DGLAP evolution and parton fits

S. I. Alekhin, J. Blümlein, H. Böttcher, L. Del Debbio, S. Forte, A. Glazov, A. Guffanti, J. Huston, G. Ingelman J. I. Latorre, S. Moch, A. Piccione, J. Pumplin, V. Ravindran, J. Rojo G.P. Salam, R.S. Thorne, J.A.M. Vermaseren, A. Vogt

1 DGLAP evolution and parton fits [1]

1.1 Introduction

The high-precision data from HERA and the anticipated data from LHC open the possibility for a precise determination of parton distributions. This, however, requires an improvement in the theoretical description of DIS and hard hadronic scattering processes, as well as an improvement of the techniques used to extract parton distributions from the data.

The determination of perturbative QCD corrections has undergone substantial progress recently. The key ingredient of a complete next-to-next-to-leading order (NNLO) prediction in perturbative QCD are the recently calculated three-loop splitting functions which govern the scale dependence of PDFs. Extensions in the accuracy of the perturbative predictions yet beyond NNLO are given by the three-loop coefficient functions for F_2, while the coefficient functions for F_L at this order are actually required to complete the NNLO predictions. Section 2 briefly discusses the recent results and their phenomenological implications. Certain mathematical aspects, which are important in the calculation of higher order corrections in massless QCD are presented in section 3. In particular, algebraic relations in Mellin-N space are pointed out, which are of importance for harmonic sums, harmonic polylogarithms and multiple ζ-values.

These calculation of the PDF evolution to NNLO in perturbative QCD are used in section 4 to provide an update and extension of a set of benchmark tables for the evolution of parton distributions of hadrons. These benchmark tables were first presented in the report of the QCD/SM working group at the 2001 Les Houches workshop, but based on approximate NNLO splitting functions, which are superseded by the exact results which are now available. In addition, section 4 now includes also reference tables for the case of polarized PDF evolution.

Whereas in principle the x-shapes of PDFs at low scales can be determined from first principles using non-perturbative methods, in practice at present this is only possible using models (briefly touched in in section 5). Therefore, an accurate determination of PDFs requires a global QCD fit to the data, which is the subject of sections 6–8.

Section 6 discusses in particular the impact on parton fits of NNLO corrections on the one hand, and of the inclusion of Drell-Yan data and future LHC data on the other hand. It then presents values for a benchmark fit together with a table of correlation coefficients for the parameter obtained in the fit. This benchmark fit is then re-examined in sec. 7, along with a comparison between PDFs and the associated uncertainty obtained using the approaches of Alekhin and the MRST group. The differences between these benchmark partons and the actual global fit partons are also discussed, and used to explore complications inherent in extracting PDFs with uncertainties. Finally, in section 8 the stability of PDF determinations in NLO global analyses is re-investigated and the results of the CTEQ PDF group on this issue are summarized.

An alternative approach to a completely bias-free parameterization of PDFs is presented in section 9. There, a neural network approach to global fits of parton distribution functions is introduced and work on unbiased parameterizations of deep-inelastic structure functions with faithful estimation of their uncertainties is reviewed together with a summary of the current status of neural network parton distribution fits.

[1] Subsection coordinators: A. Glazov, S. Moch

2 Precision Predictions for Deep-Inelastic Scattering [2]

With high-precision data from HERA and in view of the outstanding importance of hard scattering processes at the LHC, a quantitative understanding of deep-inelastic processes is indispensable, necessitating calculations beyond the standard next-to-leading order of perturbative QCD.

In this contribution we review recent results for the complete next-to-next-to-leading order (NNLO, N^{2}LO) approximation of massless perturbative QCD for the structure functions F_1, F_2, F_3 and F_L in DIS. These are based on the second-order coefficient functions [1–5], the three-loop splitting functions which govern the evolution of unpolarized parton distributions of hadrons [6, 7] and the three-loop coefficient functions for $F_L = F_2 - 2xF_1$ in electromagnetic (photon-exchange) DIS [8, 9]. Moreover we discuss partial N^{3}LO results for F_2, based on the corresponding three-loop coefficient functions also presented in Ref. [9]. For the splitting functions P and coefficient functions C we employ the convention

$$P(\alpha_{\rm s}) \,=\, \sum_{n=0} \left(\frac{\alpha_{\rm s}}{4\pi}\right)^{n+1} P^{(n)}\,, \qquad C(\alpha_{\rm s}) \,=\, \sum_{n=0} \left(\frac{\alpha_{\rm s}}{4\pi}\right)^{n} C^{(n)} \tag{1}$$

for the expansion in the running coupling constant $\alpha_{\rm s}$. For the longitudinal structure function F_L the third-order corrections are required to complete the NNLO predictions, since the leading contribution to the coefficient function C_L is of first order in the strong coupling constant $\alpha_{\rm s}$.

In the following we briefly display selected results to demonstrate the quality of precision predictions for DIS and their effect on the evolution. The exact (analytical) results to third order for the quantities in Eq. (1) are too lengthy, about $\mathcal{O}(100)$ pages in normalsize fonts and will not be reproduced here. Also the method of calculation is well documented in the literature [5–7, 9–11]. In particular, it proceeds via the Mellin transforms of the functions of the Bjorken variable x,

$$A(N) \;\;=\;\; \int\limits_0^1 dx\, x^{N-1} A(x)\,. \tag{2}$$

Selected mathematical aspects of Mellin transforms are discussed in section 3.

2.1 Parton evolution

The well-known $2n_f - 1$ scalar non-singlet and 2×2 singlet evolution equations for n_f flavors read

$$\frac{d}{d\ln \mu_f^2}\, q_{\rm ns}^i \;=\; P_{\rm ns}^i \,\otimes\, q_{\rm ns}^i\,, \qquad i = \pm, {\rm v}\,, \tag{3}$$

for the quark flavor asymmetries $q_{\rm ns}^{\pm}$ and the valence distribution $q_{\rm ns}^{\rm v}$, and

$$\frac{d}{d\ln \mu_f^2} \left(\begin{array}{c} q_{\rm s} \\ g \end{array} \right) \;=\; \left(\begin{array}{cc} P_{\rm qq} & P_{\rm qg} \\ P_{\rm gq} & P_{\rm gg} \end{array} \right) \otimes \left(\begin{array}{c} q_{\rm s} \\ g \end{array} \right) \tag{4}$$

for the singlet quark distribution $q_{\rm s}$ and the gluon distribution g, respectively. Eqs. (3) and (4) are governed by three independent types of non-singlet splitting functions, and by the 2×2 matrix of singlet splitting functions. Here $\otimes$ stands for the Mellin convolution. We note that benchmark numerical solutions to NNLO accuracy of Eqs. (3) and (4) for a specific set of input distributions are given in section 4. Phenomenological QCD fits of parton distributions in data analyses are extensively discussed in sections 6–8. An approach based on neural networks is described in section 9.

Let us start the illustration of the precision predictions by looking at the parton evolution and at large Mellin-N (large Bjorken-x) behavior. Fig. 1 shows the stability of the perturbative expansion which

[2] Contributing authors: S. Moch, J.A.M. Vermaseren, A. Vogt

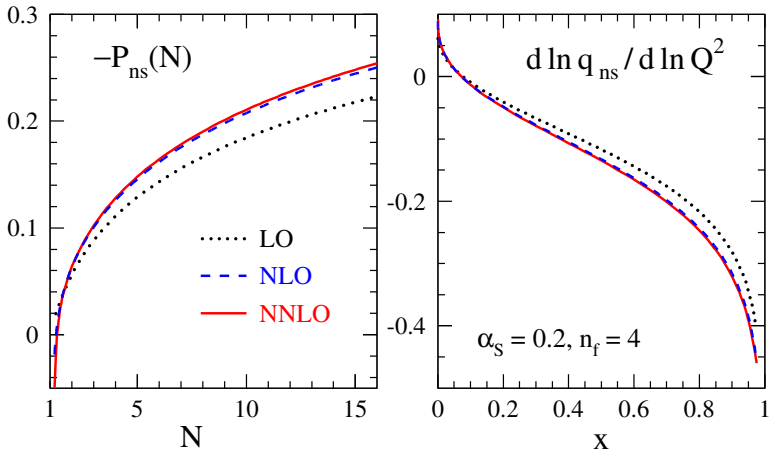

Fig. 1: On the left we show the perturbative expansion of $P_{\rm ns}^{\rm v}(N)$, and on the right the resulting perturbative expansion of the logarithmic scale derivative $d\ln q_{\rm ns}/d\ln\mu_f^2$ is displayed for a model input. See the text for details.

is very benign and indicates, for $\alpha_{\rm s} \lesssim 0.2$, corrections of less than 1% beyond NNLO. On the left we show the results for the perturbative expansion of $P_{\rm ns}$ in Mellin space, cf. Eqs. (1), (2). We employ four active flavors, $n_f = 4$, and an order-independent value for the strong coupling constant,

$$\alpha_{\rm s}(\mu_0^2) \;=\; 0.2 \;, \tag{5}$$

which corresponds to $\mu_0^2 \simeq 25\ldots50\,{\rm GeV}^2$ for $\alpha_{\rm s}(M_Z^2) = 0.114\ldots0.120$ beyond the leading order. On the right of Fig. 1 the perturbative expansion of the logarithmic derivative, cf. Eqs. (1), (3), is illustrated at the standard choice $\mu_r = \mu_f$ of the renormalization scale. We use the schematic, but characteristic model distribution,

$$xq_{\rm ns}(x,\mu_0^2) \;\;=\;\; x^{\,0.5}(1-x)^3\,. \tag{6}$$

The normalization of $q_{\rm ns}$ is irrelevant at this point, as we consider the logarithmic scale derivative only.

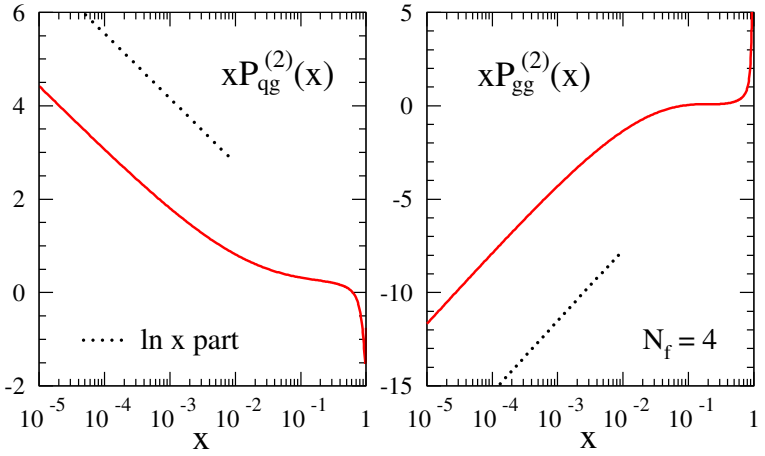

Fig. 2: The three-loop gluon-quark (left) and gluon-gluon (right) splitting functions together with the leading small-x contribution (dotted line).

Next, let us focus on the three-loop splitting functions at small momentum fractions x, where the splitting functions $P_{\rm ig}$ in the lower row of the 2×2 matrix in Eq. (4), representing ${\rm g}\to{\rm i}$ splittings, are most important. In Fig. 2 we show, again for $n_f = 4$, the three-loop splitting functions $P_{\rm qg}^{(2)}$ and $P_{\rm gg}^{(2)}$

together with the leading small-x term indicated separately for $x < 0.01$. In the present singlet case the leading logarithmic small-x limits $\sim x^{-1} \ln x$ of Refs. [12, 13] are confirmed together with the general structure of the BFKL limit [14–16]. The same holds for the leading small-x terms $\ln^4 x$ in the non-singlet sector [17, 18], with the qualification that a new, unpredicted leading logarithmic contribution is found for the color factor $d^{abc}d_{abc}$ entering at three loops for the first time.

It is obvious from Fig. 2 (see also Refs. [5–7, 11]) that the leading $x \to 0$-terms alone are insufficient for collider phenomenology at HERA or the LHC as they do not provide good approximations of the full results at experimentally relevant small values of x. Resummation of the small-x terms and various phenomenological improvements are discussed in detail in [19].

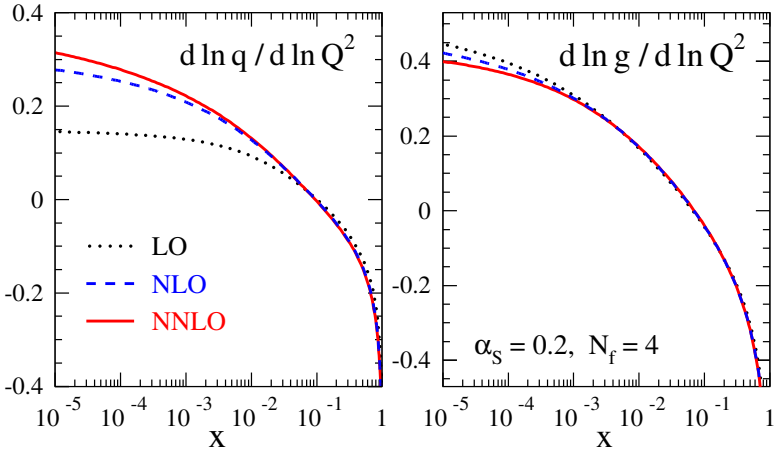

Fig. 3: The perturbative expansion of the scale derivatives (4) of the singlet distributions (7).

In the same limit of small x, it is instructive to look at the evolution of parton distributions. Again, we choose the reference scale of Eq. (5), $n_f = 4$ and the sufficiently realistic model distributions

$$\begin{aligned} xq_{\rm s}(x,\mu_0^2) &= 0.6\, x^{-0.3}(1-x)^{3.5}\,(1+5.0\, x^{0.8}) \\ xg(x,\mu_0^2) &= 1.6\, x^{-0.3}(1-x)^{4.5}\,(1-0.6\, x^{0.3}) \end{aligned} \tag{7}$$

irrespective of the order of the expansion to facilitate direct comparisons of the various contributions. Of course, this order-independence does not hold for actual data-fitted parton distributions like those in sections 6–8. In Fig. 3 we display the perturbative expansion of the scale derivative for the singlet quark and gluon densities at $\mu_f^2 = \mu_0^2$ for the initial conditions specified in Eqs. (5) and (7). For the singlet quark distribution the total NNLO corrections, while reaching 10% at $x = 10^{-4}$, remain smaller than the NLO results by a factor of eight or more over the full x-range. For the gluon distribution already the NLO corrections are small and the NNLO contribution amounts to only 3% for x as low as 10^{-4}. Thus, we see in Fig. 3 that the perturbative expansion is very stable. It appears to converge rapidly at $x > 10^{-3}$, while relatively large third-order corrections are found for very small momenta $x \lesssim 10^{-4}$.

2.2 Coefficient functions

While the previous considerations were addressing the evolution of parton distributions, we now turn to the further improvements of precision predictions due to the full third-order coefficient functions for the structure functions F_2 and F_L in electromagnetic DIS [8, 9]. The results for F_L complete the NNLO description of unpolarized electromagnetic DIS, and the third-order coefficient functions for F_2 form, at not too small values of the Bjorken variable x, the dominant part of the next-to-next-to-next-to-leading order ($\rm N^3LO$) corrections. Thus, they facilitate improved determinations of the strong coupling $\alpha_{\rm s}$ from scaling violations.

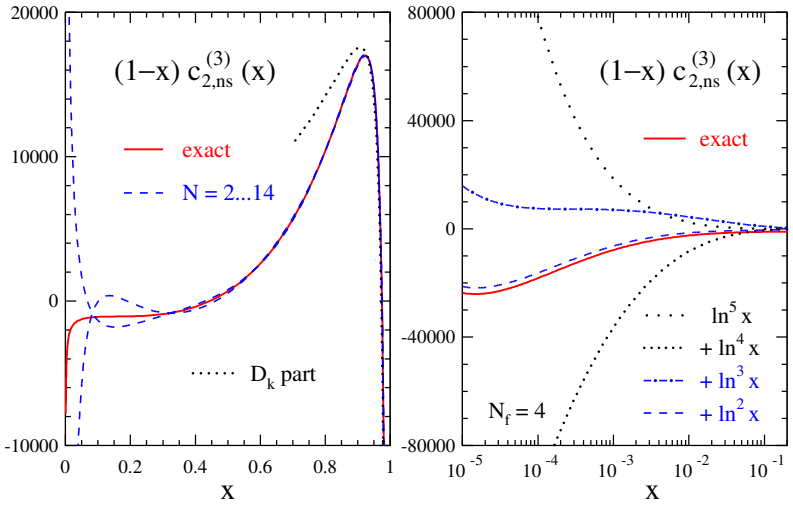

Fig. 4: The three-loop non-singlet coefficient function $c^{(3)}_{2,\rm ns}(x)$ in the large-x (left) and the small-x (right) region, multiplied by $(1-x)$ for display purposes.

Let us start with the three-loop coefficient functions for F_2 in the non-singlet case. In Fig. 4 we display the three-loop non-singlet coefficient function $c^{(3)}_{2,\rm ns}(x)$ for $n_f = 4$ flavors. We also show the soft-gluon enhanced terms $\mathcal{D}_k$ dominating the large-x limit,

$$\mathcal{D}_k \;=\; \frac{\ln^{2k-1}(1-x)}{(1-x)_+}\,, \tag{8}$$

and the small-x approximations obtained by successively including enhanced logarithms $\ln^k x$. However the latter are insufficient for an accurate description of the exact result. The dashed band in Fig. 4 shows the uncertainty of previous estimates [20] mainly based on the calculation of fixed Mellin moments [21–23]. For a detailed discussion of the soft-gluon resummation of the the $\mathcal{D}_k$ terms, we refer to [19].

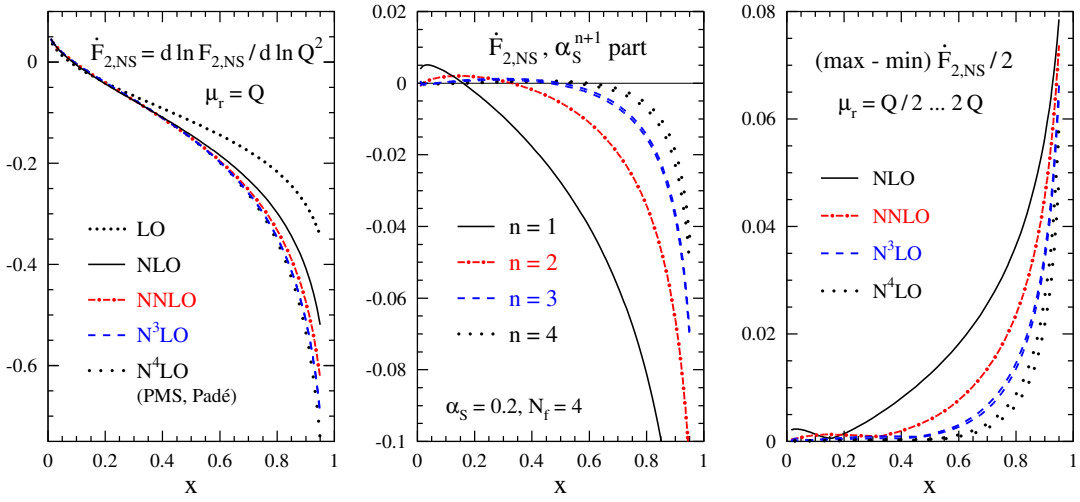

Fig. 5: The perturbative expansion of the logarithmic scale derivative of the non-singlet structure function $F_{2,\rm ns}$. The results up to NNLO are exact, while those at N^{3}LO are very good approximations. The N^{4}LO corrections have been estimated by various methods.

Building on the coefficient functions, it is interesting to study the perturbative expansion of the logarithmic scale derivative for the non-singlet structure function $F_{2,\rm ns}$. To that end we use in Fig. 5

again the input shape Eq. (6) (this time for $F_{2,\rm ns}$ itself) irrespective of the order of the expansion, $n_f = 4$ flavors and the reference scale of Eq. (5). The N^{4}LO approximation based on Padé summations of the perturbation series can be expected to correctly indicate at least the rough size of the four-loop corrections, see Ref. [9] for details. From Fig. 5 we see that the three-loop results for F_2 can be employed to effectively extend the main part of DIS analyses to the N^{3}LO at $x > 10^{-2}$ where the effect of the unknown fourth-order splitting functions is expected to be very small. This has, for example, the potential for a 'gold-plated' determination of $\alpha_{\rm s}(M_Z)$ with an error of less than 1% from the truncation of the perturbation series. On the right hand side of Fig. 5 the scale uncertainty which is conventionally estimated by

$$\Delta \dot{f} \;\equiv\; \frac{1}{2}\left(\max\left[\dot{f}(x,\mu_r^2)\right] - \min\left[\dot{f}(x,\mu_r^2)\right]\right) , \tag{9}$$

is plotted, where the scale varies $\mu_r \in [Q/2, 2Q]$.

