a determination of the NS parton distribution $q_{\text{NS}}(x, Q_0^2) = (u + \bar{u} - d + \bar{d})(x, Q_0^2)$ from the DIS structure function data of the NMC and BCDMS collaborations, was presented in Ref. [10]. Results for this PDF were obtained at LO, NLO, NNLO for different values of $\alpha_s(M_Z^2)$.

In Ref. [10] we have implemented a new fast and efficient method for solving the evolution equations up to NNLO. This method combines the advantages of x -space and N -space evolution codes: an x dependent Green function (evolution factor) is determined by inverse Mellin transformation of the exact N -space expression and stored. Evolution of PDFs is then performed by convoluting this Green function with any given boundary condition. The accuracy of this method has been benchmarked up to NNLO with the help of the tables of Refs. [2, 3].

Also, we have implemented a criterion to determine the convergence of the fitting procedure in a way which is free of bias related to the choice of parametrization. To this purpose, the dataset is randomly divided into two sets, of which only one is used in the fit. Convergence is achieved when the quality of the fit to data which are *not* used for minimization stops improving.

An important feature of our approach is that it is possible to check quantitatively the statistical features of results using suitable estimators. For example, one can check that the results do not depend on choices made during the fitting procedure, such as the choice of architecture of neural networks, which is analogous to the choice of parton parametrization in conventional fits. Namely, we repeat the fit with a different choice, and compute the distance

$$d[q] = \sqrt{\left\langle \frac{\left(q_i^{(1)} - q_i^{(2)}\right)^2}{\left(\sigma_i^{(1)}\right)^2 + \left(\sigma_i^{(2)}\right)^2} \right\rangle_{\text{dat}}}, \quad (1)$$

where $q_i^{(1)}$, $q_i^{(2)}$ are the predictions for the i -th data point in the two fits, and $\sigma_i^{(1)}$, $\sigma_i^{(2)}$ the predictions for the corresponding statistical uncertainties, and the average is performed over all data. The results of the first and second fit are the same if $d[q] = 1$ on average. This also checks that the statistical uncertainties are correctly estimated. One can similarly check stability of the uncertainty estimate. In Ref. [10] this comparison has been performed successfully.

In Fig. 1 we compare our results for the NS structure function F_2^{NS} to other published determinations. These results are available through the webpage of the NNPDF Collaboration: <http://sophia.ecm.ub.es/nnpdf>. The large uncertainty that we find is a genuine feature of the determination of the nonsinglet quark distribution from the data included in our fit, and, especially at small x , it appears to reflect the current knowledge of the nonsinglet quark distribution. Indeed, for $x \leq 0.05$ the only data which constrain the q_{NS} combination in global fits are the data used in the determination of Ref. [10]. Hence, our results suggest that standard fits might be underestimating PDF uncertainties.

In recent work on PDF uncertainties [5, 6] it has been suggested that, mostly because of inconsistencies between data, the variation of the total χ^2 which corresponds to a one-sigma variation of the underlying PDFs is of order of $\Delta\chi^2 \sim 50$ for the global fits presented in those references instead of $\Delta\chi^2 = 1$ of a statistically consistent fit [4]. In our approach, this quantity can be computed. We get $\Delta\chi^2 \approx 1.7$ (preliminary). This implies that the NMC

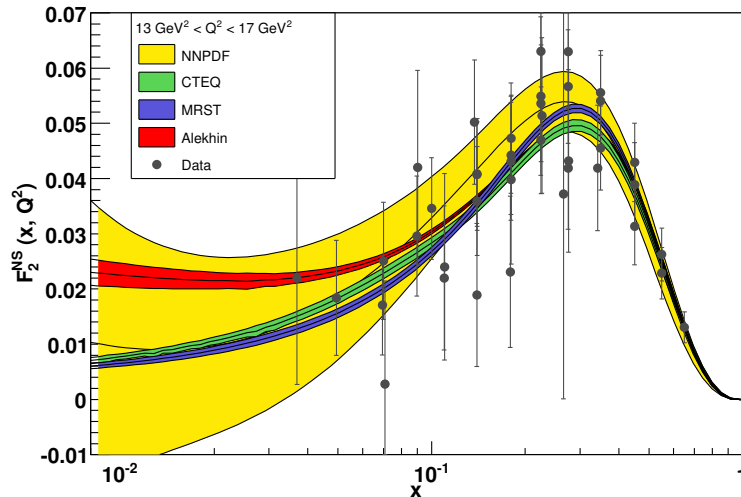


Figure 1:
The nonsinglet structure function F_2^{NS} as determined by the NNPDF collaboration [10] from 229 NMC and 254 BCDMS data points, compared to data and various other determinations.

and BCDMS data are mostly consistent, though some inconsistent data are present [7, 9]. An extensive discussion of the way the published [10] and forthcoming [11] fits based on the neural network approach can be used for the determination of physical parameters (such as α_s) and statistical properties of the data (such as $\Delta\chi^2$) will be presented in a forthcoming publication.

