DESY 10-164

arXiv:1009.6170v1 [hep-ph] 30 Sep 2010

Parton Distribution Uncertainties using Smoothness Prior

A. Glazov[a], S. Moch[b] and V. Radescu[c]

[a]*Deutsches Elektronensynchrotron DESY*
Notkestraße 85 D–22607 Hamburg, Germany

[b]*Deutsches Elektronensynchrotron DESY*
Platanenallee 6, D–15738 Zeuthen, Germany

[c] *Physikalisches Institut, Universität Heidelberg*
Philosophenweg 12, D–69120 Heidelberg, Germany

Abstract

A study of the parameterization uncertainty at low Bjorken $x \leq 0.1$ for the parton distribution functions of the proton is presented. The study is based on the HERA I combined data using a flexible parameterization form based on Chebyshev polynomials with and without an additional regularization constraint. The accuracy of the data allows to determine the gluon density in the kinematic range of $0.0005 \leq x \leq 0.05$ with a small parameterization uncertainty. An additional regularization prior leads to a significantly reduced uncertainty for $x \leq 0.0005$.

1 Introduction

The accurate knowledge of the parton distribution functions (PDFs) plays an important role for predictions of hard scattering cross sections at pp and $p\bar{p}$ colliders. The latter are computed in the perturbative approach including higher order radiative corrections, e.g. at next-to-leading order (NLO), which results in reduced theoretical uncertainties. Particular cross sections, such as Drell-Yan production of W,Z bosons at the LHC are even calculated to next-to-next-to-leading order (NNLO), see [1, 2], and exhibit a small theoretical uncertainty of $\sim 2\%$. For these processes, the accuracy of the prediction is presently limited by the uncertainties of the PDFs.

The PDFs being non-perturbative by definition can be determined from fits to data from DIS e- and ν-scattering, and from Drell-Yan experiments. These fits are performed using the well-known QCD evolution equations at NLO and NNLO [3–8]. The data are provided at discrete values of Bjorken x and absolute four momentum transfer squared Q^2 with their statistical and systematic uncertainties. With this given input, the uncertainty of the PDFs due to experimental errors are estimated using Hessian [9,10] and Monte Carlo [11] methods. Additional theoretical uncertainties arise, e.g. from unknown higher orders in the evolution or the treatment and scheme choice for heavy flavor contributions [12, 13]. These need to be considered separately (see e.g. [14, 15]).

PDF fits require an ansatz for the parameterization by a certain function of x at the starting scale Q_0^2 of the evolution. Fitting of experimental data at discrete points with an, in general, arbitrary function is an ill-posed problem which requires regularization. Typically Regge-theory inspired parameterizations are used with a small number of parameters which implicitly contain smooth and regular behavior requirements for the PDFs. For these parameterizations, it is difficult to estimate the PDF uncertainty arising from the choice of a particular ansatz. Alternatively, flexible parameterizations based on a neural net approach were used recently [16]. The number of parameters in this approach is determined by the data using an over-fitting protection technique which is an implicit regularization.

In this note, a new study of the parameterization uncertainty at low Bjorken $x < 0.1$ is performed. An explicit regularization prior is introduced which disfavors resonant-like behavior of PDFs at low x and the impact of the prior on the parameterization uncertainty is evaluated with particular emphasis on the gluon density in the range $0.0005 \leq x \leq 0.05$. We choose a flexible ansatz for the PDFs at low x using Chebyshev polynomials. The analysis is based on the combined HERA I data [17].

2 QCD Analysis Settings

The QCD analysis presented here is performed using as a sole input the combined H1 and ZEUS data on neutral and charged current $e^{\pm}p$ scattering double-differential cross sections collected during the HERA I run period of 1994-2000 [17]. The kinematic range of the data extends from $0.045 < Q^2 < 30000$ GeV^2 and $0.000006 < x < 0.65$, however in the QCD fit analysis only data with $Q^2 \geq Q^2_{\rm min} = 3.5$ GeV^2 are considered in order to minimize the non-perturbative higher twist effects.