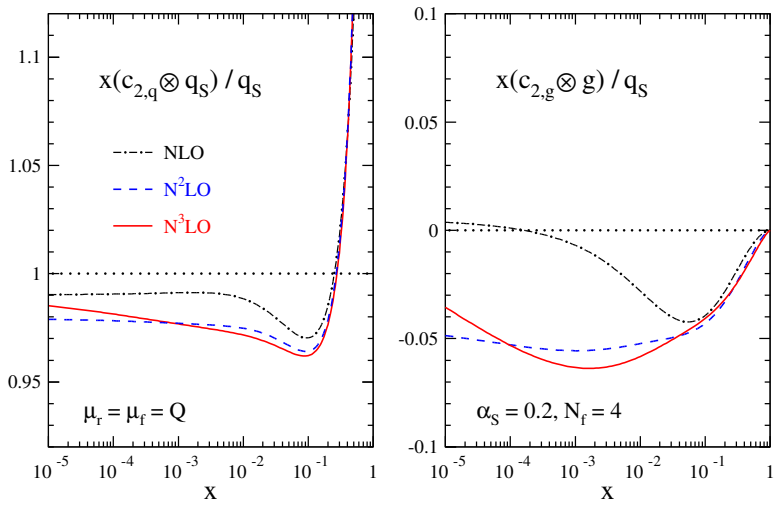

Fig. 6: The perturbative expansion up to three loops (N^{3}LO) of the quark (left) and gluon (right) contributions to singlet structure function F_2.

In the singlet case, we can study the quark and gluon contributions to the structure function F_2. In Fig. 6 we plot the perturbative expansion up to N^{3}LO of the quark and gluon contributions to structure function $F_{2,\rm s}$ at the scale (5) using the distributions (7). All curves have been normalized to the leading-order result $F_{2,\rm s}^{\rm LO} = \langle e^2 \rangle \, q_{\rm s}$. Fig. 6 nicely illustrates the perturbative stability of the structure function F_2.

Finally, we address the longitudinal structure function F_L at three loops. In the left part of Fig. 7 we plot the singlet-quark and gluon coefficient functions $c_{L,\rm q}$ and $c_{L,\rm g}$ for F_L up to the third order for four flavors and the $\alpha_{\rm s}$-value of Eq. (5). The curves have been divided by $a_{\rm s} = \alpha_s/(4\pi)$ to account for the leading contribution being actually of first order in the strong coupling constant $\alpha_{\rm s}$. Both the second-order and the third-order contributions are rather large over almost the whole x-range. Most striking, however, is the behavior at very small values of x, where the anomalously small one-loop parts are negligible against the (negative) constant two-loop terms, which in turn are completely overwhelmed by the (positive) new three-loop corrections $xc_{L,a}^{(3)} \sim \ln x + const$, which we have indicated in Fig. 7.

To assess the effect for longitudinal structure function F_L, we convolute in Fig. 7 on the right the coefficient functions with the input shapes Eq. (7) for $n_f = 4$ flavors and the reference scale of Eq. (5). A comparison of the left and right plots in Fig. 7 clearly reveals the smoothening effect of the Mellin convolutions. For the chosen input conditions, the (mostly positive) NNLO corrections to the flavor-singlet F_L amount to less than 20% for $5 \cdot 10^{-5} < x < 0.3$. In data fits we expect that the parton

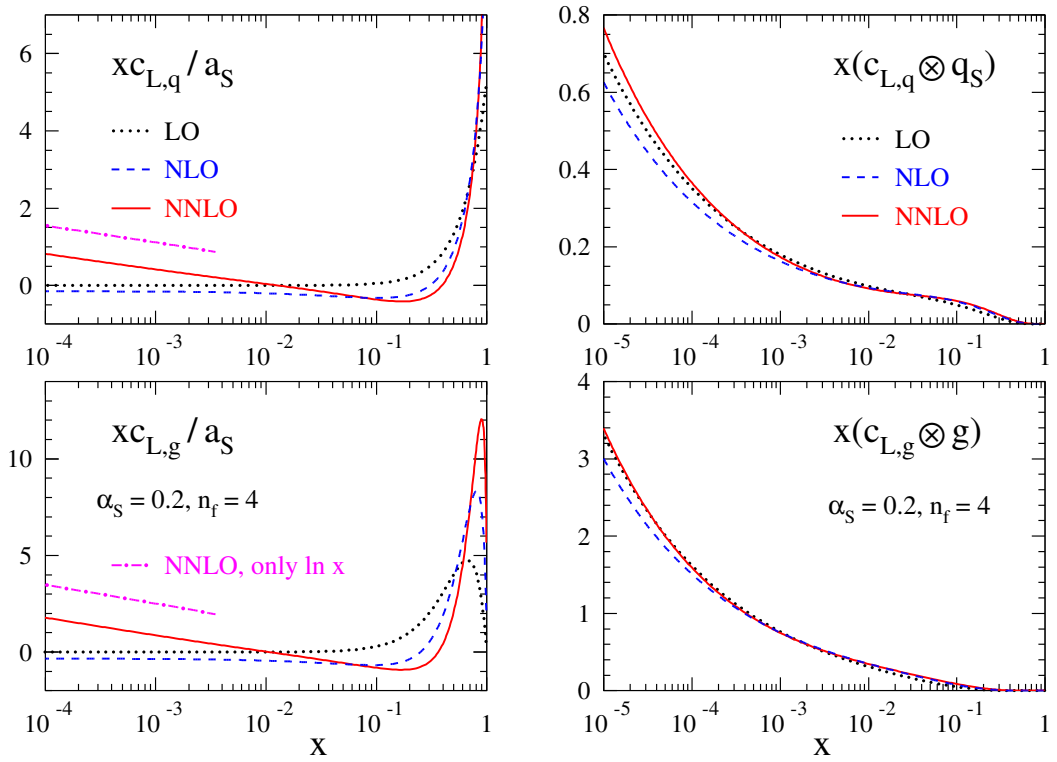

Fig. 7: The perturbative expansion to N^{2}LO of the longitudinal singlet-quark and gluon coefficient functions to third order multiplied by x for display purposes (left) and of the quark and gluon contributions to singlet structure function F_L (right).

Table 1: Number of alternating and non-alternating harmonic sums in dependence of their weight, [28].

	Number of					
Weight	Sums	a-basic sums	Sums $\neg\{-1\}$	a-basic sums	Sums $i > 0$	a-basic sums
1	2	2	1	1	1	1
2	6	3	3	2	2	1
3	18	8	7	4	4	2
4	54	18	17	7	8	3
5	162	48	41	16	16	6
6	486	116	99	30	32	9
7	1458	312	239	68	64	18

distributions, in particular the gluon distribution, will further stabilize the overall NNLO/NLO ratio. Thus, at not too small scales, F_L is a quantity of good perturbative stability, for the x-values accessible at HERA, see Ref. [8] for more details.

3 Mathematical Structure of Higher Order Corrections [3]

The QCD anomalous dimensions and Wilson coefficients for structure functions are single scale quantities and may be expressed in simple form in Mellin space in terms of polynomials of harmonic sums

[3] Contributing authors: J. Blümlein, H. Böttcher, A. Guffanti, V. Ravindran

and ration functions of the Mellin variable. Unlike the case in various calculations using representations in momentum-fraction (z-) space the use of multiple nested harmonic sums leads to a synchronization in language. Furthermore, significant simplifications w.r.t. the number of functions needed can be achieved. This is due to algebraic [24,25] relations between these quantities, which in a similar way are also present between harmonic polylogarithms [26] and multiple ζ-values [27]. These relations result from the the specific index pattern of the objects considered and their multiplication relation and do not refer to further more specific properties. In Table 1 we illustrate the level of complexity which one meets in case of harmonic sums. To three-loop order weight w=6 harmonic sums occur. The algebraic relations for the whole class of harmonic sums lead to a reduction by a factor of ~ 4 (column 3). As it turns out, physical pseudo-observables, as anomalous dimensions and Wilson-coefficients in the $\overline{\rm MS}$ scheme, to 2-, resp. 3-loop order depend on harmonic sums only, in which the index $\{-1\}$ never occurs. The algebraic reduction for this class is illustrated in column 5. We also compare the complexity of only non-alternating harmonic sums and their algebraic reduction, which is much lower. This class of sums is, however, not wide enough to describe the above physical quantities. In addition to the algebraic relations of harmonic sums structural relations exist, which reduces the basis further [28]. Using all these relations one finds that 5 basic functions are sufficient to describe all 2-loop Wilson coefficients for deep-inelastic scattering [29] and further 8 [30] for the 3-loop anomalous dimensions. Their analytic continuations to complex values of the Mellin variable are given in [31, 32]. These functions are the (regularized) Mellin transforms of :

$$\begin{array}{llll} \dfrac{\ln(1+x)}{1+x}, & \dfrac{{\rm Li}_2(x)}{1\pm x}, & \dfrac{S_{1,2}(x)}{1\pm x}, & \dfrac{{\rm Li}_4(x)}{x\pm 1}, \\ \dfrac{S_{1,3}(x)}{1+x}, & \dfrac{S_{2,2}(x)}{x\pm 1}, & \dfrac{{\rm Li}_2^2(x)}{1+x}, & \dfrac{S_{2,2}(-x)-{\rm Li}_2^2(-x)/2}{x\pm 1}\,. \end{array} \tag{10}$$

It is remarkable, that the numerator-functions in (10) are Nielsen integrals [33] and polynomials thereof, although one might expect harmonic polylogarithms [26] outside this class in general. The representation of the Wilson coefficients and anomalous dimensions in the way described allows for compact expressions and very fast and precise numerical evaluation well suited for fitting procedures to experimental data.

3.1 Two-loop Processes at LHC in Mellin Space

Similar to the case of the Wilson coefficients in section 3 one may consider the Wilson coefficients for inclusive hard processes at hadron colliders, as the Drell–Yan process to $O(\alpha_s^2)$ [34–36], scalar or pseudoscalar Higgs-boson production to $O(\alpha_s^3)$ in the heavy-mass limit [37–42], and the 2-loop time-like Wilson coefficients for fragmentation [43–45]. These quantities have been analyzed in [46, 47] w.r.t. their general structure in Mellin space. The cross section for the Drell–Yan process and Higgs production is given by

$$\sigma\left(\frac{\hat{s}}{s},Q^2\right) = \int_x^1 \frac{dx_1}{x_1}\int_{x/x_1}^1 \frac{dx_2}{x_2} f_a(x_1,\mu^2) f_b(x_2,\mu^2)\hat{\sigma}\left(\frac{x}{x_1x_2},\frac{Q^2}{\mu^2}\right)\,, \tag{3.11}$$

with $x=\hat{s}/s$. Here, $f_c(x,\mu^2)$ are the initial state parton densities and μ^2 denotes the factorization scale. The Wilson coefficient of the process is $\hat{\sigma}$ and Q^2 is the time-like virtuality of the s-channel boson. Likewise, for the fragmentation process of final state partons into hadrons in pp–scattering one considers the double differential final state distribution

$$\frac{d^2\sigma^H}{dx d\cos\theta} = \frac{3}{8}(1+\cos^2\theta)\frac{d\sigma_T^H}{dx} + \frac{3}{4}\sin^2\theta\frac{d\sigma_L^H}{dx}\,. \tag{3.12}$$

Here,

$$\frac{d\sigma_k^H}{dx} = \int_x^1 \frac{dz}{z}\Bigg[\sigma_{\rm tot}^{(0)}\left\{D_S^H\left(\frac{x}{z},M^2\right)C_{k,q}^{\rm S}(z,Q^2/M^2) + D_g^H\left(\frac{x}{z},M^2\right)C_{k,q}^{\rm S}(z,Q^2/M^2)\right\}$$

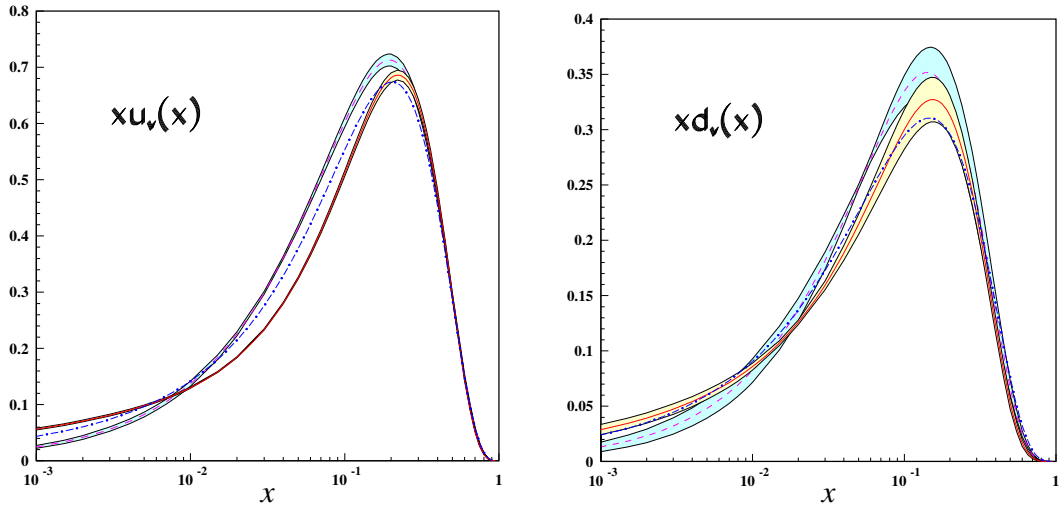

Fig. 8: xu_v and xd_v at $Q_0^2 = 4\mathrm{GeV}^2$ (full lines) [48]; dashed lines [50]; dash-dotted lines [51].

$$+\sum_{p=1}^{N_f} \sigma_p^{(0)} D^H_{\mathrm{NS},p}\left(\frac{x}{z}, M^2\right) C^{\mathrm{NS}}_{k,q}(z, Q^2/M^2)\Bigg] . \tag{3.13}$$

In the subsystem cross-sections σ the initial state parton distributions are included. D_k^H denote the non-perturbative fragmentation functions and $C^{\mathrm{S,NS}}_{k,i}(z, Q^2/M^2)$ the respective time-like Wilson coefficients describing the fragmentaion process for a parton i into the hadron H.

Although these Wilson coefficients are not directly related to the 2-loop Wilson coefficients for deeply inelastic scattering, one finds for these functions at most the same set of basic functions as given above. Again one obtains very fast and concise numerical programs also for these processes working in Mellin space, which will be well suited for inclusive analyses of experimental collider data at LHC in the future.

3.2 Non-Singlet Parton Densities at $O(\alpha_s^3)$

The precision determination of the QCD-scale Λ_{QCD} and of the idividual parton densities is an important issue for the whole physics programme at LHC since all measurements rely on the detailed knwoledge of this parameter and distribution functions. In Ref. [48] first results were reported of a world data analysis for charged lepton-$p(d)$ scattering w.r.t. the flavor non-singlet sector at $O(\alpha_s^3)$ accuracy. The flavor non-singlet distributions $xu_v(x, Q^2)$ and $xd_v(x, Q^2)$ were determined along with fully correlated error bands giving parameterizations both for the values and errors of these distributions for a wide range in x and Q^2. In Figure 8 these distributions including their error are shown. The value of the strong coupling constant $\alpha_s(M_Z^2)$ was determined as $0.1135 + 0.0023 - 0.0026$ (exp.) The full analysis is given in [49], including the determination of higher twist contributions in the large x region both for $F_2^p(x, Q^2)$ and $F_2^d(x, Q^2)$.

3.3 Scheme-invariant evolution for unpolarzed DIS structure functions

The final HERA-II data on unpolarized DIS structure functions, combined with the present world data from other experiments, will allow to reduce the experimental error on the strong coupling constant, $\alpha_s(M_Z^2)$, to the level of 1% [52]. On the theoretical side the NLO analyzes have intrinsic limitations which allow no better than 5% accuracy in the determination of α_s [53]. In order to match the expected experimental accuracy, analyzes of DIS structure functions need then to be carried out at the NNLO-level. To perform a full NNLO analysis the knowledge of the 3-loop β-function coefficient, β_2, the 2-

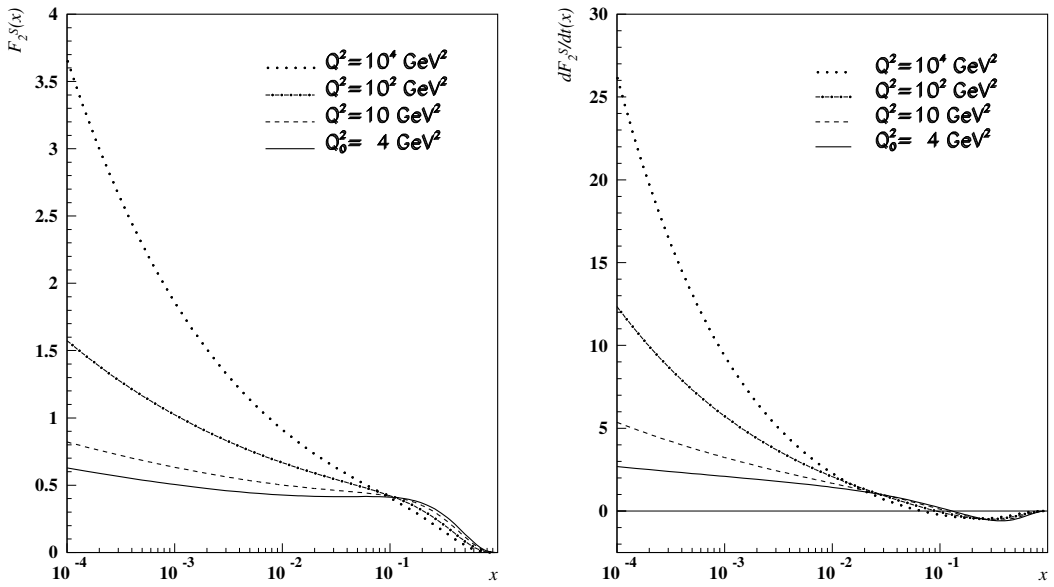

Fig. 9: NNLO scheme invariant evolution for the singlet part of the structure function F_2 and its slope $\partial F_2/\partial t$ for four massless flavours, [54].

resp. 3-loop Wilson coefficients and the 3-loop anomalous dimensions is required. With the calculation of the latter [6, 7], the whole scheme-independent set of quantities is known, thus allowing a complete NNLO study of DIS structure functions.

Besides the standard approach solving the QCD evolution equations for parton densities in the $\overline{\rm MS}$ scheme it appears appealing to study scheme-invariant evolution equations [54]. Within this approach the input distributions at a scale Q_0^2 are measured experimentally. The only parameter to be determined by a fit to data is the QCD-scale $\Lambda_{\rm QCD}$. To perform an analysis in the whole kinematic region the non-singlet [48] contribution has to be separated from the singlet terms of two measured observables. In practice these can be chosen to be $F_2(x,Q^2)$ and $\partial F_2(x,Q^2)/\partial\ln(Q^2)$ or $F_2(x,Q^2)$ and $F_L(x,Q^2)$ if the latter structure function is measured well enough. Either $\partial F_2(x,Q^2)/\partial\ln(Q^2)$ or $F_L(x,Q^2)$ play a role synonymous to the gluon distribution while $F_2(x,Q^2)$ takes the role of the singlet-quark distribution compared to the standard analysis. These equations do no longer describe the evolution of universal quantities depending on the choice of a scheme but of process-dependent quantities which are observables and thus factorization scheme-indedependent. Since the respective evolution kernels are calculated in perturbation theory the dependence on the renormalization scale remains and becomes smaller with the order in the coupling constant included.

Physical evolution kernels have been studied before in [55–57]. The 3-loop scheme-invariant evolution equations were solved in the massless case in [54]. This analysis is extended including the heavy flavor contributions at present [49]. The large complexity of the evolution kernels can only be handeled in Mellin space since in z-space various inverse and direct Mellin convolutions would be required numerically, causing significant accuracy and run-time problems. The inclusion of the heavy flavor contributions is possible using the parameterizations [58].

In Fig. 8 we present the scheme invariant evolution for the structure functions F_2 and $\partial F_2/\partial t$ to NNLO with $t = -2/\beta_0 \ln(\alpha_s(Q^2)/\alpha_s(Q_0^2))$. The input distribution at the reference scale are not extracted from data, but rather built up as a convolution of Wilson coefficients and PDFs, the latter being parametrised according to [59].