In [10] the strong coupling $\alpha_s(M_Z^2)$ was fixed, but we could also extract it from the fit. The results of a preliminary analysis, shown in Fig. 2, suggest that nonsinglet data determine $\alpha_s(M_Z^2)$ with an uncertainty which is rather larger than that ($\Delta\alpha_s(M_Z^2) \sim 0.002$) obtained in comparable determinations (see e.g. Ref. [12]). This preliminary result is consistent with that obtained using the same data in Ref. [13], with a method which eliminates the need to choose a parton parametrization. This strengthens the conclusion that uncertainties in available PDF fits might be underestimated.

3 Towards a full parton set

The extension of the results described in Ref. [10] to a full global PDF fit has benefited from the increased manpower of the NNPDF Collaboration, and is at a rather advanced stage [11]. In particular, the evolution formalism of Ref. [10] has been extended to the computation of a full set of neutral-current and charged-current structure functions and fully benchmarked. A first full neural parton fit is in preparation. It will at first be based on DIS data only, including all available F_2^p and F_2^d fixed target data and the full NC and CC HERA reduced cross sections.

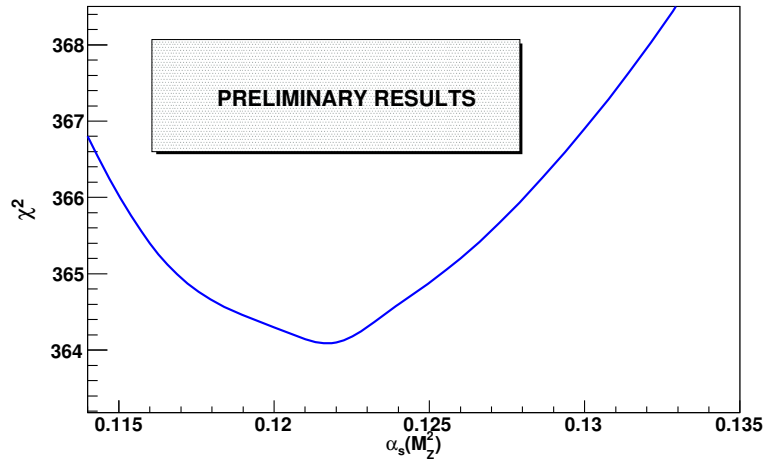


Figure 2:

The χ^2 profile for a preliminary NNLO determination of $\alpha_s(M_Z^2)$ from NS data. The number of data points included in the fit is $N_{\text{dat}} = 483$.

References

- [1] Slides:
<http://indico.cern.ch/contributionDisplay.py?contribId=32&sessionId=8&confId=9499>
- [2] W. Giele et al. The QCD/SM working group: Summary report. [hep-ph/0204316](#).
- [3] M. Dittmar et al. Parton distributions: Summary report. [hep-ph/0511119](#).
- [4] Sergey Alekhin. Parton distributions from deep-inelastic scattering data. *Phys. Rev.*, D68:014002, 2003.
- [5] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.
- [6] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne. Uncertainties of predictions from parton distributions. I: Experimental errors. *Eur. Phys. J.*, C28:455–473, 2003.
- [7] Stefano Forte, Lluis Garrido, Jose I. Latorre, and Andrea Piccione. Neural network parametrization of deep-inelastic structure functions. *JHEP*, 05:062, 2002.
- [8] W. T. Giele, S. A. Keller and D. A. Kosower. Parton distribution function uncertainties. [hep-ph/0104052](#).
- [9] Luigi Del Debbio, Stefano Forte, Jose I. Latorre, Andrea Piccione, and Joan Rojo. Unbiased determination of the proton structure function F_2^p with faithful uncertainty estimation. *JHEP*, 03:080, 2005.
- [10] Luigi Del Debbio, Stefano Forte, Jose I. Latorre, Andrea Piccione, and Juan Rojo. Neural network determination of parton distributions: The nonsinglet case. *JHEP*, 03:039, 2007.
- [11] Richard D. Ball, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Andrea Piccione, Juan Rojo, and Maria Ubiali. Neural network determination of parton distributions: The singlet case. *In preparation*, 2007.
- [12] J. Blumlein, H. Bottcher and A. Guffanti. Non-singlet QCD analysis of deep inelastic world data at $O(\alpha_s^3)$. *Nucl. Phys. B* **774** (2007) 182
- [13] Stefano Forte, Jose I. Latorre, Lorenzo Magnea, and Andrea Piccione. Determination of α_s from scaling violations of truncated moments of structure functions. *Nucl. Phys.*, B643:477–500, 2002.