The QCD fit is performed within the framework of the QCDNUM program implemented at NLO in QCD [18] and using a Zero-Mass-Variable-Flavor scheme. The fit minimizes a χ^2 function as specified in [17]. The PDFs are parameterized at the starting scale of $Q_0^2 = 1.9$ GeV^2. We use a

flavor decomposition similar to [19] as follows: xd_{val}, xu_{val}, $x\Delta = x\bar{u} - x\bar{d}$, $xS = 2(x\bar{u} + x\bar{d} + x\bar{s} + x\bar{c} + x\bar{b})$ where c and b quark densities are zero at the scales below their corresponding thresholds. The PDFs are evolved in Q^2 using the NLO equations in the massless $\overline{MS}$-scheme and the charm and beauty quark PDFs are generated by evolution for scales above the respective thresholds. The renormalization and factorization scales are set to Q^2.

Since this study is focused on the low $x < 0.1$ region, the set-up for the QCD analysis is special if somewhat simplified compared to modern high precision determinations of PDFs, see e.g. [17, 20–22]. At low x the PDFs are dominated by the gluon and sea-quark densities while at high x the valence-quark densities give larger contribution. Thus, regarding the functional form for the PDFs, standard Regge-theory inspired parameterizations are used for the valence quarks:

$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}}, \quad (1)$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}. \quad (2)$$

The low-x behavior of the valence densities is assumed to be the same for u and d quarks by setting $B_{u_v} = B_{d_v}$. The normalizations A_{u_v} and A_{d_v} are determined by the fermion number sum rules. Therefore the valence sector is described by three parameters.

For the gluon and sea densities a flexible Chebyshev polynomials based parameterization is used. The polynomials use $\log x$ as an argument to emphasize the low x behavior. The parameterization is valid for $x > x_{\min} = 1.7 \times 10^{-5}$ which covers the x range of the HERA measurements for $Q^2 \geq Q^2_{\min}$. The PDFs are multiplied by $(1-x)$ to ensure that they vanish as $x \to 1$. The resulting parameterization form is

$$xg(x) = A_g (1-x) \sum_{i=0}^{N_g-1} A_{g_i} T_i \left(-\frac{2\log x - \log x_{\min}}{\log x_{\min}} \right), \quad (3)$$

$$xS(x) = (1-x) \sum_{i=0}^{N_S-1} A_{S_i} T_i \left(-\frac{2\log x - \log x_{\min}}{\log x_{\min}} \right), \quad (4)$$

where T_i denote Chebyshev polynomials of the first type and the sum over i runs up to $N_{g,S} = 15$ for the gluon and sea-quark densities. The Chebyshev polynomials are given by the well-known recurrence relation:

$$T_0(x) = 1, \quad (5)$$

$$T_1(x) = x, \quad (6)$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \quad (7)$$

The normalization A_g is determined by the momentum sum rule. The advantage of the parameterization given by Eqs. (3), (4) is that momentum sum rule can be evaluated analytically. Moreover, already for $N_{g,S} \geq 5$ the fit quality is similar to that of a standard Regge-inspired parameterization with a similar number of parameters.

The PDF uncertainties are estimated using the Monte Carlo technique [11]. The method consists in preparing replicas of data sets by allowing the central values of the cross sections to fluctuate within their systematic and statistical uncertainties taking into account all point-to-point correlations. The preparation of the data is repeated for $N > 100$ times and for each of these replicas a complete NLO QCD fit is performed to extract the PDF set. The PDF central values and uncertainties are estimated using the mean values and root-mean-squared (RMS) over the PDF sets obtained for each replica.