Scheme-invariant evolution equations allow a widely un-biased approach to determine the initial conditions for QCD evolution, which in general is a source of systematic effects which are difficult to control. On the other hand, their use requires to consider all correlations of the input measurements in

a detailed manner experimentally. At any scale Q^2 mappings are available to project the observables evolved onto the quark-singlet and the gluon density in whatever scheme. In this way the question whether sign changes in the unpolarized gluon distribution in the $\overline{\rm MS}$ scheme do occur or do not occur in the small x region can be answered uniquely. As in foregoing analyses [48, 60] correlated error propagation throughout the evolution is being performed.

4 Updated reference results for the evolution of parton distributions [4]

In this contribution we update and extend our benchmark tables, first presented in the report of the QCD/SM working group at the 2001 Les Houches workshop [59], for the evolution of parton distributions of hadrons in perturbative QCD. Since then the complete next-to-next-to-leading order (NNLO) splitting functions have been computed [6,7], see also section 2. Thus we can now replace the NNLO results of 2001 which were based on the approximate splitting functions of Ref. [61]. Furthermore we now include reference tables for the polarized case treated in neither Ref. [59] nor the earlier study during the 1995/6 HERA workshop [62]. Since the spin-dependent NNLO splitting functions are still unknown, we have to restrict ourselves to the polarized leading-order (LO) and next-to-leading-order (NLO) evolution.

As in Ref. [59], we employ two entirely independent and conceptually different FORTRAN programs. At this point, the x-space code of G.S. is available from the author upon request, while the Mellin-space program of A.V. has been published in Ref. [63]. The results presented below correspond to a direct iterative solution of the N$^{\rm m}$LO evolution equations for the parton distributions $f_p(x,\mu_{\rm f}^2) \equiv p(x,\mu_{\rm f}^2)$, where $p = q_i,\, \bar{q}_i\,, g$ with $i = 1,\, \ldots, N_{\rm f}$,

$$\frac{d\, f_p(x,\mu_{\rm f}^2)}{d \ln \mu_{\rm f}^2} \;=\; \sum_{l=0}^{m} a_{\rm s}^{l+1}(\mu_{\rm r}^2) \int_x^1 \frac{dy}{y} \sum_{p'} P_{pp'}^{(l)}\left(\frac{x}{y}, \frac{\mu_{\rm f}^2}{\mu_{\rm r}^2}\right) f_{p'}(y,\mu_{\rm f}^2) \tag{4.14}$$

with the strong coupling, normalized as $a_{\rm s} \equiv \alpha_{\rm s}/(4\pi)$, given in terms of

$$\frac{d\, a_{\rm s}}{d \ln \mu_r^2} \;=\; \beta_{\rm N^mLO}(a_{\rm s}) \;=\; -\sum_{l=0}^{m} a_{\rm s}^{l+2}\, \beta_l \tag{4.15}$$

with $\beta_0 = 11 - 2/3\, N_{\rm f}$ etc. $\mu_{\rm r}$ and $\mu_{\rm f}$ represent the renormalization and mass-factorization scales in the $\overline{\rm MS}$ scheme. The reader is referred to Refs. [59, 63] for the scale dependence of the splitting functions $P^{(l)}$ and a further discussion of our solutions of Eqs. (4.14) and (4.15).

For the unpolarized case we retain the initial conditions as set up at the Les Houches meeting: The evolution is started at

$$\mu_{\rm f,0}^2 \;=\; 2\ {\rm GeV}^2 \;. \tag{4.16}$$

Roughly along the lines of the CTEQ5M parametrization [64], the input distributions are chosen as

$$\begin{aligned}
x u_{\rm v}(x,\mu_{\rm f,0}^2) &= 5.107200\, x^{0.8}\, (1-x)^3 \\
x d_{\rm v}(x,\mu_{\rm f,0}^2) &= 3.064320\, x^{0.8}\, (1-x)^4 \\
x g\,(x,\mu_{\rm f,0}^2) &= 1.700000\, x^{-0.1} (1-x)^5 \\
x \bar{d}\,(x,\mu_{\rm f,0}^2) &= .1939875\, x^{-0.1} (1-x)^6 \\
x \bar{u}\,(x,\mu_{\rm f,0}^2) &= (1-x)\, x\bar{d}\,(x,\mu_{\rm f,0}^2) \\
x s\,(x,\mu_{\rm f,0}^2) &= x\bar{s}\,(x,\mu_{\rm f,0}^2) \;=\; 0.2\, x(\bar{u}+\bar{d})(x,\mu_{\rm f,0}^2)
\end{aligned} \tag{4.17}$$

where, as usual, $q_{i,{\rm v}} \equiv q_i - \bar{q}_i$. The running couplings are specified by Eq. (4.15) and

$$\alpha_{\rm s}(\mu_{\rm r}^2 = 2\ {\rm GeV}^2) \;=\; 0.35 \;. \tag{4.18}$$

[4] Contributing authors: G.P. Salam, A. Vogt

For simplicity initial conditions (4.17) and (4.18) are employed regardless of the order of the evolution and the (fixed) ratio of the renormalization and factorization scales.

For the evolution with a fixed number $N_{\rm f} > 3$ of quark flavours the quark distributions not specified in Eq. (4.17) are assumed to vanish at $\mu^2_{\rm f,0}$, and Eq. (4.18) is understood to refer to the chosen value of $N_{\rm f}$. For the evolution with a variable $N_{\rm f} = 3 \ldots 6$, Eqs. (4.16) and (4.17) always refer to three flavours. $N_{\rm f}$ is then increased by one unit at the heavy-quark pole masses taken as

$$m_{\rm c} \,=\, \mu_{\rm f,0}\,, \quad m_{\rm b} \,=\, 4.5\,{\rm GeV}^2\,, \quad m_{\rm t} \,=\, 175\,{\rm GeV}^2\,, \tag{4.19}$$

i.e., Eqs. (4.14) and (4.15) are solved for a fixed number of flavours between these thresholds, and the respective matching conditions are invoked at $\mu^2_{\rm f} = m^2_{\rm h}$, $h = c$, b, t. The matching conditions for the unpolarized parton distributions have been derived at NNLO in Ref. [65], and were first implemented in an evolution program in Ref. [66]. Note that, while the parton distributions are continuous up to NLO due to our choice of the matching scales, $\alpha_{\rm s}$ is discontinuous at these flavour thresholds already at this order for $\mu_{\rm r} \neq \mu_{\rm f}$, see Refs. [67, 68]. Again the reader is referred to Refs. [59, 63] for more details.

Since the exact NNLO splitting functions $P^{(2)}$ are rather lengthy and not directly suitable for use in a Mellin-space program (see, however, Ref. [32]), the reference tables shown below have been computed using the parametrizations (4.22)–(4.24) of Ref. [6] and (4.32)–(4.35) of Ref. [7]. Likewise, the operator matrix element $\widetilde{A}^{\rm S,2}_{\rm hg}$ entering the NNLO flavour matching is taken from Eq. (3.5) of Ref. [63]. The relative error made by using the parametrized splitting functions is illustrated in Fig. 10. It is generally well below 10^{-4}, except for the very small sea quark distributions at very large x.

Eqs. (4.16), (4.18) and (4.19) are used for the (longitudinally) polarized case as well, where Eq. (4.17) replaced by the sufficiently realistic toy input [63]

$$\begin{aligned}
x u_{\rm v} &= +1.3\, x^{0.7}\,(1-x)^3\,(1+3x) \\
x d_{\rm v} &= -0.5\, x^{0.7}\,(1-x)^4\,(1+4x) \\
x g &= +1.5\, x^{0.5}\,(1-x)^5 \\
x \bar{d} &= x \bar{u} \,=\, -0.05\, x^{0.3}\,(1-x)^7 \\
x s &= x \bar{s} \,=\, +0.5\, x \bar{d}\ .
\end{aligned} \tag{4.20}$$

As Eq. (4.17) in the unpolarized case, this input is employed regardless of the order of the evolution.

As in Ref. [59], we have compared the results of our two evolution programs, under the conditions specified above, at 500 x-$\mu^2_{\rm f}$ points covering the range $10^{-8} \leq x \leq 0.9$ and $2\ {\rm GeV}^2 \leq \mu^2_{\rm f} \leq 10^6\ {\rm GeV}^2$. A representative subset of our results at $\mu^2_{\rm f} = 10^4\ {\rm GeV}^4$, a scale relevant to high-E_T jets and close to $m^2_{\rm W}$, $m^2_{\rm Z}$ and, possibly, $m^2_{\rm Higgs}$, is presented in Tables 2 – 6. These results are given in terms of the valence distributions, defined below Eq. (4.17), $L_\pm \equiv \bar{d} \pm \bar{u}$, and the quark–antiquark sums $q_+ \equiv q - \bar{q}$ for $q = s$, c and, for the variable-$N_{\rm f}$ case, b.

For compactness an abbreviated notation is employed throughout the tables, i.e., all numbers $a \cdot 10^b$ are written as a^b. In the vast majority of the x-$\mu^2_{\rm f}$ points our results are found to agree to all five figures displayed, except for the tiny NLO and NNLO sea-quark distributions at $x = 0.9$, in the tables. Entries where the residual offsets between our programs lead to a different fifth digit after rounding are indicated by the subscript '$*$'. In these cases the number with the smaller modulus is given in the tables.

The approximate splitting functions [61], as mentioned above employed in the previous version [59] of our reference tables, have been used in (global) NNLO fits of the unpolarized parton distributions [51, 69], which in turn have been widely employed for obtaining NNLO cross sections, in particular for W and Higgs production. The effect of replacing the approximate results by the full splitting functions [6, 7] is illustrated in Figure 11. Especially at scales relevant to the above-mentioned processes, the previous approximations introduce an error of less than 0.2% for $x \gtrsim 10^{-3}$, and less than 1% even down to $x \simeq 10^{-5}$. Consequently the splitting-function approximations used for the evolution the parton distributions of Refs. [51, 69] are confirmed to a sufficient accuracy for high-scale processes at the LHC.

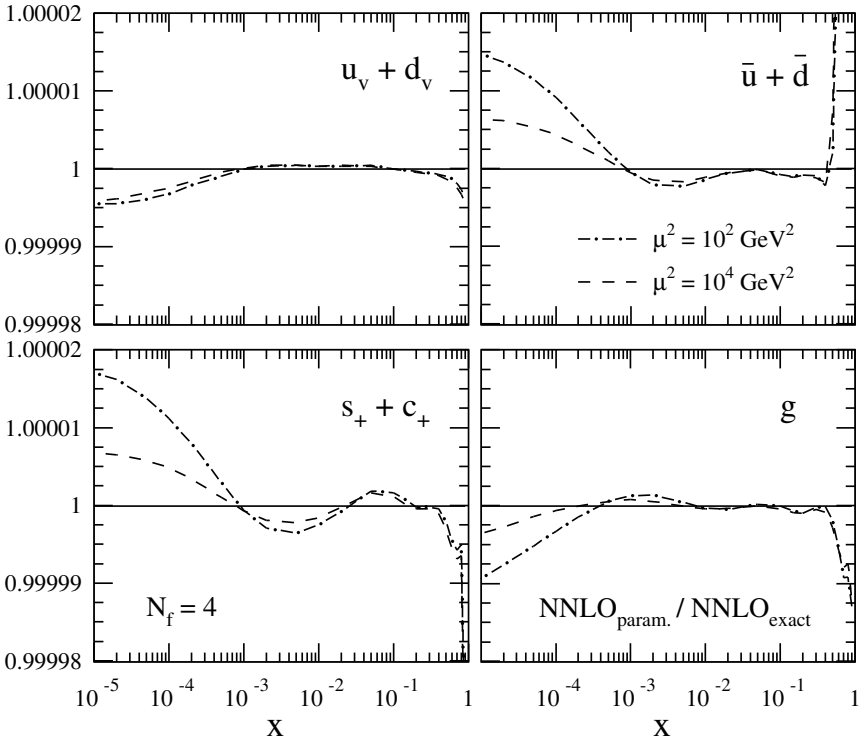

Fig. 10: Relative effects of using the parametrized three-loop splitting functions of Refs. [6,7], instead of the exact expressions from the same source, on the NNLO evolution for the input (4.16)–(4.18) at two representative values of $\mu = \mu_r = \mu_f$.

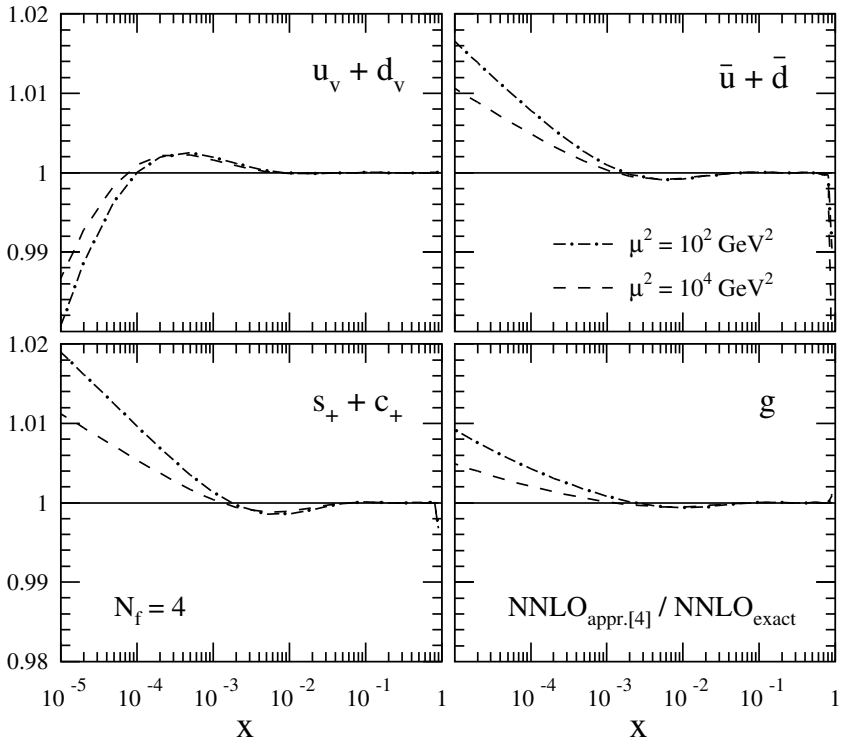

Fig. 11: Relative errors made by using the previous average approximations [61] for the three-loop splitting functions (used, e.g., in Refs. [51,69]) instead of the full results [6,7], on the NNLO evolution of the input (4.16)–(4.18) at $\mu_r = \mu_f$.

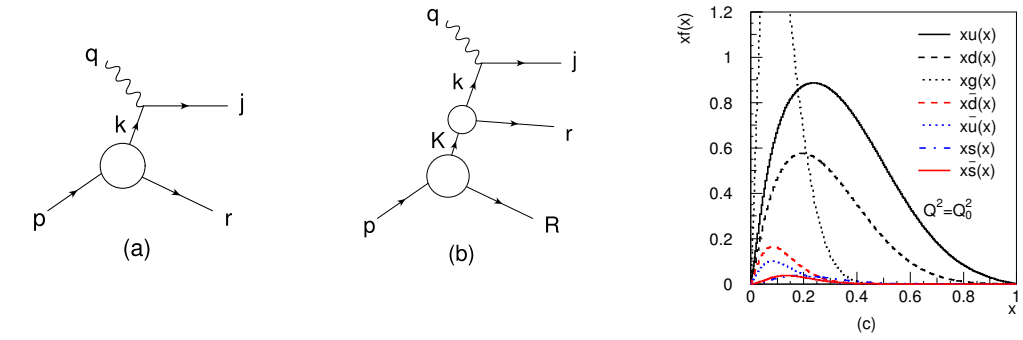

Fig. 12: Probing (a) a valence parton in the proton and (b) a sea parton in a hadronic fluctuation (letters are four-momenta) resulting in (c) parton distributions at the starting scale Q_0^2.

The unchanged unpolarized LO and NLO reference tables of Ref. [59] are not repeated here. Note that the one digit of the first (FFN) α_s value was mistyped in the header of Table 1 in that report [5], the correct value can be found in Table 3 below.

5 Non-perturbative x-shape of PDFs [6]

The x-shape of parton density functions at a low scale Q_0^2 is due to the dynamics of the bound state proton and is hence an unsolved problem of non-perturbative QCD. Usually this is described by parameterizations of data using more or less arbitrary functional forms. More understanding can be obtained by a recently developed physical model [70], which is phenomenologically successful in describing data.

The model gives the four-momentum k of a single probed valence parton (Fig. 12a) by assuming that, in the nucleon rest frame, the shape of the momentum distribution for a parton of type i and mass m_i can be taken as a Gaussian $f_i(k) = N(\sigma_i, m_i) \exp\left\{-\left[(k_0 - m_i)^2 + k_x^2 + k_y^2 + k_z^2\right]/2\sigma_i^2\right\}$, which may be motivated as a result of the many interactions binding the parton in the nucleon. The width of the distribution should be of order hundred MeV from the Heisenberg uncertainty relation applied to the nucleon size, *i.e.* $\sigma_i = 1/d_N$. The momentum fraction x of the parton is then defined as the light-cone fraction $x = k_+/p_+$ and is therefore invariant under longitudinal boosts (e.g. to the infinite momentum frame). Constraints are imposed on the final-state momenta to obtain a kinematically allowed final state, which also ensures that $0 < x < 1$ and $f_i(x) \to 0$ for $x \to 1$.

The sea partons are obtained using a hadronic basis for the non-perturbative dynamics of the bound state proton and considering hadronic fluctuations

$$|p\rangle = \alpha_0|p_0\rangle + \alpha_{p\pi^0}|p\pi^0\rangle + \alpha_{n\pi^+}|n\pi^+\rangle + \ldots + \alpha_{\Lambda K}|\Lambda K^+\rangle + \ldots \tag{5.21}$$

Probing a parton i in a hadron H of a baryon-meson fluctuation $|BM\rangle$ (Fig. 12b) gives a sea parton with light-cone fraction $x = x_H\, x_i$ of the target proton. The momentum of the probed hadron is given by a similar Gaussian, but with a separate width parameter σ_H. Also here, kinematic constraints ensure physically allowed final states.

Using a Monte Carlo method the resulting valence and sea parton x-distributions are obtained without approximations. These apply at a low scale Q_0^2 and the distributions at higher Q^2 are obtained using perturbative QCD evolution at next-to-leading order. To describe all parton distributions (Fig. 12c),

[5] We thank H. Böttcher and J. Blümlein for pointing this out to us.

[6] Contributing author: G. Ingelman

Table 2: Reference results for the $N_{\rm f} = 4$ next-next-to-leading-order evolution for the initial conditions (4.16)–(4.18). The corresponding value of the strong coupling is $\alpha_{\rm s}(\mu_{\rm r}^2 = 10^4\ \text{GeV}^2) = 0.110141$. The valence distributions s_v and c_v are equal for the input (4.17). The notation is explained below Eq. (4.17) and in the paragraph below Eq. (4.20).

NNLO, $N_{\rm f} = 4$, $\mu_{\rm f}^2 = 10^4\ \text{GeV}^2$								
x	xu_v	xd_v	xL_-	$2xL_+$	xs_v	xs_+	xc_+	xg
$\mu_{\rm r}^2 = \mu_{\rm f}^2$								
10^{-7}	1.5287^{-4}	1.0244^{-4}	5.7018^{-6}	1.3190^{+2}	3.1437^{-5}	6.4877^{+1}	6.4161^{+1}	9.9763^{+2}
10^{-6}	6.9176^{-4}	4.4284^{-4}	2.5410^{-5}	6.8499^{+1}	9.4279^{-5}	3.3397^{+1}	3.2828^{+1}	4.9124^{+2}
10^{-5}	3.0981^{-3}	1.8974^{-3}	1.0719^{-4}	3.3471^{+1}	2.2790^{-4}	1.6059^{+1}	1.5607^{+1}	2.2297^{+2}
10^{-4}	1.3722^{-2}	8.1019^{-3}	4.2558^{-4}	1.5204^{+1}	3.6644^{-4}	7.0670^{+0}	6.7097^{+0}	9.0668^{+1}
10^{-3}	5.9160^{-2}	3.4050^{-2}	1.6008^{-3}	6.3230^{+0}	1.4479^{-4}	2.7474^{+0}	2.4704^{+0}	3.1349^{+1}
10^{-2}	2.3078^{-1}	1.2919^{-1}	5.5688^{-3}	2.2752^{+0}	-5.7311^{-4}	8.5502^{-1}	6.6623^{-1}	8.1381^{+0}
0.1	5.5177^{-1}	2.7165^{-1}	1.0023^{-2}	3.9019^{-1}	-3.0627^{-4}	1.1386^{-1}	5.9773^{-2}	9.0563^{-1}
0.3	3.5071^{-1}	1.3025^{-1}	3.0098^{-3}	3.5358^{-2}	-3.1891^{-5}	9.0480^{-3}	3.3061^{-3}	8.4186^{-2}
0.5	1.2117^{-1}	3.1528^{-2}	3.7742^{-4}	2.3867^{-3}	-2.7215^{-6}	5.7965^{-4}	1.7170^{-4}	8.1126^{-3}
0.7	2.0077^{-2}	3.0886^{-3}	1.3434^{-5}	5.4244^{-5}	-1.0106^{-7}	1.2936^{-5}	3.5304^{-6}	3.8948^{-4}
0.9	3.5111^{-4}	1.7783^{-5}	8.651^{-9}	2.695^{-8}	-1.476^{-10}	7.132^{-9}	2.990^{-9}	1.2136^{-6}
$\mu_{\rm r}^2 = 2\,\mu_{\rm f}^2$								
10^{-7}	1.3416^{-4}	8.7497^{-5}	4.9751^{-6}	1.3020^{+2}	2.1524^{-5}	6.4025^{+1}	6.3308^{+1}	1.0210^{+3}
10^{-6}	6.2804^{-4}	3.9406^{-4}	2.2443^{-5}	6.6914^{+1}	6.5149^{-5}	3.2602^{+1}	3.2032^{+1}	4.9626^{+2}
10^{-5}	2.9032^{-3}	1.7575^{-3}	9.6205^{-5}	3.2497^{+1}	1.5858^{-4}	1.5570^{+1}	1.5118^{+1}	2.2307^{+2}
10^{-4}	1.3206^{-2}	7.7673^{-3}	3.9093^{-4}	1.4751^{+1}	2.5665^{-4}	6.8388^{+0}	6.4807^{+0}	9.0162^{+1}
10^{-3}	5.8047^{-2}	3.3434^{-2}	1.5180^{-3}	6.1703^{+0}	1.0388^{-4}	2.6695^{+0}	2.3917^{+0}	3.1114^{+1}
10^{-2}	2.2930^{-1}	1.2857^{-1}	5.4626^{-3}	2.2492^{+0}	-3.9979^{-4}	8.4058^{-1}	6.5087^{-1}	8.0993^{+0}
0.1	5.5428^{-1}	2.7326^{-1}	1.0072^{-2}	3.9297^{-1}	-2.1594^{-4}	1.1439^{-1}	5.9713^{-2}	9.0851^{-1}
0.3	3.5501^{-1}	1.3205^{-1}	3.0557^{-3}	3.6008^{-2}	-2.2632^{-5}	9.2227^{-3}	3.3771^{-3}	8.5022^{-2}
0.5	1.2340^{-1}	3.2166^{-2}	3.8590^{-4}	2.4459^{-3}	-1.9420^{-6}	5.9487^{-4}	1.7699^{-4}	8.2293^{-3}
0.7	2.0597^{-2}	3.1751^{-3}	1.3849^{-5}	5.5722^{-5}	-7.2616^{-8}	1.3244^{-5}	3.5361^{-6}	3.9687^{-4}
0.9	3.6527^{-4}	1.8544^{-5}	9.050^{-9}	2.663^{-8}	-1.075^{-10}	6.713^{-9}	2.377^{-9}	1.2489^{-6}
$\mu_{\rm r}^2 = 1/2\,\mu_{\rm f}^2$								
10^{-7}	1.7912^{-4}	1.2521^{-4}	6.4933^{-6}_{*}	1.2714^{+2}	4.9649^{-5}	6.2498^{+1}	6.1784^{+1}	9.2473^{+2}
10^{-6}	7.7377^{-4}	5.1222^{-4}	2.8719^{-5}	6.7701^{+1}	1.4743^{-4}	3.2999^{+1}	3.2432^{+1}	4.6863^{+2}
10^{-5}	3.3184^{-3}	2.0760^{-3}	1.1977^{-4}	3.3644^{+1}	3.5445^{-4}	1.6147^{+1}	1.5696^{+1}	2.1747^{+2}
10^{-4}	1.4184^{-2}	8.4455^{-3}	4.6630^{-4}	1.5408^{+1}	5.6829^{-4}	7.1705^{+0}	6.8139^{+0}	8.9820^{+1}_{*}
10^{-3}	5.9793^{-2}	3.4418^{-2}	1.6996^{-3}	6.4042^{+0}	2.2278^{-4}	2.7892^{+0}	2.5128^{+0}	3.1336^{+1}
10^{-2}	2.3106^{-1}	1.2914^{-1}	5.7016^{-3}	2.2876^{+0}	-8.9125^{-4}	8.6205^{-1}	6.7377^{-1}	8.1589^{+0}
0.1	5.5039^{-1}	2.7075^{-1}	1.0031^{-2}	3.8850^{-1}	-4.7466^{-4}	1.1332^{-1}	5.9489^{-2}	9.0795^{-1}
0.3	3.4890^{-1}	1.2949^{-1}	2.9943^{-3}	3.5090^{-2}	-4.9304^{-5}	8.9667^{-3}	3.2670^{-3}	8.4309^{-2}
0.5	1.2026^{-1}	3.1269^{-2}	3.7428^{-4}	2.3729^{-3}	-4.1981^{-6}	5.7783^{-4}	1.7390^{-4}	8.1099^{-3}_{*}
0.7	1.9867^{-2}	3.0534^{-3}	1.3273^{-5}	5.4635^{-5}	-1.5541^{-7}	1.3275^{-5}	3.9930^{-6}	3.8824^{-4}
0.9	3.4524^{-4}	1.7466^{-5}	8.489^{-9}	3.030^{-8}	-2.255^{-10}	8.863^{-9}	4.803^{-9}	1.2026^{-6}

Table 3: As Table 2, but for the variable-$N_{\rm f}$ evolution using the flavour matching conditions of Ref. [65, 67, 68]. The corresponding values for the strong coupling $\alpha_{\rm s}(\mu_{\rm r}^2 = 10^4\ {\rm GeV}^2)$ are given by 0.115818, 0.115605 and 0.115410 for $\mu_{\rm r}^2/\mu_{\rm f}^2 = 0.5$, 1 and 2, respectively. For brevity the small, but non-vanishing valence distributions s_v, c_v and b_v are not displayed.

NNLO, $N_{\rm f} = 3 \ldots 5$, $\mu_{\rm f}^2 = 10^4$ GeV^2								
x	xu_v	xd_v	xL_-	$2xL_+$	xs_+	xc_+	xb_+	xg
$\mu_{\rm r}^2 = \mu_{\rm f}^2$								
10^{-7}	1.5978^{-4}	1.0699^{-5}	6.0090^{-6}	1.3916^{+2}	6.8509^{+1}	6.6929^{+1}	5.7438^{+1}	9.9694^{+3}
10^{-6}	7.1787^{-4}	4.5929^{-4}	2.6569^{-5}	7.1710^{+1}	3.5003^{+1}	3.3849^{+1}	2.8332^{+1}	4.8817^{+2}
10^{-5}	3.1907^{-3}	1.9532^{-3}	1.1116^{-4}	3.4732^{+1}	1.6690^{+1}	1.5875^{+1}	1.2896^{+1}	2.2012^{+2}
10^{-4}	1.4023^{-2}	8.2749^{-3}	4.3744^{-4}	1.5617^{+1}	7.2747^{+0}	6.7244^{+0}	5.2597^{+0}	8.8804^{+1}
10^{-3}	6.0019^{-2}	3.4519^{-2}	1.6296^{-3}	6.4173^{+0}	2.7954^{+0}	2.4494^{+0}	1.8139^{+0}	3.0404^{+1}
10^{-2}	2.3244^{-1}	1.3000^{-1}	5.6100^{-3}	2.2778^{+0}	8.5749^{-1}	6.6746^{-1}	4.5073^{-1}	7.7912^{+0}
0.1	5.4993^{-1}	2.7035^{-1}	9.9596^{-3}	3.8526^{-1}	1.1230^{-1}	6.4466^{-2}	3.7280^{-2}	8.5266^{-1}
0.3	3.4622^{-1}	1.2833^{-1}	2.9572^{-3}	3.4600^{-2}	8.8410^{-3}	4.0134^{-3}	2.1047^{-3}	7.8898^{-2}
0.5	1.1868^{-1}	3.0811^{-2}	3.6760^{-4}	2.3198^{-3}	5.6309^{-4}	2.3752^{-4}	1.2004^{-4}	7.6398^{-3}
0.7	1.9486^{-2}	2.9901^{-3}	1.2957^{-5}	5.2352^{-5}	1.2504^{-5}	5.6038^{-6}	2.8888^{-6}	3.7080^{-4}
0.9	3.3522^{-4}	1.6933^{-5}	8.209^{-9}	2.574^{-8}	6.856^{-9}	4.337^{-9}	2.679^{-9}	1.1721^{-6}
$\mu_{\rm r}^2 = 2\,\mu_{\rm f}^2$								
10^{-7}	1.3950^{-4}	9.0954^{-5}	5.2113^{-6}	1.3549^{+2}	6.6672^{+1}	6.5348^{+1}	5.6851^{+1}	1.0084^{+3}
10^{-6}	6.4865^{-4}	4.0691^{-4}	2.3344^{-5}	6.9214^{+1}	3.3753^{+1}	3.2772^{+1}	2.7818^{+1}	4.8816^{+2}
10^{-5}	2.9777^{-3}	1.8020^{-3}	9.9329^{-5}	3.3385^{+1}	1.6015^{+1}	1.5306^{+1}	1.2601^{+1}	2.1838^{+2}
10^{-4}	1.3452^{-2}	7.9078^{-3}	4.0036^{-4}	1.5035^{+1}	6.9818^{+0}	6.4880^{+0}	5.1327^{+0}	8.7550^{+1}
10^{-3}	5.8746^{-2}	3.3815^{-2}	1.5411^{-3}	6.2321^{+0}	2.7012^{+0}	2.3747^{+0}	1.7742^{+0}	3.0060^{+1}
10^{-2}	2.3063^{-1}	1.2923^{-1}	5.4954^{-3}	2.2490^{+0}	8.4141^{-1}	6.5083^{-1}	4.4354^{-1}	7.7495^{+0}
0.1	5.5279^{-1}	2.7222^{-1}	1.0021^{-2}	3.8897^{-1}	1.1312^{-1}	6.2917^{-2}	3.7048^{-2}	8.5897^{-1}
0.3	3.5141^{-1}	1.3051^{-1}	3.0134^{-3}	3.5398^{-2}	9.0559^{-3}	3.8727^{-3}	2.0993^{-3}	8.0226^{-2}
0.5	1.2140^{-1}	3.1590^{-2}	3.7799^{-4}	2.3919^{-3}	5.8148^{-4}	2.2376^{-4}	1.1918^{-4}	7.8098^{-3}
0.7	2.0120^{-2}	3.0955^{-3}	1.3462^{-5}	5.4194^{-5}	1.2896^{-5}	5.0329^{-6}	2.8153^{-6}	3.8099^{-4}
0.9	3.5230^{-4}	1.7849^{-5}	8.687^{-9}	2.568^{-8}	6.513^{-9}	3.390^{-9}	2.407^{-9}	1.2188^{-6}
$\mu_{\rm r}^2 = 1/2\,\mu_{\rm f}^2$								
10^{-7}	1.8906^{-4}	1.3200^{-4}	6.9268^{-6}	1.3739^{+2}	6.7627^{+1}	6.5548^{+1}	5.5295^{+1}	9.4403^{+2}
10^{-6}	8.1001^{-4}	5.3574^{-4}	3.0345^{-5}	7.2374^{+1}	3.5337^{+1}	3.3846^{+1}	2.7870^{+1}	4.7444^{+2}
10^{-5}	3.4428^{-3}	2.1524^{-3}	1.2531^{-4}	3.5529^{+1}	1.7091^{+1}	1.6065^{+1}	1.2883^{+1}	2.1802^{+2}
10^{-4}	1.4580^{-2}	8.6744^{-3}	4.8276^{-4}	1.6042^{+1}	7.4886^{+0}	6.8276^{+0}	5.3044^{+0}	8.9013^{+1}
10^{-3}	6.0912^{-2}	3.5030^{-2}	1.7393^{-3}	6.5544^{+0}	2.8656^{+0}	2.4802^{+0}	1.8362^{+0}	3.0617^{+1}
10^{-2}	2.3327^{-1}	1.3022^{-1}	5.7588^{-3}	2.2949^{+0}	8.6723^{-1}	6.7688^{-1}	4.5597^{-1}	7.8243^{+0}_{*}
0.1	5.4798^{-1}	2.6905^{-1}	9.9470^{-3}	3.8192^{-1}	1.1124^{-1}	6.7091^{-2}	3.7698^{-2}	8.4908^{-1}
0.3	3.4291^{-1}	1.2693^{-1}	2.9239^{-3}	3.4069^{-2}	8.6867^{-3}	4.3924^{-3}	2.1435^{-3}	7.8109^{-2}
0.5	1.1694^{-1}	3.0310^{-2}	3.6112^{-4}	2.2828^{-3}	5.5537^{-4}	2.7744^{-4}	1.2416^{-4}	7.5371^{-3}
0.7	1.9076^{-2}	2.9217^{-3}	1.2635^{-5}	5.2061^{-5}	1.2677^{-5}	7.2083^{-6}	3.0908^{-6}	3.6441^{-4}
0.9	3.2404^{-4}	1.6333^{-5}	7.900^{-9}	2.850^{-8}	8.407^{-9}	6.795^{-9}	3.205^{-9}	1.1411^{-6}

Table 4: Reference results for the $N_{\rm f} = 4$ (FFN) and the variable-$N_{\rm f}$ (VFN) polarized leading-order evolution of the initial distributions (4.20), shown together with these boundary conditions. The respective values for $\alpha_{\rm s}(\mu_{\rm r}^2 = \mu_{\rm f}^2 = 10^4\ \text{GeV}^2)$ read 0.117574 (FFN) and 0.122306 (VFN). The notation is the same as for the unpolarized case.

x	xu_v	$-xd_v$	$-xL_-$	$-2xL_+$	xs_+	xc_+	xb_+	xg
Pol. input, $\mu_{\rm f}^2 = 2\ \text{GeV}^2$								
10^{-7}	1.6366^{-5}	6.2946^{-6}	7.9433^{-5}	1.5887^{-3}	-3.9716^{-4}	0.0^{+0}	0.0^{+0}	4.7434^{-4}
10^{-6}	8.2024^{-5}	3.1548^{-5}	1.5849^{-4}	3.1698^{-3}	-7.9244^{-4}	0.0^{+0}	0.0^{+0}	1.5000^{-3}
10^{-5}	4.1110^{-4}	1.5811^{-4}	3.1621^{-4}	6.3241^{-3}	-1.5810^{-3}	0.0^{+0}	0.0^{+0}	4.7432^{-3}
10^{-4}	2.0604^{-3}	7.9245^{-4}	6.3052^{-4}	1.2610^{-2}	-3.1526^{-3}	0.0^{+0}	0.0^{+0}	1.4993^{-2}
10^{-3}	1.0326^{-2}	3.9716^{-3}	1.2501^{-3}	2.5003^{-2}	-6.2507^{-3}	0.0^{+0}	0.0^{+0}	4.7197^{-2}
10^{-2}	5.1723^{-2}	1.9886^{-2}	2.3412^{-3}	4.6825^{-2}	-1.1706^{-2}	0.0^{+0}	0.0^{+0}	1.4265^{-1}
0.1	2.4582^{-1}	9.1636^{-2}	2.3972^{-3}	4.7943^{-2}	-1.1986^{-2}	0.0^{+0}	0.0^{+0}	2.8009^{-1}
0.3	3.6473^{-1}	1.1370^{-1}	5.7388^{-4}	1.1478^{-2}	-2.8694^{-3}	0.0^{+0}	0.0^{+0}	1.3808^{-1}
0.5	2.5008^{-1}	5.7710^{-2}	6.3457^{-5}	1.2691^{-3}	-3.1729^{-4}	0.0^{+0}	0.0^{+0}	3.3146^{-2}
0.7	8.4769^{-2}	1.1990^{-2}	1.9651^{-6}	3.9301^{-5}	-9.8254^{-6}	0.0^{+0}	0.0^{+0}	3.0496^{-3}
0.9	4.4680^{-3}	2.1365^{-4}	9.689^{-10}	1.9378^{-8}	-4.8444^{-9}	0.0^{+0}	0.0^{+0}	1.4230^{-5}
LO, $N_{\rm f} = 4$, $\mu_{\rm f}^2 = 10^4\ \text{GeV}^2$								
10^{-7}	4.8350_{*}^{-5}	1.8556^{-5}	1.0385^{-4}	3.5124^{-3}	-1.2370^{-3}	-7.1774^{-4}	0.0^{+0}	1.4116^{-2}
10^{-6}	2.3504^{-4}	9.0090^{-5}	2.0700^{-4}	7.7716^{-3}	-2.8508^{-3}	-1.8158^{-3}	0.0^{+0}	4.2163^{-2}
10^{-5}	1.1220^{-3}	4.2916^{-4}	4.1147^{-4}	1.6007^{-2}	-5.9463^{-3}	-3.8889^{-3}	0.0^{+0}	1.0922^{-1}
10^{-4}	5.1990^{-3}	1.9818^{-3}	8.0948^{-4}	2.8757^{-2}	-1.0331^{-2}	-6.2836^{-3}	0.0^{+0}	2.4069^{-1}
10^{-3}	2.2900^{-2}	8.6763^{-3}	1.5309^{-3}	4.0166^{-2}	-1.2428^{-2}	-4.7739^{-3}	0.0^{+0}	4.2181^{-1}
10^{-2}	9.1489^{-2}	3.4200^{-2}	2.4502^{-3}	3.3928^{-2}	-4.7126^{-3}	7.5385^{-3}	0.0^{+0}	4.9485^{-1}
0.1	2.6494^{-1}	9.1898^{-2}	1.5309^{-3}	8.5427^{-3}	3.3830^{-3}	1.1037^{-2}	0.0^{+0}	2.0503^{-1}
0.3	2.2668^{-1}	6.2946^{-2}	2.1104^{-4}	6.6698^{-4}	7.2173^{-4}	1.7769^{-3}	0.0^{+0}	3.3980^{-2}
0.5	9.7647^{-2}	1.9652^{-2}	1.4789^{-5}	-1.8850^{-5}	8.3371^{-5}	1.5732^{-4}	0.0^{+0}	4.3802^{-3}
0.7	1.9545^{-2}	2.3809^{-3}	2.7279^{-7}	-4.1807^{-6}	3.4543^{-6}	4.8183^{-6}	0.0^{+0}	2.6355^{-4}
0.9	4.1768^{-4}	1.7059^{-5}	5.494^{-11}	-7.6712^{-9}	4.1103^{-9}	4.3850^{-9}	0.0^{+0}	9.8421^{-7}
LO, $N_{\rm f} = 3 \ldots 5$, $\mu_{\rm f}^2 = 10^4\ \text{GeV}^2$								
10^{-7}	4.9026^{-5}	1.8815^{-5}	1.0422^{-4}	3.5315^{-3}	-1.2447^{-3}	-7.2356^{-4}	-6.2276^{-4}	1.3726^{-2}
10^{-6}	2.3818^{-4}	9.1286^{-5}	2.0774^{-4}	7.8108^{-3}	-2.8667^{-3}	-1.8280^{-3}	-1.5301^{-3}	4.1011^{-2}
10^{-5}	1.1359^{-3}	4.3445^{-4}	4.1289^{-4}	1.6070^{-2}	-5.9705^{-3}	-3.9060^{-3}	-3.1196^{-3}	1.0615^{-1}
10^{-4}	5.2567^{-3}	2.0035^{-3}	8.1206^{-4}	2.8811^{-2}	-1.0345^{-2}	-6.2849^{-3}	-4.5871^{-3}	2.3343^{-1}
10^{-3}	2.3109^{-2}	8.7537^{-3}	1.5345^{-3}	4.0125^{-2}	-1.2390^{-2}	-4.7174^{-3}	-2.4822^{-3}	4.0743^{-1}
10^{-2}	9.2035^{-2}	3.4391^{-2}	2.4501^{-3}	3.3804^{-2}	-4.6512^{-3}	7.5994^{-3}	6.4665^{-3}	4.7445^{-1}
0.1	2.6478^{-1}	9.1762^{-2}	1.5206^{-3}	8.5181^{-3}	3.3438^{-3}	1.0947^{-2}	6.5223^{-3}	1.9402^{-1}
0.3	2.2495^{-1}	6.2376^{-2}	2.0811^{-4}	6.6195^{-4}	7.0957^{-4}	1.7501^{-3}	9.2045^{-4}	3.1960^{-2}
0.5	9.6318^{-2}	1.9353^{-2}	1.4496^{-5}	-1.8549^{-5}	8.1756^{-5}	1.5424^{-4}	7.8577^{-5}	4.1226^{-3}
0.7	1.9147^{-2}	2.3281^{-3}	2.6556^{-7}	-4.0936^{-6}	3.3746^{-6}	4.7024^{-6}	2.4901^{-6}	2.4888^{-4}
0.9	4.0430^{-4}	1.6480^{-5}	5.285^{-11}	-7.4351^{-9}	3.9818^{-9}	4.2460^{-9}	2.6319^{-9}	9.2939^{-7}

Table 5: Reference results for the polarized next-to-leading-order polarized evolution of the initial distributions (4.20) with $N_{\rm f} = 4$ quark flavours. The corresponding value of the strong coupling is $\alpha_{\rm s}(\mu_{\rm r}^2 = 10^4\,{\rm GeV}^2) = 0.110902$. As in the leading-order case, the valence distributions s_v and c_v vanish for the input (4.20).