3 Choice of the Smoothness Constraint

Fitting an arbitrary function to a discrete number of measurements is an ill-posed problem which requires regularization. This regularization should have a physical motivation and be flexible enough to cover the space of solutions compatible with QCD. At low x, the sea-quark PDF closely corresponds to a measurement of the structure function F_2 in a DIS process. For DIS at low x, the invariant mass of the hadronic final state W is calculated from Q^2 and x as

$$W = \sqrt{Q^2 \frac{1-x}{x}}. \tag{8}$$

Experimentally, it is well-known that for low values of W and Q^2 the structure function F_2 displays resonances [23]. These resonances, however, disappear for high $W > 5$ GeV. The smooth behavior of F_2 for high W can be explained phenomenologically by high particle multiplicity of the hadronic final state. A prior which disfavors resonant structures in W, for W exceeding a certain value $W_{\min}$, has therefore a strong phenomenological motivation. This prior can be introduced as an additional penalty to the likelihood function for the PDFs which are *longer* in W. Note that a prior using the length in W as opposed to the length in x enhances sensitivity to the low x region. For the χ^2 function the prior corresponds to an extra penalty term of a form

$$\chi^2_{\text{prior}} = \alpha \left[\int_{W_{\min}}^{W_{\max}} \sqrt{1 + (xf'(W))^2} dW - (W_{\max} - W_{\min}) \right], \tag{9}$$

where α is the relative weight of this PDF-length prior and the PDF $xf = xg, xS$, respectively. The prior χ^2_{prior} has a minimum for the shortest PDF in W which corresponds to a condition for the derivative, $xf'(W) = 0$. In this case, $\chi^2_{\text{prior}} = 0$ holds irrespective of the value of α. The total χ^2_{tot} is given by the sum of the χ^2, for the data versus theory comparison, and the penalty term

$$\chi^2_{\text{tot}} = \chi^2 + \chi^2_{\text{prior}}. \tag{10}$$

We choose $W_{\max} = 320$ GeV which is the maximum value achievable at HERA. To stay far away from the resonance region, $W_{\min} = 10$ GeV is used which for $Q^2 = 1.9$ GeV^2 corresponds to $x \approx 0.02$. The prior is applied to both gluon and sea-quark densities at the starting scale $Q_0^2 = 1.9$ GeV^2 of the evolution.

4 Results

The Monte Carlo procedure of extracting PDFs is illustrated for $N_{par} \equiv N_g = N_S = 9$ in Fig. 1 which shows the gluon PDF at the starting scale $Q_0^2 = 1.9$ GeV^2 for each replica, with their RMS band. The distributions are compared to those obtained using the standard parameterization form: $xg(x) = A_g x^{B_g}(1-x)^{C_g}$, $xS(x) = A_s x^{B_s}(1-x)^{C_s}$, where A_g is determined by the momentum sum rule. The flexible parameterization does not suppress minima and maxima of the distribution as a function of x, as a result, several of them are observed.

Introducing the length prior to the fit by changing the weight of the penalty term from 0 GeV^{-1} to 1000 GeV^{-1} increases the χ^2 of the fit by 66 units, see Table 1. Further increase of the penalty term to $\alpha = 5000$ GeV^{-1} reduces fit quality considerably with an additional increase of χ^2 by 141