Pol. NLO, $N_{\rm f} = 4$, $\mu_{\rm f}^2 = 10^4\,{\rm GeV}^2$							
x	xu_v	xd_v	xL_-	$2xL_+$	xs_+	xc_+	xg
$\mu_{\rm r}^2 = \mu_{\rm f}^2$							
10^{-7}	6.7336^{-5}	-2.5747^{-5}	-1.1434^{-4}	-5.2002^{-3}	-2.0528^{-3}	-1.5034^{-3}	2.6955^{-2}
10^{-6}	3.1280^{-4}	-1.1938^{-4}	-2.3497^{-4}	-1.0725^{-2}	-4.2774^{-3}	-3.1845^{-3}	6.5928^{-2}
10^{-5}	1.4180^{-3}	-5.3982^{-4}	-4.8579^{-4}	-1.9994^{-2}	-7.8594^{-3}	-5.6970^{-3}	1.4414^{-1}
10^{-4}	6.2085^{-3}	-2.3546^{-3}	-9.8473^{-4}	-3.1788^{-2}	-1.1749^{-2}	-7.5376^{-3}	2.7537^{-1}
10^{-3}	2.5741^{-2}	-9.7004^{-3}	-1.8276^{-3}	-3.8222^{-2}	-1.1427^{-2}	-3.6138^{-3}	4.3388^{-1}
10^{-2}	9.6288^{-2}	-3.5778^{-2}	-2.6427^{-3}	-2.6437^{-2}	-1.2328^{-3}	1.0869^{-2}	4.8281^{-1}
0.1	2.5843^{-1}	-8.9093^{-2}	-1.4593^{-3}	-7.5546^{-3}	3.4258^{-3}	1.0639^{-2}	2.0096^{-1}
0.3	2.1248^{-1}	-5.8641^{-2}	-1.9269^{-4}	-1.2210^{-3}	3.5155^{-4}	1.3138^{-3}	3.4126^{-2}
0.5	8.9180^{-2}	-1.7817^{-2}	-1.3125^{-5}	-9.1573^{-5}	1.9823^{-5}	8.5435^{-5}	4.5803^{-3}
0.7	1.7300^{-2}	-2.0885^{-3}	-2.3388^{-7}	-1.9691^{-6}	1.8480^{-7}	1.3541^{-6}	2.9526^{-4}
0.9	3.4726^{-4}	-1.4028^{-5}	-4.407^{-11}	-4.247^{-9}	-1.903^{-9}	-1.683^{-9}	1.2520^{-6}
$\mu_{\rm r}^2 = 2\,\mu_{\rm f}^2$							
10^{-7}	6.1781^{-5}	-2.3641^{-5}	-1.1137^{-4}	-4.6947^{-3}	-1.8092^{-3}	-1.2695^{-3}	2.2530^{-2}
10^{-6}	2.8974^{-4}	-1.1068^{-4}	-2.2755^{-4}	-9.8528^{-3}	-3.8580^{-3}	-2.7838^{-3}	5.7272^{-2}_{*}
10^{-5}	1.3281^{-3}	-5.0612^{-4}	-4.6740^{-4}	-1.8799^{-2}	-7.2908^{-3}	-5.1629^{-3}	1.2975^{-1}
10^{-4}	5.8891^{-3}	-2.2361^{-3}	-9.4412^{-4}	-3.0787^{-2}	-1.1292^{-2}	-7.1363^{-3}	2.5644^{-1}
10^{-3}	2.4777^{-2}	-9.3502^{-3}	-1.7632^{-3}	-3.8610^{-2}	-1.1658^{-2}	-3.9083^{-3}	4.1725^{-1}
10^{-2}	9.4371^{-2}	-3.5129^{-2}	-2.6087^{-3}	-2.8767^{-2}	-2.3430^{-3}_{*}	9.7922^{-3}_{*}	4.7804^{-1}
0.1	2.6008^{-1}	-8.9915^{-2}	-1.4923^{-3}	-8.3806^{-3}	3.1932^{-3}	1.0585^{-2}	2.0495^{-1}
0.3	2.1837^{-1}	-6.0497^{-2}	-2.0143^{-4}	-1.2157^{-3}	3.9810^{-4}	1.4042^{-3}	3.5366^{-2}
0.5	9.3169^{-2}	-1.8699^{-2}	-1.3954^{-5}	-7.9331^{-5}	3.0091^{-5}	9.9849^{-5}	4.7690^{-3}
0.7	1.8423^{-2}	-2.2357^{-3}	-2.5360^{-7}	-1.0062^{-6}	7.6483^{-7}	2.0328^{-6}	3.0796^{-4}
0.9	3.8293^{-4}	-1.5559^{-5}	-4.952^{-11}	-1.955^{-9}	-7.298^{-10}	-4.822^{-10}	1.3247^{-6}
$\mu_{\rm r}^2 = 1/2\,\mu_{\rm f}^2$							
10^{-7}	7.4443^{-5}	-2.8435^{-5}	-1.1815^{-4}	-5.7829^{-3}	-2.3341^{-3}	-1.7739^{-3}	3.2071^{-2}
10^{-6}	3.4143^{-4}	-1.3016^{-4}	-2.4482^{-4}	-1.1668^{-2}	-4.7305^{-3}	-3.6168^{-3}	7.5123^{-2}
10^{-5}	1.5256^{-3}	-5.8002^{-4}	-5.1085^{-4}	-2.1193^{-2}	-8.4295^{-3}	-6.2295^{-3}_{*}	1.5788^{-1}
10^{-4}	6.5726^{-3}	-2.4891^{-3}	-1.0409^{-3}	-3.2697^{-2}	-1.2166^{-2}	-7.8952^{-3}	2.9079^{-1}
10^{-3}	2.6766^{-2}	-1.0070^{-2}	-1.9171^{-3}	-3.7730^{-2}	-1.1160^{-2}	-3.2890^{-3}_{*}	4.4380^{-1}
10^{-2}	9.8073^{-2}	-3.6370^{-2}	-2.6942^{-3}	-2.4056^{-2}	-1.2354^{-4}_{*}	1.1929^{-2}	4.8272^{-1}
0.1	2.5628^{-1}	-8.8133^{-2}	-1.4304^{-3}	-6.9572^{-3}	3.5561^{-3}	1.0604^{-2}	1.9831^{-1}
0.3	2.0709^{-1}	-5.6988^{-2}	-1.8541^{-4}	-1.3308^{-3}	2.5993^{-4}	1.1855^{-3}	3.3524^{-2}
0.5	8.5835^{-2}	-1.7089^{-2}	-1.2463^{-5}	-1.1920^{-4}	2.6972^{-6}_{*}	6.4995^{-5}	4.5044^{-3}
0.7	1.6405^{-2}	-1.9723^{-3}	-2.1859^{-7}_{*}	-3.6817^{-6}	-7.4795^{-7}_{*}	3.4496^{-7}	2.9100^{-4}
0.9	3.2011^{-4}	-1.2870^{-5}	-4.000^{-11}	-8.173^{-9}	-3.886^{-9}	-3.686^{-9}	1.2230^{-6}

Table 6: As Table 5, but for the variable-$N_{\rm f}$ evolution using Eqs. (4.16), (4.17) and (4.20). The corresponding values for the strong coupling $\alpha_{\rm s}(\mu_{\rm r}^2 = 10^4\ {\rm GeV}^2)$ are given by 0.116461, 0.116032 and 0.115663 for $\mu_{\rm r}^2/\mu_{\rm f}^2 = 0.5$, 1 and 2, respectively.

Pol. NLO, $N_{\rm f} = 3 \ldots 5$, $\mu_{\rm f}^2 = 10^4$ GeV^2								
x	xu_v	$-xd_v$	$-xL_-$	$-2xL_+$	xs_+	xc_+	xb_+	xg
$\mu_{\rm r}^2 = \mu_{\rm f}^2$								
10^{-7}	6.8787^{-5}_{*}	2.6297^{-5}	1.1496^{-4}	5.2176^{-3}	-2.0592^{-3}	-1.5076^{-3}	-1.2411^{-3}	2.5681^{-2}
10^{-6}	3.1881^{-4}	1.2165^{-4}	2.3638^{-4}	1.0770^{-2}	-4.2953^{-3}	-3.1979^{-3}	-2.4951^{-3}	6.3021^{-2}
10^{-5}	1.4413^{-3}	5.4856^{-4}	4.8893^{-4}	2.0077^{-2}	-7.8934^{-3}	-5.7228^{-3}	-4.1488^{-3}	1.3809^{-1}
10^{-4}	6.2902^{-3}	2.3849^{-3}	9.9100^{-4}	3.1883^{-2}	-1.1785^{-2}	-7.5596^{-3}	-4.8420^{-3}	2.6411^{-1}
10^{-3}	2.5980^{-2}	9.7872^{-3}	1.8364^{-3}	3.8224^{-2}	-1.1416^{-2}	-3.5879^{-3}	-1.1723^{-3}	4.1601^{-1}
10^{-2}	9.6750^{-2}	3.5935^{-2}	2.6452^{-3}	2.6306^{-2}	-1.1774^{-3}	1.0917^{-2}	8.1196^{-3}	4.6178^{-1}
0.1	2.5807^{-1}	8.8905^{-2}	1.4509^{-3}	7.4778^{-3}	3.4207^{-3}	1.0591^{-2}	6.1480^{-3}	1.9143^{-1}
0.3	2.1104^{-1}	5.8186^{-2}	1.9054^{-4}	1.2026^{-3}	3.4999^{-4}	1.3015^{-3}	7.2795^{-4}	3.2621^{-2}
0.5	8.8199^{-2}	1.7601^{-2}	1.2924^{-5}	8.9668^{-5}	1.9771^{-5}	8.4378^{-5}	5.2125^{-5}	4.4207^{-3}
0.7	1.7027^{-2}	2.0531^{-3}	2.2921^{-7}	1.9243^{-6}	1.8384^{-7}	1.3298^{-6}	1.2157^{-6}	2.8887^{-4}
0.9	3.3898^{-4}	1.3676^{-5}	4.284^{-11}	4.260^{-9}	-1.916^{-9}	-1.701^{-9}	-7.492^{-11}	1.2435^{-6}
$\mu_{\rm r}^2 = 2\,\mu_{\rm f}^2$								
10^{-7}	6.2819^{-5}_{*}	2.4035^{-5}	1.1180^{-4}	4.6896^{-3}	-1.8050^{-3}	-1.2637^{-3}	-1.0544^{-3}	2.1305^{-2}
10^{-6}	2.9408^{-4}	1.1232^{-4}	2.2855^{-4}	9.8538^{-3}	-3.8554^{-3}	-2.7780^{-3}	-2.2077^{-3}	5.4411^{-2}
10^{-5}	1.3450^{-3}	5.1245^{-4}	4.6965^{-4}	1.8815^{-2}	-7.2936^{-3}	-5.1597^{-3}	-3.8359^{-3}	1.2368^{-1}
10^{-4}	5.9485^{-3}	2.2582^{-3}	9.4866^{-4}	3.0816^{-2}	-1.1297^{-2}	-7.1323^{-3}	-4.7404^{-3}	2.4503^{-1}
10^{-3}	2.4951^{-2}	9.4134^{-3}	1.7698^{-3}	3.8618^{-2}	-1.1654^{-2}	-3.8925^{-3}	-1.5608^{-3}	3.9912^{-1}
10^{-2}	9.4706^{-2}	3.5243^{-2}	2.6108^{-3}	2.8761^{-2}	-2.3471^{-3}	9.7827^{-3}	7.5188^{-3}	4.5698^{-1}
0.1	2.5982^{-1}	8.9780^{-2}	1.4862^{-3}	8.3807^{-3}	3.1615^{-3}	1.0522^{-2}	6.1973^{-3}	1.9561^{-1}
0.3	2.1732^{-1}	6.0165^{-2}	1.9984^{-4}	1.2086^{-3}	3.9371^{-4}	1.3919^{-3}	7.6929^{-4}	3.3906^{-2}
0.5	9.2445^{-2}	1.8539^{-2}	1.3804^{-5}	7.8411^{-5}	2.9799^{-5}	9.8805^{-5}	5.7333^{-5}	4.6166^{-3}
0.7	1.8219^{-2}	2.2090^{-3}	2.5004^{-7}_{*}	9.8927^{-7}_{*}	7.5552^{-7}	2.0057^{-6}	1.4438^{-6}	3.0231^{-4}
0.9	3.7653^{-4}	1.5285^{-5}	4.855^{-11}	2.005^{-9}	-7.599^{-10}	-5.171^{-10}	3.809^{-10}	1.3232^{-6}
$\mu_{\rm r}^2 = 1/2\,\mu_{\rm f}^2$								
10^{-7}	7.6699^{-5}	2.9289^{-5}	1.1912^{-4}	5.8548^{-3}	-2.3667^{-3}	-1.8030^{-3}	-1.4521^{-3}	3.1009^{-2}
10^{-6}	3.5067^{-4}	1.3364^{-4}	2.4707^{-4}	1.1806^{-2}	-4.7934^{-3}	-3.6731^{-3}	-2.7846^{-3}	7.2690^{-2}
10^{-5}	1.5611^{-3}	5.9329^{-4}	5.1593^{-4}	2.1406^{-2}	-8.5248^{-3}	-6.3125^{-3}	-4.4072^{-3}	1.5274^{-1}
10^{-4}	6.6957^{-3}	2.5346^{-3}	1.0509^{-3}	3.2903^{-2}	-1.2252^{-2}	-7.9608^{-3}	-4.8402^{-3}	2.8097^{-1}
10^{-3}	2.7125^{-2}	1.0200^{-2}	1.9310^{-3}	3.7698^{-2}	-1.1127^{-2}	-3.2334^{-3}	-7.5827^{-4}	4.2756^{-1}
10^{-2}	9.8758^{-2}	3.6602^{-2}	2.6980^{-3}	2.3675^{-2}	5.1386^{-5}	1.2092^{-2}	8.6053^{-3}	4.6241^{-1}
0.1	2.5572^{-1}	8.7847^{-2}	1.4179^{-3}	6.7523^{-3}	3.5944^{-3}	1.0578^{-2}	6.0904^{-3}	1.8838^{-1}
0.3	2.0497^{-1}	5.6318^{-2}	1.8228^{-4}	1.2965^{-3}	2.6142^{-4}	1.1713^{-3}	6.8941^{-4}	3.1884^{-2}
0.5	8.4404^{-2}	1.6775^{-2}	1.2174^{-5}	1.1604^{-4}	2.8309^{-6}	6.3682^{-5}	4.7009^{-5}	4.3221^{-3}
0.7	1.6013^{-2}	1.9215^{-3}	2.1196^{-7}_{*}	3.6047^{-6}	-7.4260^{-7}	3.1714^{-7}	9.6419^{-7}	2.8268^{-4}
0.9	3.0848^{-4}	1.2377^{-5}	3.829^{-11}	8.129^{-9}	-3.873^{-9}	-3.681^{-9}	-6.816^{-10}	1.2009^{-6}

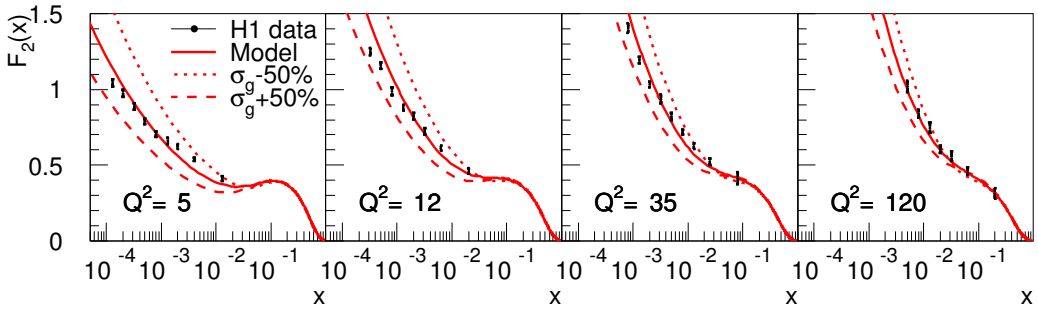

Fig. 13: $F_2(x, Q^2)$ from H1 compared to the model with $\pm 50\%$ variation of the width parameter σ_g of the gluon distribution.

the model has only four shape parameters and three normalization parameters, plus the starting scale:

$$\begin{array}{cccc} \sigma_u = 230 \text{ MeV} & \sigma_d = 170 \text{ MeV} & \sigma_g = 77 \text{ MeV} & \sigma_H = 100 \text{ MeV} \\ \alpha^2_{\text{p}\pi^0} = 0.45 & \alpha^2_{n\pi^+} = 0.14 & \alpha^2_{\Lambda K} = 0.05 & Q_0 = 0.75 \text{ GeV} \end{array} \tag{5.22}$$

These are determined from fits to data as detailed in [70] and illustrated in Fig. 13. The model reproduces the inclusive proton structure function and gives a natural explanation of observed quark asymmetries, such as the difference between the up and down valence distributions and between the anti-up and anti-down sea quark distributions. Moreover, its asymmetry in the momentum distribution of strange and anti-strange quarks in the nucleon is large enough to reduce the NuTeV anomaly to a level which does not give a significant indication of physics beyong the Standard Model.

Recent fits of PDF's at very low x and Q^2 have revealed problems with the gluon density, which in some cases even becomes negative. The reason for this is that the DGLAP evolution, driven primarily by the gluon at small x, otherwise gives too large parton densities and thereby a poor fit to F_2 in the genuine DIS region at larger Q^2. It has been argued [71] that the root of the problem is the application of the formalism for DIS also in the low-Q^2 region, where the momentum transfer is not large enough that the parton structure of the proton is clearly resolved. The smallest distance that can be resolved is basically given by the momentum transfer of the exchanged photon through $d = 0.2/\sqrt{Q^2}$, where d is in Fermi if Q^2 is in GeV^2. This indicates that partons are resolved only for $Q^2 \gtrsim 1\,\text{GeV}^2$. For $Q^2 \lesssim 1\,\text{GeV}^2$, there is no hard scale involved and a parton basis for the description is not justified. Instead, the interaction is here of a soft kind between the nearly on-shell photon and the proton. The cross section is then dominated by the process where the photon fluctuates into a virtual vector meson state which then interacts with the proton in a strong interaction. The quantum state of the photon can be expressed as $|\gamma\rangle = C_0|\gamma_0\rangle + \sum_V \frac{e}{f_V}|V\rangle + \int_{m_0} dm(\cdots)$. The sum is over $V = \rho^0, \omega, \phi \ldots$ as in the original vector meson dominance model (VDM), whereas the generalised vector meson dominance model (GVDM) also includes the integral over a continuous mass spectrum (not written out explicitly here).

Applied to ep at low Q^2 this leads to the expression [71]

$$\begin{aligned} F_2(x,Q^2) \;=\; & \frac{(1-x)Q^2}{4\pi^2\alpha}\left\{\sum_{V=\rho,\omega,\phi} r_V \left(\frac{m_V^2}{Q^2+m_V^2}\right)^2 \left(1+\xi_V\frac{Q^2}{m_V^2}\right)\right. \\ & \left. + r_C\left[(1-\xi_C)\frac{m_0^2}{Q^2+m_0^2} + \xi_C\frac{m_0^2}{Q^2}\ln\left(1+\frac{Q^2}{m_0^2}\right)\right]\right\} A_\gamma \frac{Q^{2\epsilon}}{x^\epsilon} \end{aligned} \tag{5.23}$$

where the hadronic cross-section $\sigma(ip \to X) = A_i s^\epsilon + B_i s^{-\eta} \approx A_i s^\epsilon \approx A_i (Q^2/x)^\epsilon$ has been used for the small-x region of interest. The parameters involved are all essentially known from GVDM phe-

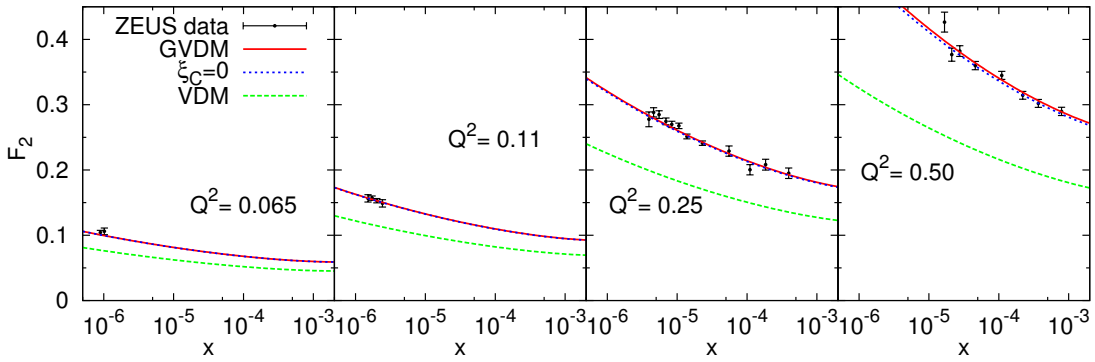

Fig. 14: F_2 data at low Q^2 from ZEUS compared to the full GVDM in eq. (5.23) (full curves), when excluding the longitudinal contribution of the continuum ($\xi_C = 0$) and excluding the continuous contribution altogether (setting $r_C = 0$) giving VDM.

nomenology. With $\epsilon = 0.091$, $\xi = 0.34$, $m_0 = 1.5$ GeV and $A_\gamma = 71\,\mu$b, this GVDM model gives a good fit ($\chi^2/\text{d.o.f.} = 87/66 = 1.3$) as illustrated in Fig. 14. Using this model at very low Q^2 in combination with the normal parton density approach at larger Q^2 it is possible to obtain a good description of data over the full Q^2 range [71]. This involves, however, a phenomenological matching of these two approaches, since a theoretically well justified combination is an unsolved problem.

Neglecting the GVDM component when fitting PDF's to data at small Q^2 may thus lead to an improper gluon distribution, which is not fully universal and therefore may give incorrect results when used for cross section calculations at LHC.

6 Towards precise determination of the nucleon PDFs [7]

The nucleon parton distribution functions (PDFs) available to the moment are extracted from the rather limited set of experimental distributions (the deep-inelastic scattering (DIS) structure functions, the Drell-Yan (DY) and jet production cross sections). Other high-energy processes potentially could provide additional constraints on PDFs, however insufficient theoretical understanding does not allow to use those data without risk of having uncontrolled theoretical inaccuracies. Even for the case of the existing global fits of the PDFs performed by the MRST and CTEQ groups missing next-to-next-to-leading (NNLO) order QCD corrections to the Drell-Yan and jet production cross sections are not small as compared to the accuracy of the corresponding data used and therefore might give non-negligible effect. In this section we outline progress in the QCD fits with consistent account of the NNLO corrections.

6.1 Impact of the NNLO evolution kernel fixation on PDFs

In order to allow account of the NNLO corrections in the fit of PDFs one needs analytical expressions for the 3-loop corrections to the QCD evolution kernel. Until recent times these expressions were known only in the approximate form of Ref. [61] derived from the partial information about the kernel, including the set of its Mellin moments and the low-x asymptotics [12,22,23] However with the refined calculations of Ref. [6, 7] the exact expression for the NNLO kernel has been available. These improvement is of particular importance for analysis of the low-x data including the HERA ones due to general rise of the high-order QCD correction at low x. We illustrate impact of the NNLO evolution kernel validation on PDFs using the case of fit to the global DIS data [72–77]. The exact NNLO corrections to the DIS coefficient functions are know [4, 78] that allowed to perform approximate NNLO fit of PDFs to these data [69] using the approximate NNLO corrections to the evolution kernel of Ref. [61]. Taking into

[7] Contributing author: S. I. Alekhin

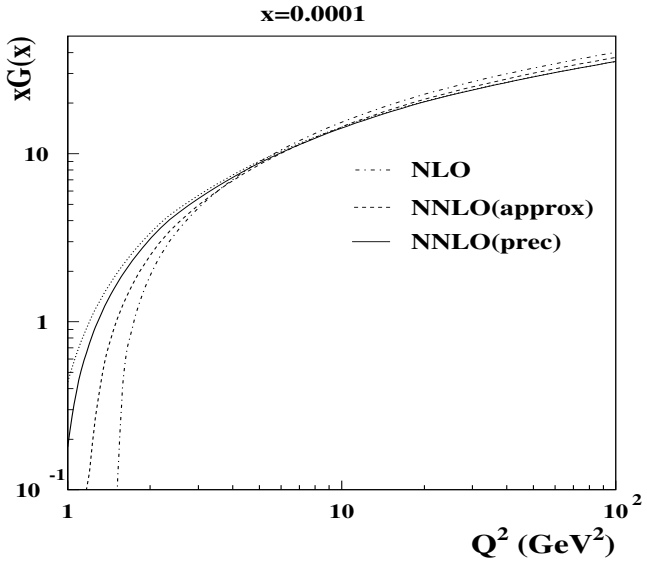

Fig. 15: The gluon distributions obtained in the different variants of PDFs fit to the DIS data (solid: the fit with exact NNLO evolution; dashes: the fit with approximate NNLO evolution; dots: the approximate NNLO gluons evolved with the exact NNLO kernel; dashed-dots: the NLO fit).

account exact NNLO evolution kernel the analysis of Ref. [69] was updated recently to the exact NNLO case [79].

The gluon distributions at small x obtained in these two variants of the fit are compared in Fig.15. With the exact NNLO corrections the QCD evolution of gluon distribution at small x gets weaker and as a result at small x/Q the gluon distribution obtained using the precise NNLO kernel is quite different from the approximate one. In particular, the approximate NNLO gluon distribution is negative at $Q^2 \lesssim 1.3$ GeV^2, while the precise one remains positive even below $Q^2 = 1$ GeV^2. For the NLO case the positivity of gluons at small x/Q is even worse than for the approximate NNLO case due to the approximate NNLO corrections dampen the gluon evolution at small x too, therefore account of the NNLO corrections is crucial in this respect. (cf. discussion of Ref. [80]). Positivity of the PDFs is not mandatory beyond the QCD leading order, however it allows probabilistic interpretation of the parton model and facilitates modeling of the soft processes, such as underlying events in the hadron-hadron collisions at LHC. The change of gluon distribution at small x/Q as compared to the fit with approximate NNLO evolution is rather due the change in evolution kernel than due to shift in the fitted parameters of PDFs. This is clear from comparison of the exact NNLO gluon distribution to one obtained from the approximate NNLO fit and evolved to low Q using the exact NNLO kernel (see Fig.15). In the vicinity of crossover in the gluon distribution to the negative values its relative change due to variation of the evolution kernel is quite big and therefore further fixation of the kernel at small x discussed in Ref. [81] might be substantial for validation of the PDFs at low x/Q. For the higher-mass kinematics at LHC numerical impact of the NNLO kernel update is not dramatic. Change in the Higgs and W/Z bosons production cross sections due to more precise definition of the NNLO PDFs is comparable to the errors coming from the PDFs uncertainties, i.e. at the level of several percent.

6.2 NNLO fit of PDFs to the combined DIS and Drell-Yan data

The DIS process provide very clean source of information about PDFs both from experimental and theoretical side, however very poorly constrains the gluon and sea distributions at $x \gtrsim 0.3$. The well known way to improve precision of the sea distributions is to combine DIS data with the Drell-Yan ones.

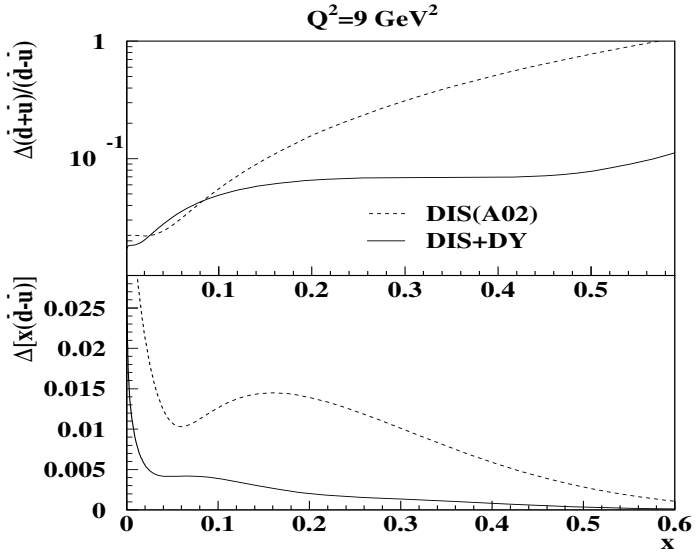

Fig. 16: Uncertainties in the non-strange sea distributions obtained from NNLO QCD fit to the DIS data combined with the fixed target Drell-Yan data (solid curves). The same uncertainties obtained in fit to the DIS data only [8] are given for comparison by dashes.

The cross section of process $NN \to l^+l^-$ reads

$$\sigma_{DY} \propto \sum_i [q_i(x_1)\bar{q}_i(x_2) + q_i(x_2)\bar{q}_i(x_1)] + \text{higher-order terms},$$

where $q(\bar{q})_i$ are the quarks(antiquarks) distribution and $x_{1,2}$ give the momentum fractions carried by each of the colliding partons. The quark distributions are determined by the DIS data with the precision of several percent in the wide region of x and therefore precision of the sea distribution extracted from the combined fit to the DIS and DY data is basically determined by the latter. The Fermilab fixed-target experiments provide measurements of the DY cross sections for the isoscalar target [82] and the ratio of cross sections for the deuteron and proton targets [83] with the accuracy better than 20% at $x \lesssim 0.6$. Fitting PDFs to these data combined with the global DIS data of Ref. [72–77] we can achieve comparable precision in the sea distributions. Recent calculations of Ref. [84] allow to perform this fit with full account of the NNLO correction. Using these calculations the DY data of Refs. [82, 83] were included into the NNLO fit of Ref. [79] that leads to significant improvement in the precision of sea distributions (see Fig. 16). Due to the DY data on the deuteron/proton ratio the isospin asymmetry of sea is also improved. It is worth to note that the precision achieved for the total sea distribution is in good agreement to the rough estimates given above. The value of χ^2/NDP obtained in the fit is 1.1 and the spread of χ^2/NDP over separate experiments used in the fit is not dramatic, its biggest value is 1.4. We rescaled the errors in data for experiments with $\chi^2/\text{NDP} > 1$ in order to bring χ^2/NDP for this experiments to 1 and found that overall impact of this rescaling on the PDFs errors is marginal. This proofs sufficient statistical consistency of the data sets used in the fit and disfavors huge increase in the value of $\Delta\chi^2$ criterion suggested by the CTEQ collaboration for estimation of errors in the global fit of PDFs. A particular feature of the PDFs obtained is good stability with respect to the choice of factorization/renormalization scale in the DY cross section: Variation of this scale from $M_{\mu^+\mu^-}/2$ to $2M_{\mu^+\mu^-}$ leads to variation of PDFs comparable to their uncertainties due to errors in data.

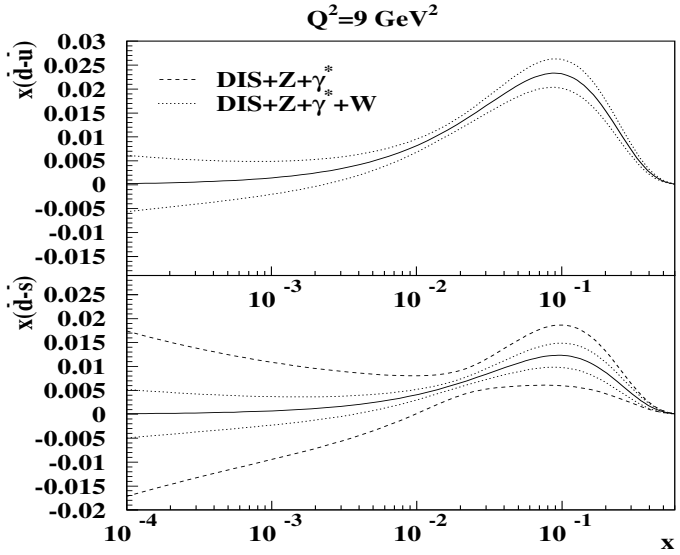

Fig. 17: The 1σ error band for $x(\bar{d}-\bar{u})$ (upper panel) and $x(\bar{d}-\bar{s})$ (lower panel) expected for the fit of PDFs to the LHC data combined with the global DIS ones. Dashed curves correspond to the case of $Z/\gamma*$-production, dots are for the combination $Z/\gamma*$- with the W^+/W^--production. Solid curves are for the central values obtained from the reference fit to the global DIS data

.

6.3 LHC data and flavor separation of the sea at small x

Combination of the existing DIS and fixed-target DY data provide good constraint on the total sea quarks distribution and allows separation of the $\bar{u}$- and $\bar{d}$-quark distributions up to the values of x sufficient for most practical applications at the LHC. At small x the total sea is also well constrained by the precise HERA data on the inclusive structure functions, however $\bar{u}/\bar{d}$ separation is poor in this region due to lack of the deuteron target data at HERA. The problem of the sea flavor separation is regularly masked due to additional constraints imposed on PDFs. In particular, most often the Regge-like behavior of the sea isospin asymmetry $x(\bar{d}-\bar{u}) \propto x^{a_{ud}}$ is assumed with a_{ud} selected around value of 0.5 motivated by the intercept of the meson trajectories. This assumption automatically provides constraint $\bar{d}=\bar{u}$ at $x \to 0$ and therefore leads to suppression of the uncertainties both in $\bar{u}$ and $\bar{d}$ at small x. If we do not assume the Regge-like behavior of $x(\bar{d}-\bar{u})$ its precision determined from the NNLO fit to the combined DIS and DY data of Section 1.2 is about 0.04 at $x = 10^{-4}$ furthermore this constraint is defined rather by assumption about the shape of PDFs at small x than by data used in the fit. The strange sea distribution is known much worse than the non-strange ones. It is essentially defined only by the CCFR experiment from the cross section of dimuon production in the neutrino nucleus collisions [85]. In this experiment the strange sea distribution was probed at $x = 0.01 \div 0.2$ and the shape obtained is similar to one of the non-strange sea with the strangeness suppression factor about 0.5. This is in clear disagreement with the Regge-like constraint on $x(\bar{d}-\bar{s})$ or $x(\bar{u}-\bar{s})$ and therefore we cannot use even this assumption to predict the strange sea at small x.

The LHC data on $\mu^+\mu^-$ production cross section can be used for further validation of the sea distributions at small x. Study of this process at the lepton pair masses down to 15 GeV will allow to probe PDFs at x down to 10^{-4}, while with both leptons detected full kinematics can be reliably reconstructed. In order to check impact of the foreseen LHC data on the sea flavor separation we generated sample of pseudo-data for the process $pp \to \mu^+\mu^- X$ at $\sqrt{s} = 14$ TeV with integral luminosity of 10 1/fb corresponding to the first stage of the LHC operation. In order to meet typical limitations of the LHC

Table 7: Values of the parameters obtained in the benchmark fit.

Valence	a_u	0.718 ± 0.085
	b_u	3.81 ± 0.16
	ϵ_u	-1.56 ± 0.46
	γ_u	3.30 ± 0.49
	a_d	1.71 ± 0.20
	b_d	10.00 ± 0.97
	ϵ_d	-3.83 ± 0.23
	γ_d	4.64 ± 0.41
Sea	A_S	0.211 ± 0.016
	a_s	-0.048 ± 0.039
	b_s	2.20 ± 0.20
Glue	a_G	0.356 ± 0.095
	b_G	10.9 ± 1.4
	$\alpha_{\rm s}(M_{\rm Z})$	0.1132 ± 0.0015

detectors only events with the lepton pair absolute rapidity less than 2.5 were accepted; other detector effects were not taken into account. For generation of these pseudo-data we used PDFs obtained in the dedicated version of fit [79] with the sea distributions parameterized as $xS_{u,d,s} = \eta_{u,d,s}x^a(1-x)^{b_{u,d,s}}$ with the constraints $\eta_u = \eta_d = \eta_s$ and $b_s = (b_u + b_d)/2$ imposed. These constraints are necessary for stability of the fit in view of limited impact of the DIS data on the flavor separation and, besides, the former one guarantees SU(3) symmetry in the sea distributions at small x. The generated pseudo-data were added to the basic DIS data sample and the errors in PDFs parameters were re-estimated with no constraints on the sea distributions imposed at this stage. Since dimuon data give extra information about the PDFs products they allow to disentangle the strange distribution, if an additional constraint on the non-strange sea distributions is set. The dashed curves in the lower panel of Fig.17 give the 1σ bands for $x(\bar{d}-\bar{s})$ as they are defined by the LHC simulated data combined with the global DIS ones given $(\bar{d}-\bar{u})$ is fixed. One can see that $\bar{d}/\bar{s}$ (and $\bar{u}/\bar{s}$) separation at the level of several percents would be feasible down to x=10^{-4} in this case. The supplementary constraint on $(\bar{d}-\bar{u})$ can be obtained from study of the W-boson charge asymmetry. To estimate impact of this process we simulated the single W^+- and W^--production data similarly to the case of the $\mu^+\mu^-$-production and took into account this sample too. In this case one can achieve separation of all three flavors with the precision better than 0.01 (see Fig.17). Note that strange sea separation is also improved due to certain sensitivity of the W-production cross section to the strange sea contribution. The estimates obtained refer to the ideal case of full kinematical reconstruction of the W-bosons events. For the case of using the charge asymmetry of muons produced from the W-decays the precision of the PDFs would be worse. Account of the backgrounds and the detector effects would also deteriorate it, however these losses can be at least partially compensated by rise of the LHC luminosity at the second stage of operation.

6.4 Benchmarking of the PDFs fit

For the available nucleon PDFs the accuracy at percent level is reached in some kinematical regions. For this reason benchmarking of the codes used in these PDFs fits is becoming important issue. A tool for calibration of the QCD evolution codes was provided by Les Houches workshop [59]. To allow benchmarking of the PDFs errors calculation we performed a test fit suggested in Les Houches workshop too. This fit reproduces basic features of the existing global fits of PDFs, but is simplified a lot to facilitate its reproduction. We use for the analysis data on the proton DIS structure functions F_2 obtained by the

BCDMS, NM, H1, and ZEUS collaborations and ratio of the deuteron and proton structure functions F_2 obtained by the NMC. The data tables with full description of experimental errors taken into account are available online[8]. Cuts for the momentum transferred $Q^2 > 9$ GeV^2 and for invariant mass of the hadronic system $W^2 > 15$ GeV^2 are imposed in order to avoid influence of the power corrections and simplify calculations. The contribution of the Z-boson exchange at large Q is not taken into account for the same purpose. The PDFs are parameterized in the form

$$xp_i(x, 1\ GeV) = N_i x^{a_i}(1-x)^{b_i}(1+\epsilon_i\sqrt{x}+\gamma_i x),$$

to meet choice common for many popular global fits of PDFs. Some of the parameters ϵ_i and γ_i are set to zero since they were found to be consistent to zero within the errors. We assume isotopic symmetry for sea distribution and the strange sea is the same as the non-strange ones suppressed by factor of 0.5. Evolution of the PDFs is performed in the NLO QCD approximation within the $\overline{\rm MS}$ scheme. The heavy quarks contribution is accounted in the massless scheme with the variable number of flavors (the thresholds for c- and b-quarks are 1.5 GeV and 4.5 GeV correspondingly). All experimental errors including correlated ones are taken into account for calculation of the errors in PDFs using the covariance matrix approach [86] and assuming linear propagation of errors. The results of the benchmark fit obtained with the code used in analysis of Refs. [69,79] are given in Tables 7 and 8. The total number of the fitted PDF parameters left is 14. The normalization parameters N_i for the gluon and valence quark distributions are calculated from the momentum and fermion number conservation. The remaining normalization parameter A_S gives the total momentum carried by the sea distributions. Important note is that in view of many model assumptions made in the fit these results can be used mainly for the purposes of benchmarking rather for the phenomenological studies.