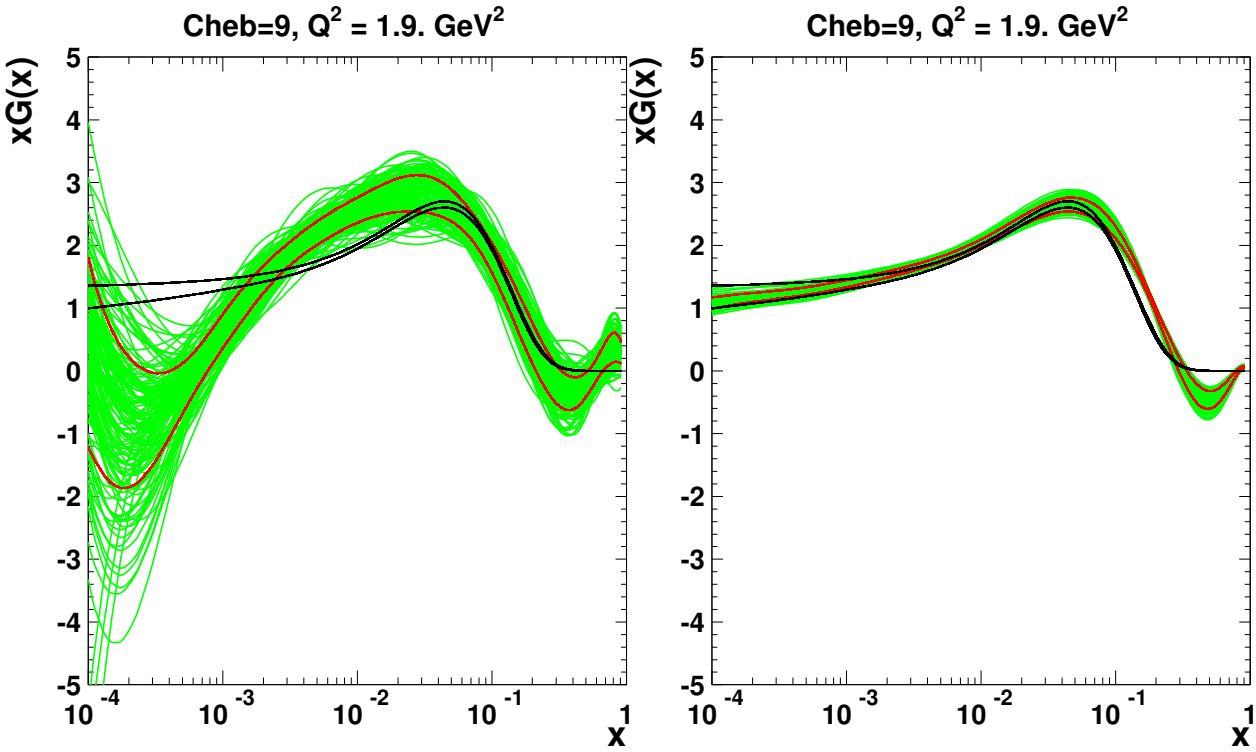

Figure 1: The gluon PDF $xg(x)$ at the starting scale $Q_0^2 = 1.9$ GeV^2. The green lines show fits to individual replicas of the data, the red lines show the RMS over the replicas. The black lines correspond to the error band of the gluon distribution using a standard parameterization and it is to be compared to the case of the Chebyshev parameterization. On the left hand side, the gluon distribution is shown using an unconstrained Chebyshev expansion to order nine, see Eq. (3), while on the right hand side the same distribution is displayed but with a tight length penalty $\alpha = 5000$ GeV^{-1} applied.

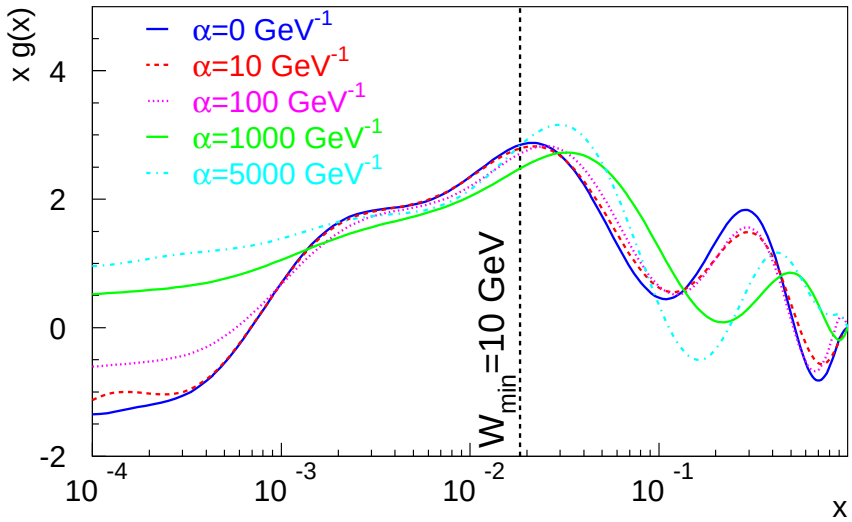

Figure 2: The central value of the gluon PDF for various values of the length-prior weight α at the evolution starting scale $Q^2 = 1.9$ GeV^2 using the Chebyshev parameterization expanded to the $15th$ order. The vertical dashed line shows the x value corresponding to the $W_{\rm min} = 10$ GeV limit.