7 Benchmark Partons from DIS data and a Comparison with Global Fit Partons [9]

In this article I consider the uncertainties on partons arising from the errors on the experimental data that are used in a parton analysis. Various groups [87], [88], [69], [89], [76], [90], [91] have concentrated on the experimental errors and have obtained estimates of the uncertainties on parton distributions within a NLO QCD framework, using a variety of competing procedures. Here the two analyses, performed by myself and S. Alekhin (see Sec. 6) minimise the differences one obtains for the central values of the partons and the size of the uncertainties by fitting to exactly the same data sets with the same cuts, and using the same theoretical prescription. In order to be conservative we use only DIS data — BCDMS proton [73] and deuterium [74] fixed target data, NMC data on proton DIS and on the ratio $F_2^n(x,Q^2)/F_2^p(x,Q^2)$ [75], and H1 [76] and ZEUS [77] DIS data. We also apply cuts of $Q^2 = 9\text{GeV}^2$ and $W^2 = 15\text{GeV}^2$ in order to avoid the influence of higher twist. We each use NLO perturbative QCD in the $\overline{\rm MS}$ renormalization and factorization scheme, with the zero-mass variable flavour number scheme and quark masses of $m_c = 1.5\text{GeV}$ and $m_b = 4.5\text{GeV}$. There is a very minor difference between $\alpha_S(\mu^2)$ used in the two fitting programs due to the different methods of implementing heavy quark thresholds (the differences being formally of higher order), as observed in the study by M. Whalley for this workshop [92]. If the couplings in the two approaches have the same value at $\mu^2 = M_Z^2$, then the MRST value is $\sim 1\%$ higher for $Q^2 \sim 20\text{GeV}^2$.

We each input our parton distributions at $Q_0^2 = 1\text{GeV}^2$ with a parameterization of the form

$$xf_i(x,Q_0^2) = A_i(1-x)^{b_i}(1+\epsilon_i x^{0.5}+\gamma_i x)x^{a_i}. \tag{7.24}$$

The input sea is constrained to be 40% up and anti-up quarks, 40% down and anti-down quarks, and 20% strange and antistrange. No difference between $\bar{u}$ and $\bar{d}$ is input. There is no negative term for the gluon, as introduced in [90], since this restricted form of data shows no strong requirement for it in order

[8] https://mail.ihep.ru/~alekhin/benchmark/TABLE

[9] Contributing author: R.S. Thorne.

Table 8: Correlation coefficients for the parameters obtained in the benchmark fit.

	a_u	b_u	ϵ_u	γ_u	a_d	b_d	ϵ_d	γ_d	A_S	a_s	b_s	a_G	b_G	$\alpha_s(M_Z)$
a_u	1.000	0.728	-0.754	-0.708	0.763	0.696	-0.444	0.215	-0.216	-0.473	-0.686	0.593	0.777	-0.006
b_u	0.728	1.000	-0.956	-0.088	0.377	0.620	-0.420	0.387	0.175	-0.182	-0.713	0.067	0.505	-0.337
ϵ_u	-0.754	-0.956	1.000	0.105	-0.388	-0.662	0.503	-0.485	-0.229	0.059	0.600	-0.047	-0.503	0.276
γ_u	-0.708	-0.088	0.105	1.000	-0.741	-0.390	0.219	0.107	0.597	0.591	0.310	-0.716	-0.675	-0.088
a_d	0.763	0.377	-0.388	-0.741	1.000	0.805	-0.622	0.248	-0.367	-0.509	-0.528	0.652	0.664	0.101
b_d	0.696	0.620	-0.662	-0.390	0.805	1.000	-0.904	0.728	0.017	-0.193	-0.512	0.272	0.576	-0.136
ϵ_d	-0.444	-0.420	0.503	0.219	-0.622	-0.904	1.000	-0.896	-0.132	-0.019	0.245	-0.038	-0.362	0.173
γ_d	0.215	0.387	-0.485	0.107	0.248	0.728	-0.896	1.000	0.346	0.240	-0.107	-0.241	0.120	-0.228
A_S	-0.216	0.175	-0.229	0.597	-0.367	0.017	-0.132	0.346	1.000	0.708	0.127	-0.375	-0.026	0.047
a_s	-0.473	-0.182	0.059	0.591	-0.509	-0.193	-0.019	0.240	0.708	1.000	0.589	-0.595	-0.241	-0.011
b_s	-0.686	-0.713	0.600	0.310	-0.528	-0.512	0.245	-0.107	0.127	0.589	1.000	-0.508	-0.402	-0.109
a_G	0.593	0.067	-0.047	-0.716	0.652	0.272	-0.038	-0.241	-0.375	-0.595	-0.508	1.000	0.565	0.587
b_G	0.777	0.505	-0.503	-0.675	0.664	0.576	-0.362	0.120	-0.026	-0.241	-0.402	0.565	1.000	-0.138
$\alpha_s(M_Z)$	-0.006	-0.337	0.276	-0.088	0.101	-0.136	0.173	-0.228	0.047	-0.011	-0.109	0.587	-0.138	1.000

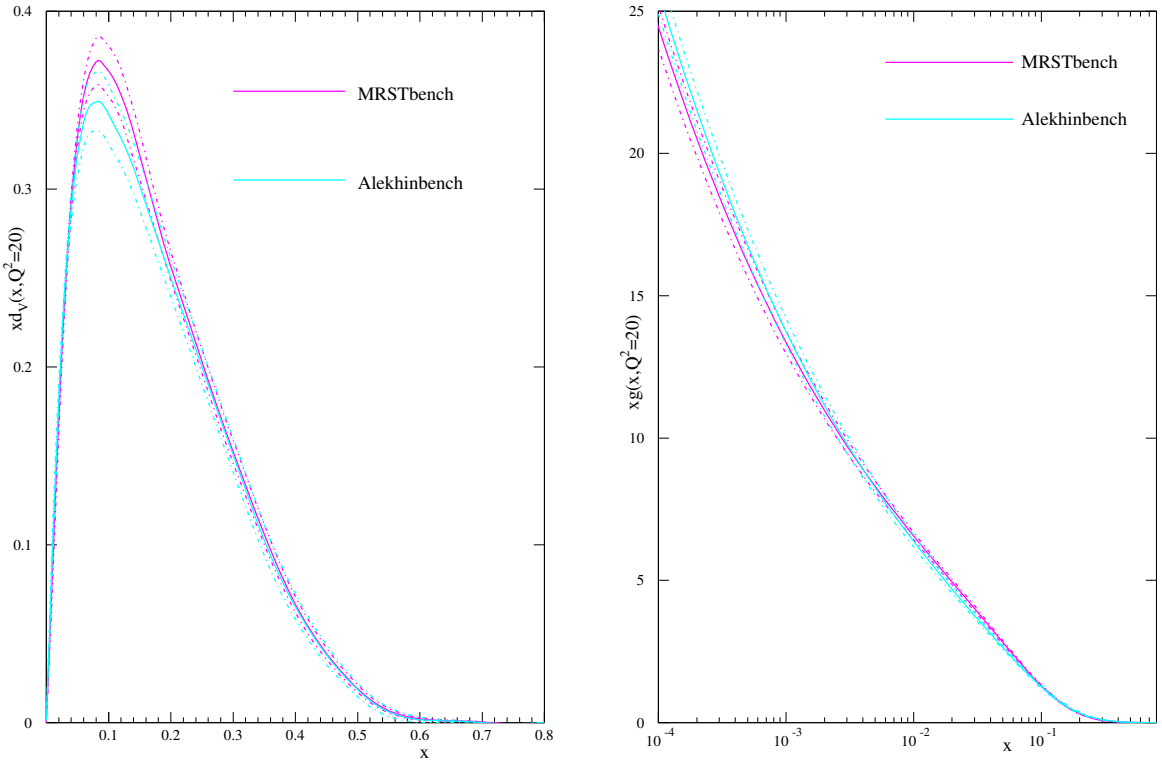

Fig. 18: Left plot: $xd_V(x, 20)$ from the MRST benchmark partons compared to that from the Alekhin benchmark partons. Right plot: $xg(x, 20)$ from the MRST benchmark partons compared to that from the Alekhin benchmark partons.

to obtain the best fit. Similarly we are able to set ϵ_g, γ_g, ϵ_S and γ_S all equal to zero. A_g is set by the momentum sum rule and A_{u_V} and A_{d_V} are set by valence quark number. Hence, there are nominally 13 free parton parameters. However, the MRST fitting program exhibited instability in the error matrix due to a very high correlation between u_V parameters, so ϵ_u was set at its best fit value of $\epsilon_u = -1.56$, while 12 parameters were free to vary. The coupling was also allowed to vary in order to obtain the best fit. The treatment of the errors on the data was exactly as for the published partons with uncertainties for each group, i.e. as in [69] and [93]. This means that all detail on correlations between errors is included for the Alekhin fit (see Sec. 6), assuming that these errors are distributed in the Gaussian manner. The errors in the MRST fit are treated as explained in the appendix of [93], and the correlated errors are not allowed to move the central values of the data to as great an extent for the HERA data, and cannot do so at all for the fixed target data, where the data used are averaged over the different beam energies. The Alekhin approach is more statistically rigorous. The MRST approach is more pragmatic, reducing the ability of the data to move relative to the theory comparison by use of correlated errors (other than normalization), and is in some ways similar to the offset method [91]. The danger of this movement of data relative to theory has been suggested by the joint analysis of H1 and ZEUS data at this workshop (see [94]), where letting the joint data sets determine the movement due to correlated errors gives different results from when the data sets are compared to theoretical results.

7.1 Comparison Between the Benchmark Parton Distributions.

I compare the results of the two approaches to fitting the restricted data chosen for the benchmarking. The input parameters for the Alekhin fit are presented in Sec. 6. Those for the MRST type fit are similar, but there are some differences which are best illustrated by comparing the partons at a typical Q^2 for the data, e.g. $Q^2 = 20\text{GeV}^2$. A comparison is shown for the d_V quarks and the gluon in Fig. 18.

From the plots it is clear that there is generally good agreement between the parton distributions. The central values are usually very close, and nearly always within the uncertainties. The difference in the central values is mainly due to the different treatment of correlated errors, and partially due to the difference in the coupling definition. The uncertainties are similar in the two sets, but are generally about $1.2 - 1.5$ times larger for the Alekhin partons, due to the increased freedom in the use of the correlated experimental errors. The values of $\alpha_S(M_Z^2)$ are quite different, $\alpha_S(M_Z^2) = 0.1132 \pm 0.0015$ compared to 0.1110 ± 0.0012. However, as mentioned earlier, one expects a 1% difference due to the different threshold prescriptions — the MRST α_S would be larger at $Q^2 \sim 20\mathrm{GeV}^2$, where the data are concentrated, so correspondingly to fit the data it receives a 1% shift downwards for $Q^2 = M_Z^2$. Once this systematic effect is taken into account, the values of $\alpha_S(M_Z^2)$ are very compatible. Hence, there is no surprising inconsistency between the two sets of parton distributions.

7.2 Comparison of the Benchmark Parton Distributions and Global Fit Partons.

It is also illuminating to show the comparison between the benchmark partons and the published partons from a global fit. This is done below for the MRST01 partons. For example, $u_V(x, Q^2)$ and $\bar{u}(x, Q^2)$ are shown in Fig. 19. It is striking that the uncertainties in the two sets are rather similar. This is despite the fact that the uncertainty on the benchmark partons is obtained from allowing $\Delta\chi^2 = 1$ in the fit while that for the MRST01 partons is obtained from $\Delta\chi^2 = 50$.[10] This illustrates the great improvement in precision which is obtained due to the increase in data from the relaxation of the cuts and the inclusion of types of data other than DIS. For the u_V partons, which are those most directly constrained by the DIS data in the benchmark fit, the comparison between the two sets of partons is reasonable, but hardly perfect — the central values differing by a few standard deviations. This is particularly important given that in this comparison the treatment of the data in the fit has been exactly the same in both cases. There is a minor difference in theoretical approach because of the simplistic treatment of heavy flavours in the benchmark fit. However, this would influence the gluon and sea quarks rather than valence quarks. Moreover, the region sensitive to this simplification would be $Q^2 \sim m_c^2$ (the lower charge weighting for bottom quarks greatly reducing the effect near $Q^2 = m_b^2$) which is removed by the Q^2 cut of $9\mathrm{GeV}^2$. Indeed, introducing the variable flavour number scheme usually used for the MRST partons modifies the benchmark partons only very minimally. Hence, if the statistical analysis is correct, the benchmark partons should agree with the global partons within their uncertainties (or at most 1.5 times their uncertainties, allowing for the effect of the correlated errors), which they do not. For the $\bar{u}$ partons the comparison is far worse, the benchmark partons being far larger at high x.

This disagreement in the high-x $\bar{u}$ partons can be understood better if one also looks at the high-x d_V distribution shown in Fig. 20. Here the benchmark distribution is very much smaller than for MRST01. However, the increase in the sea distribution, which is common to protons and neutrons, at high-x has allowed a good fit to the high-x BCDMS deuterium data even with the very small high-x d_V distribution. In fact it is a better fit than in [93]. However, the fit can be shown to break down with the additional inclusion of high-x SLAC data [72] on the deuterium structure function. More dramatically, the shape of the $\bar{u}$ is also completely incompatible with the Drell-Yan data usually included in the global fit, e.g. [82, 95]. Also in Fig. 20 we see that the d_V distributions are very different at smaller x. The benchmark set is markedly inconsistent with NMC data on $F_2^n(x, Q^2)/F_2^p(x, Q^2)$ which is at small x, but below the cut of $Q^2 = 9\mathrm{GeV}^2$.

The gluon from the benchmark set is also compared to the MRST01 gluon in Fig. 21. Again there is an enormous difference at high x. Nominally the benchmark gluon has little to constrain it at high x. However, the momentum sum rule determines it to be very small in this region in order to get the best fit to HERA data, similar to the gluon from [76]. As such, the gluon has a small uncertainty and is many standard deviations from the MRST01 gluon. Indeed, the input gluon at high x is so small that its value at higher Q^2 is dominated by the evolution of u_V quarks to gluons, rather than by the input gluon. Hence,

[10] Though it is meant to be interpreted as a one sigma error in the former case and a 90% confidence limit in the latter.

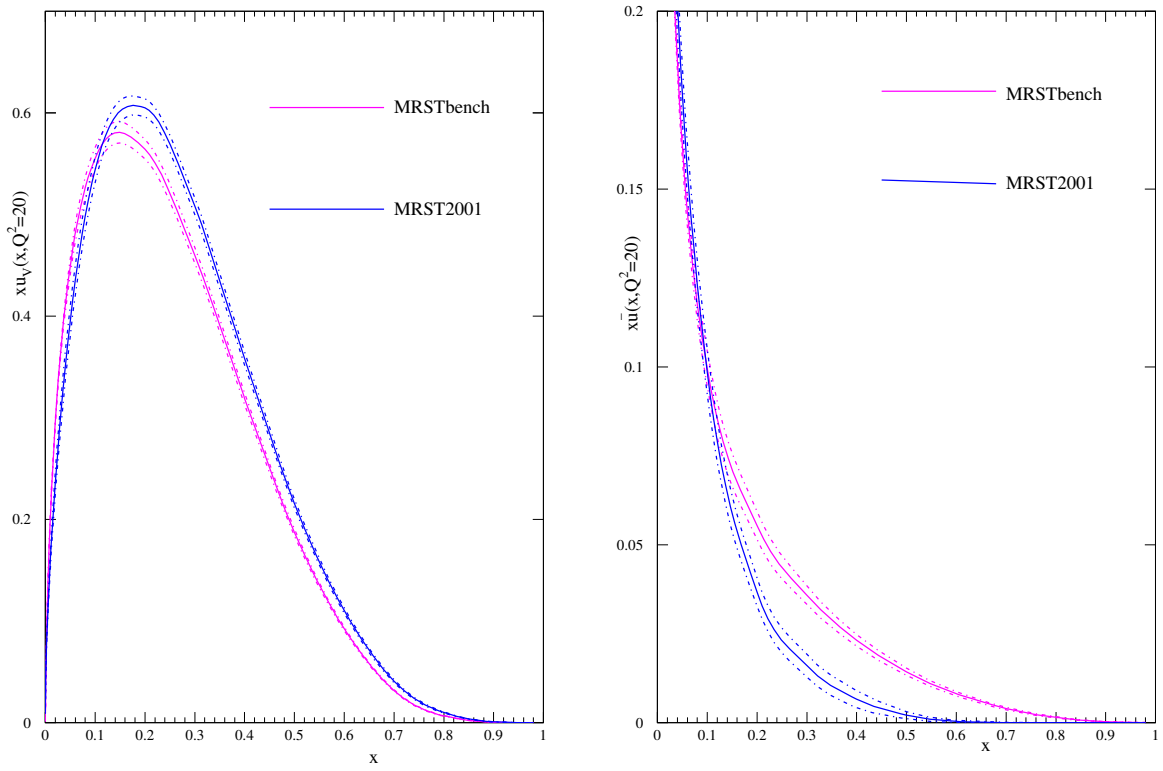

Fig. 19: Left plot: $xu_V(x, 20)$ from the MRST benchmark partons compared to that from the MRST01 partons. Right plot: $x\bar{u}(x, 20)$ from the MRST benchmark partons compared to that from the MRST01 partons with emphasis on large x.

the uncertainty is dominated by the quark parton input uncertainty rather than its own, and since the up quark is well determined the uncertainty on the high-x gluon is small for the benchmark partons. The smallness of the high-x gluon results in the benchmark partons producing a very poor prediction indeed for the Tevatron jet data [96, 97], which are the usual data that constrain the high-x gluon in global fits.

It is also illustrative to look at small x. Here the benchmark gluon is only a couple of standard deviations from the MRST01 gluon, suggesting that its size is not completely incompatible with a good fit to the HERA small-x data at Q^2 below the benchmark cut. However, the uncertainty in the benchmark gluon is much smaller than in the MRST01 gluon, despite the much smaller amount of low-x data in the fit for the benchmark partons. This comes about as a result of the artificial choice made in the gluon input at Q_0^2. Since it does not have the term introduced in [93], allowing the freedom for the input gluon to be negative at very small x, the gluon is required by the fit to be valence-like. Hence, at input it is simply very small at small x. At higher Q^2 it becomes much larger, but in a manner driven entirely by evolution, i.e. it is determined by the input gluon at moderate x, which is well constrained. In this framework the small-x gluon does not have any intrinsic uncertainty — its uncertainty is a reflection of moderate x. This is a feature of e.g. the CTEQ6 gluon uncertainty [89], where the input gluon is valence-like. In this case the percentage gluon uncertainty does not get any larger once x reaches about 0.001. The alternative treatment in [93] gives the expected increase in the gluon uncertainty as $x \to 0$, since in this case the uncertainty is determined largely by that in the input gluon at small x. The valence-like input form for a gluon is an example of fine-tuning, the form being unstable to evolution in either direction. The artificial limit on the small-x uncertainty is a consequence of this.

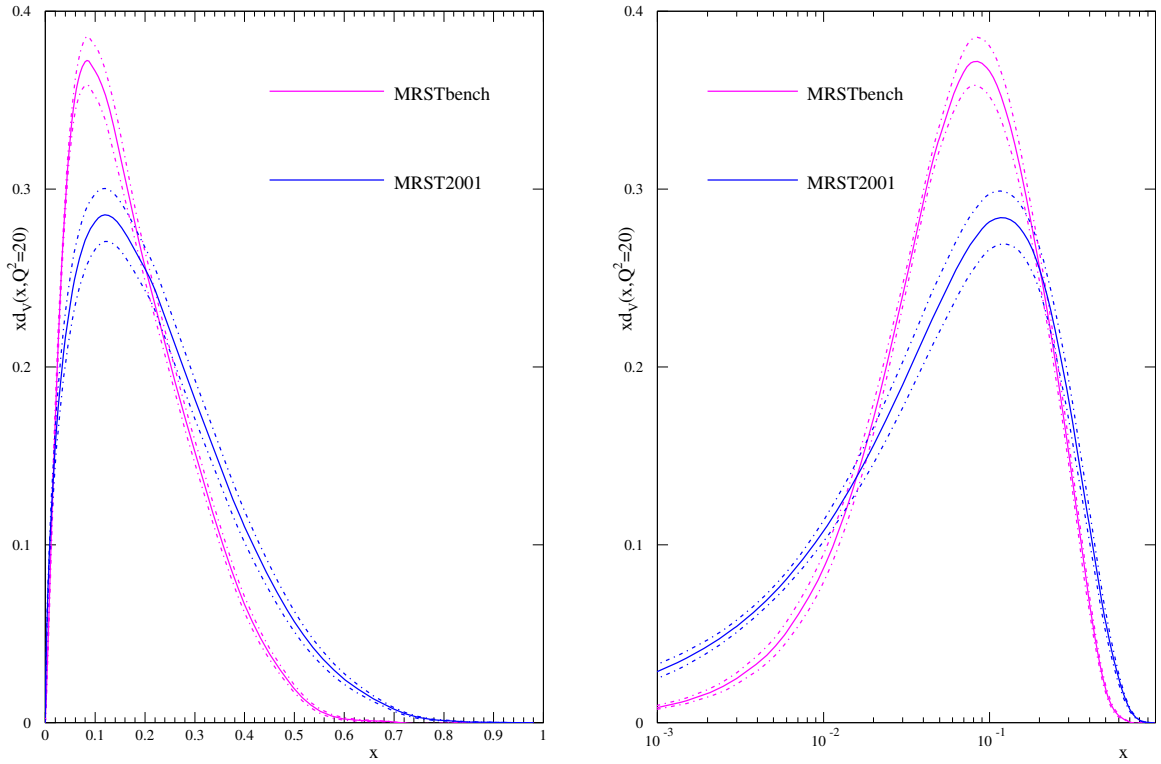

Fig. 20: Left plot: $xd_V(x, 20)$ from the MRST benchmark partons compared to that from the MRST01 partons. Right plot: $xd_V(x, 20)$ from the MRST benchmark partons compared to that from the MRST01 partons with emphasis on small x.