Table 1: The quality of the fit in terms of $\chi^2/n_{\rm df}$ as a function of the length-prior weight α.

α, GeV^{-1}	0	10	100	1000	5000
$\chi^2/n_{\rm df}$	560/557	561/557	572/557	626/557	767/557

units. For low values of $\alpha \leq 100$, the impact of the penalty term on the shape of the central value of the gluon PDF is small while for $\alpha \geq 1000$ the distribution changes significantly, see Fig. 2. In addition, the shape of the gluon distribution using a standard parameterization can be reproduced by the Chebyshev parameterization of the gluon PDF if a tight length prior is applied to the fit, as demonstrated in Fig.1.

Fig. 1 shows that the PDF uncertainty is very large for $x < 0.0002$ and $x > 0.05$ while for intermediate values of x the data constrains the gluon PDF well. To quantify the dependence of the uncertainty as a function of $N_{g,S}$, the RMS values at fixed values of x=0.0002, 0.001, 0.01 and 0.1 are examined. The results of this investigation are summarized in Fig. 3. For x=0.01 and 0.001, the uncertainty stays approximately constant as the number of parameters increases while for $x = 0.0002$ and $x = 0.1$ it increases significantly, if no penalty term is used. The impact of the penalty term on the uncertainty of $xg(x)$ is large at low x. Even low values of $\alpha \leq 100$ GeV^{-1} significantly reduce the uncertainty for $x = 0.0002$ at large values of N_{par}, while for $x = 0.1$ the penalty term has no impact. For the intermediate x values, the change of the uncertainty is moderate, indicating that for $p \leq 100$ GeV^{-1} the data provides stronger constraints on the gluon PDF than the prior.

5 Summary

The focus of this study has been on the parameterization uncertainty of PDFs at low x, especially of the gluon PDF in a fit to the HERA I data at NLO in QCD. A flexible PDF parameterization based on Chebyshev polynomials has been chosen and the impact of an additional smoothness prior on the quality of the has been investigated. We have found that the uncertainty of the fit is generally small in the $0.0005 < x < 0.05$ range. The uncertainty, however, increases significantly for larger and smaller x values. The regularization with a smoothness prior, which disfavors resonant structures for large values of W allows to significantly reduce uncertainty also for the range $x < 0.0005$.

Acknowledgments

We are thankful to W. Giele for suggesting the length prior and for valuable discussions about quantifying parton distribution function uncertainties [24]. We are also thankful to S. Alekhin for interesting discussions. This work has been supported in part by Helmholtz Gemeinschaft under contract VH-HA-101 (Alliance *Physics at the Terascale*).

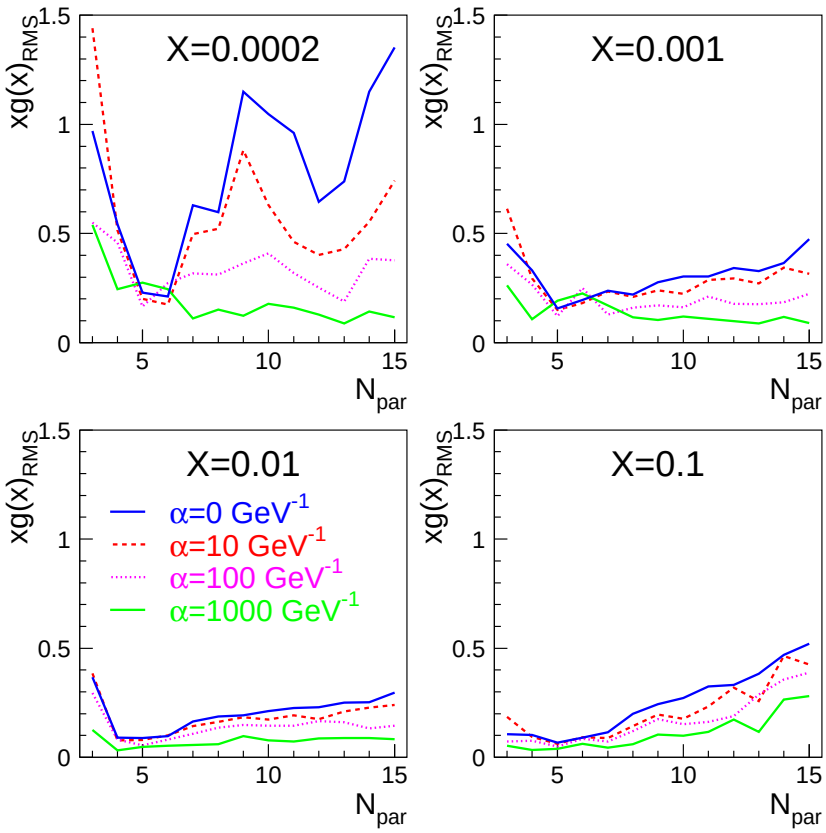

Figure 3: The uncertainty of the gluon PDF as a function of $\mathrm{N_{par}} \equiv N_g = N_S$ for $Q^2 = 1.9$ GeV^2 at fixed values of x for different values of the length-prior weight α as indicated by the figure's legend.

References

[1] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys.Rev. **D69**, 094008 (2004), arXiv:hep-ph/0312266.

[2] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, Phys.Rev.Lett. **103**, 082001 (2009), arXiv:0903.2120.

[3] V. Gribov and L. Lipatov, Sov.J.Nucl.Phys. **15**, 438 (1972).

[4] G. Altarelli and G. Parisi, Nucl.Phys. **B126**, 298 (1977).

[5] G. Curci, W. Furmanski, and R. Petronzio, Nucl.Phys. **B175**, 27 (1980).

[6] W. Furmanski and R. Petronzio, Phys.Lett. **B97**, 437 (1980).

[7] S. Moch, J. Vermaseren, and A. Vogt, Nucl.Phys. **B688**, 101 (2004), arXiv:hep-ph/0403192.

[8] A. Vogt, S. Moch, and J. Vermaseren, Nucl.Phys. **B691**, 129 (2004), arXiv:hep-ph/0404111.

[9] C. Pascaud and F. Zomer, *QCD analysis from the proton structure function F_2 measurement: Issues on fitting, statistical and systematic errors* (1995).

[10] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, and W. K. Tung, JHEP **0207**, 012 (2002), arXiv:hep-ph/0201195.

[11] Z. Ajaltouni *et al.*, (2009), arXiv:0903.3861.

[12] M. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, Phys.Rev. **D50**, 3102 (1994), arXiv:hep-ph/9312319.

[13] M. Buza, Y. Matiounine, J. Smith, and W. van Neerven, Eur.Phys.J. **C1**, 301 (1998), arXiv:hep-ph/9612398.

[14] H1 Collaboration, C. Adloff *et al.*, Eur.Phys.J. **C30**, 1 (2003), arXiv:hep-ex/0304003.

[15] H1 Collaboration, F. Aaron *et al.*, Eur.Phys.J. **C64**, 561 (2009), arXiv:0904.3513.

[16] R. D. Ball *et al.*, (2010), arXiv:1002.4407.

[17] H1 Collaboration, ZEUS Collaboration, F. Aaron *et al.*, JHEP **1001**, 109 (2010), arXiv:0911.0884.

[18] M. Botje, (2010), arXiv:1005.1481, http://www.nikef.nl/h24/qcdnum/index.html.

[19] ZEUS Collaboration, S. Chekanov *et al.*, Eur.Phys.J. **C42**, 1 (2005), arXiv:hep-ph/0503274.

[20] S. Alekhin, J. Blümlein, S. Klein, and S. Moch, Phys.Rev. **D81**, 014032 (2010), arXiv:0908.2766.

[21] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, (2009), arXiv:0901.0002.

[22] P. M. Nadolsky *et al.*, Phys. Rev. **D78**, 013004 (2008), arXiv:0802.0007.

[23] CLAS Collaboration, M. Osipenko *et al.*, Phys. Rev. **C73**, 1 (2006).

[24] W. T. Giele, S. A. Keller, and D. A. Kosower, (2001), arXiv:hep-ph/0104052.