7.3 Conclusions.

I have demonstrated that different approaches to fitting parton distributions that use exactly the same data and theoretical framework produce partons that are very similar and have comparable uncertainties. There are certainly some differences due to the alternative approaches to dealing with experimental errors, but these are relatively small. However, the partons extracted using a very limited data set are completely incompatible, even allowing for the uncertainties, with those obtained from a global fit with an identical treatment of errors and a minor difference in theoretical procedure. This implies that the inclusion of more data from a variety of different experiments moves the central values of the partons in a manner indicating either that the different experimental data are inconsistent with each other, or that the theoretical framework is inadequate for correctly describing the full range of data. To a certain extent both explanations are probably true. Some data sets are not entirely consistent with each other (even if they are seemingly equally reliable). Also, there are a wide variety of reasons why NLO perturbative QCD might require modification for some data sets, or in some kinematic regions [98]. Whatever the reason for the inconsistency between the MRST benchmark partons and the MRST01 partons, the comparison exhibits the dangers in extracting partons from a very limited set of data and taking them seriously. It also clearly illustrates the problems in determining the true uncertainty on parton distributions.

8 Stability of PDF fits [11]

One of the issues raised at the workshop is the reliability of determinations of parton distribution functions (PDFs), which might be compromised for example by the neglect of NNLO effects or non-DGLAP evolution in the standard analysis, or hidden assumptions made in parameterizing the PDFs at nonper-

[11] Contributing authors: J. Huston, J. Pumplin.

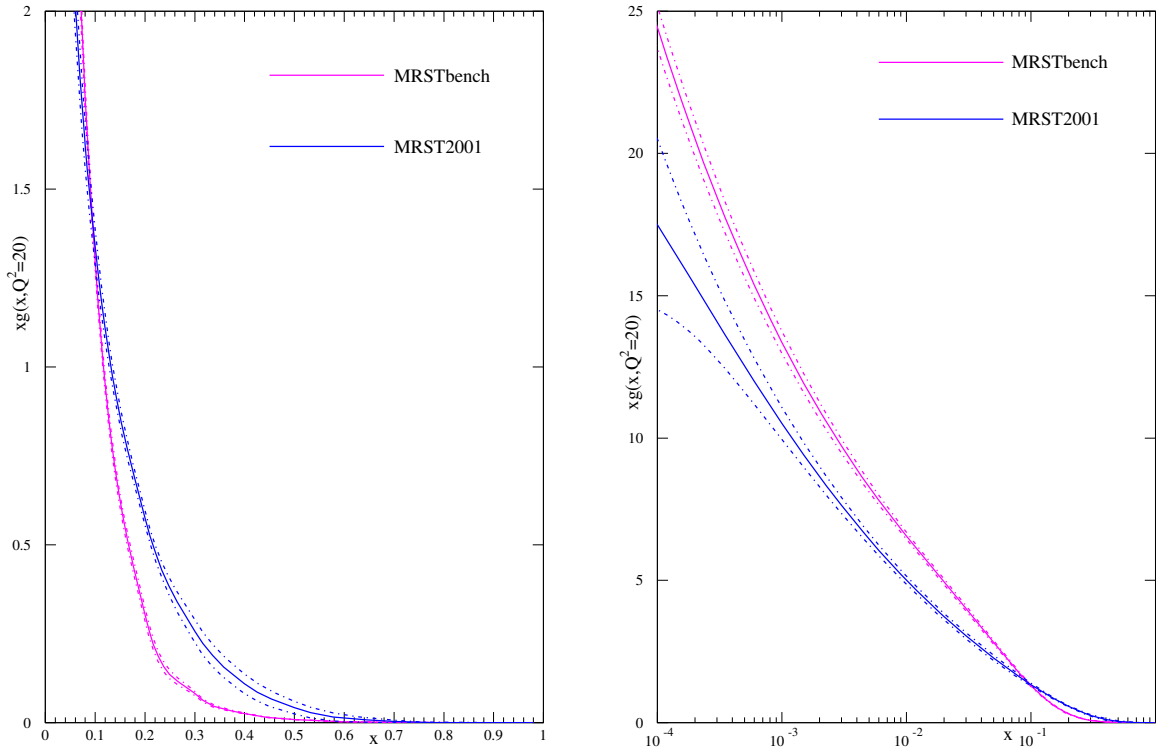

Fig. 21: Left plot: $xg(x, 20)$ from the MRST benchmark partons compared to that from the MRST2001 partons. Right plot: $xg(x, 20)$ from the MRST benchmark partons compared to that from the MRST2001 partons with emphasis on small x.

turbative scales. We summarize the results of the CTEQ PDF group on this issue. For the full story see [80].

8.1 Stability of PDF determinations

The stability of NLO global analysis was seriously challenged by an analysis [98] which found a 20% variation in the cross section predicted for W production at the LHC — a critical "standard candle" process for hadron colliders — when certain cuts on input data are varied. If this instability were confirmed, it would significantly impact the phenomenology of a wide range of physical processes for the Tevatron Run II and the LHC. The CTEQ PDF group therefore performed an independent study of this issue within their global analysis framework. In addition, to explore the dependence of the results on assumptions about the parameterization of PDFs at the starting scale $Q_0 = 1.3\,\mathrm{GeV}$, we also studied the effect of allowing a negative gluon distribution at small x — a possibility that is favored by the MRST NLO analysis, and that is closely tied to the W cross section controversy.

The stability of the global analysis was investigated by varying the inherent choices that must be made to perform the analysis. These choices include the selection of experimental data points based on kinematic cuts, the functional forms used to parameterize the initial nonperturbative parton distribution functions, and the treatment of α_s.

The stability of the results is most conveniently measured by differences in the global χ^2 for the relevant fits. To quantitatively define a change of χ^2 that characterizes a significant change in the quality of the PDF fit is a difficult issue in global QCD analysis. In the context of the current analysis, we have argued that an increase by $\Delta\chi^2 \sim 100$ (for $\sim$ 2000 data points) represents roughly a 90% confidence level uncertainty on PDFs due to the uncertainties of the current input experimental data [89, 99–101].

Table 9: Comparisons of three fits with different choices of the cuts on input data at the Q and x values indicated. In these fits, a conventional positive-definite gluon parameterization was used.

Cuts	$Q_{\min}$	$x_{\min}$	$N_{\rm pts}$	χ^2_{1926}	χ^2_{1770}	χ^2_{1588}	$\sigma_W^{\rm LHC} \times B_{\ell\nu}$ [nb]
standard	2 GeV	0	1926	2023	1850	1583	20.02
intermediate	2.5 GeV	0.001	1770	–	1849	1579	20.10
strong	3.162 GeV	0.005	1588	–	–	1573	20.34

Table 10: Same as Table 9 except that the gluon parameterization is extended to allow negative values.

Cuts	$Q_{\min}$	$x_{\min}$	$N_{\rm pts}$	χ^2_{1926}	χ^2_{1770}	χ^2_{1588}	$\sigma_W^{\rm LHC} \times B_{\ell\nu}$ [nb]
standard	2 GeV	0	1926	2011	1845	1579	19.94
intermediate	2.5 GeV	0.001	1770	–	1838	1574	19.80
strong	3.162 GeV	0.005	1588	–	–	1570	19.15

In other words, PDFs with $\chi^2 - \chi^2_{\rm BestFit} > 100$ are regarded as not tolerated by current data.

The CTEQ6 and previous CTEQ global fits imposed "standard" cuts $Q > 2\,{\rm GeV}$ and $W > 3.5\,{\rm GeV}$ on the input data set, in order to suppress higher-order terms in the perturbative expansion and the effects of resummation and power-law ("higher twist") corrections. We examined the effect of stronger cuts on Q to see if the fits are stable. We also examined the effect of imposing cuts on x, which should serve to suppress any errors due to deviations from DGLAP evolution, such as those predicted by BFKL. The idea is that any inconsistency in the global fit due to data points near the boundary of the accepted region will be revealed by an improvement in the fit to the data that remain after those near-boundary points have been removed. In other words, the decrease in χ^2 for the subset of data that is retained, when the PDF shape parameters are refitted to that subset alone, measures the degree to which the fit to that subset was distorted in the original fit by compromises imposed by the data at low x and/or low Q.

The main results of this study are presented in Table 9. Three fits are shown, from three choices of the cuts on input data as specified in the table. They are labeled 'standard', 'intermediate' and 'strong'. $N_{\rm pts}$ is the number of data points that pass the cuts in each case, and $\chi^2_{N_{\rm pts}}$ is the χ^2 value for that subset of data. The fact that the changes in χ^2 in each column are insignificant compared to the uncertainty tolerance is strong evidence that our NLO global fit results are very stable with respect to choices of kinematic cuts.

We extended the analysis to a series of fits in which the gluon distribution $g(x)$ is allowed to be negative at small x, at the scale $Q_0 = 1.3\,{\rm GeV}$ where we begin the DGLAP evolution. The purpose of this additional study is to determine whether the feature of a negative gluon PDF is a key element in the stability puzzle, as suggested by the findings of [98]. The results are presented in Table 10. Even in this extended case, we find no evidence of instability. For example, χ^2 for the subset of 1588 points that pass the *strong* cuts increases only from 1570 to 1579 when the fit is extended to include the full standard data set.

Comparing the elements of Table 9 and Table 10 shows that our fits with $g(x) < 0$ have slightly smaller values ofχ^2: e.g., 2011 versus 2023 for the standard cuts. However, the difference $\Delta\chi^2 = 12$ between these values is again not significant according to our tolerance criterion.

8.2 W cross sections at the LHC

The last columns of Tables 9 and 10 show the predicted cross section for $W^+ + W^-$ production at the LHC. This prediction is also very stable: it changes by only 1.6% for the positive-definite gluon parameterization, which is substantially less than the overall PDF uncertainty of σ_W estimated previously with the standard cuts. For the negative gluon parameterization, the change is 4%–larger, but still less than the overall PDF uncertainty. These results are explicitly displayed, and compared to the MRST results in Fig. 22. We see that this physical prediction is indeed insensitive to the kinematic cuts used for

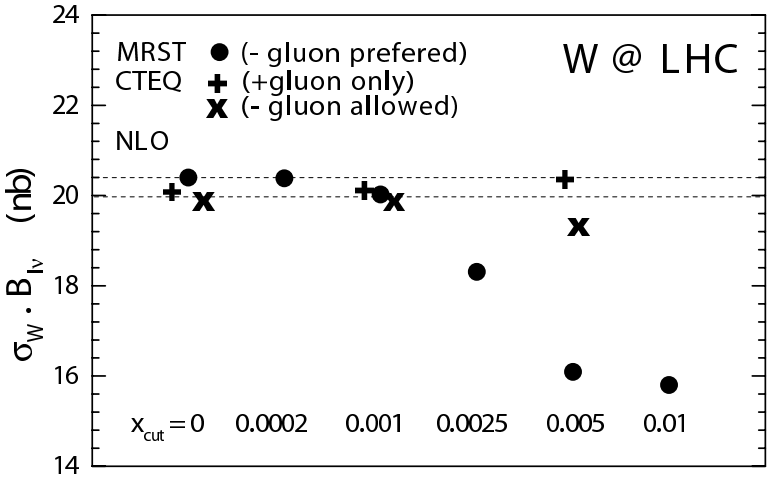

Fig. 22: Predicted total cross section of $W^+ + W^-$ production at the LHC for the fits obtained in our stability study, compared to the NLO results of Ref. [98]. The Q-cut values associated with the CTEQ points are given in the two tables. The overall PDF uncertainty of the prediction is $\sim 5\%$.

the fits, and to the assumption on the positive definiteness of the gluon distribution.

We also studied the stability of the prediction for σ_W using the Lagrange Multiplier (LM) method of Refs. [99–101]. Specifically, we performed a series of fits to the global data set that are constrained to specific values of σ_W close to the best-fit prediction. The resulting variation of χ^2 versus σ_W measures the uncertainty of the prediction. We repeated the constrained fits for each case of fitting choices (parameterization and kinematic cuts). In this way we gain an understanding of the stability of the uncertainty, in addition to the stability of the central prediction.

Figure 23 shows the results of the LM study for the three sets of kinematic cuts described in Table 9, all of which have a positive-definite gluon distribution. The χ^2 shown along the vertical axis is normalized to its value for the best fit in each series. In all three series, χ^2 depends almost quadratically on σ_W. We observe several features:

- The location of the minimum of each curve represents the best-fit prediction for $\sigma_W^{\rm LHC}$ for the corresponding choice of cuts. The fact that the three minima are close together displays the stability of the predicted cross section already seen in Table 9.
- Although more restrictive cuts make the global fit less sensitive to possible contributions from resummation, power-law and other nonperturbative effects, the loss of constraints caused by the removal of precision HERA data points at small x and low Q results directly in increased uncertainties on the PDF parameters and their physical predictions. This is shown in Fig. 23 by the increase of the width of the curves with stronger cuts. The uncertainty of the predicted σ_W increases by more than a factor of 2 in going from the standard cuts to the strong cuts.

Figure 24 shows the results of the LM study for the three sets of kinematic cuts described in Table 10, all of which have a gluon distribution which is allowed to go negative.

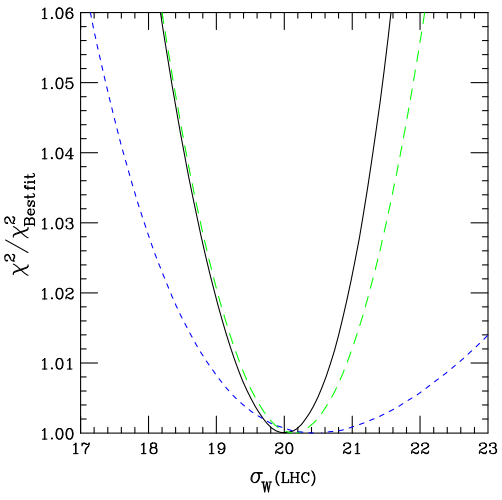

Fig. 23: Lagrange multiplier results for the W cross section (in nb) at the LHC using a positive-definite gluon. The three curves, in order of decreasing steepness, correspond to the three sets of kinematic cuts labeled standard/intermediate/strong in Table 9.

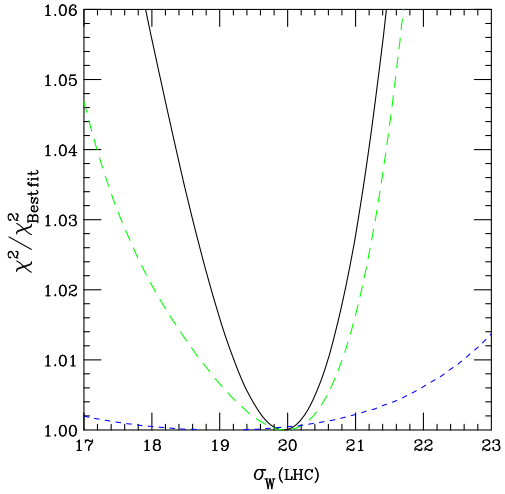

Fig. 24: Lagrange multiplier results for the W cross section (in nb) at the LHC using a functional form where the gluon is not required to be positive-definite. The three curves, in order of decreasing steepness, correspond to the three sets of kinematic cuts labeled standard/intermediate/strong in Table 10.

We observe:

- Removing the positive definiteness condition necessarily lowers the value of χ^2, because more possibilities are opened up in the χ^2 minimization procedure. But the decrease is insignificant compared to other sources of uncertainty. Thus, a negative gluon PDF is allowed, but not required.
- The minima of the two curves occur at approximately the same σ_W. Allowing a negative gluon makes no significant change in the central prediction — merely a decrease of about $1\,\%$, which is small compared to the overall PDF uncertainty.
- For the standard set of cuts, allowing a negative gluon PDF would expand the uncertainty range only slightly. For the intermediate and strong cuts, allowing a negative gluon PDF would significantly expand the uncertainty range.

We examined a number of aspects of our analysis that might account for the difference in conclusions between our stability study and that of [98]. A likely candidate seems to be that in order to obtain stability, it is necessary to allow a rather free parametrization of the input gluon distribution. This suspicion is seconded by recent work by MRST [102], in which a different gluon parametrization appears to lead to a best-fit gluon distribution that is close to that of CTEQ6. In summary, we found that the NLO PDFs and their physical predictions at the Tevatron and LHC are quite stable with respect to variations of the kinematic cuts and the PDF parametrization after all.

8.3 NLO and NNLO

In recent years, some preliminary next-to-next-leading-order (NNLO) analyses for PDFs have been carried out either for DIS alone [103], or in a global analysis context [51] — even if all the necessary hard cross sections, such as inclusive jet production, are not yet available at this order. Determining the parton distributions at NNLO is obviously desirable on theoretical grounds, and it is reasonable to plan for having a full set of tools for a true NNLO global analysis in place by the time LHC data taking begins. At the moment, however, NNLO fitting is not a matter of pressing necessity, since the difference between NLO and NNLO appears to be very small compared to the other uncertainties in the PDF analysis. This

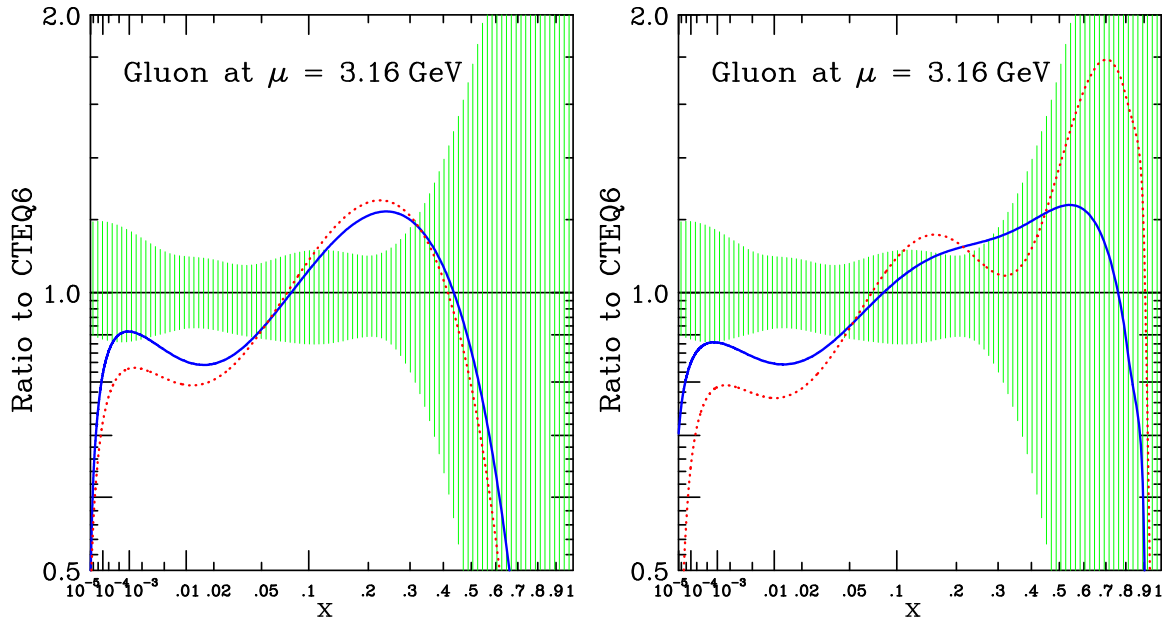

Fig. 25: Left: mrst2002 NLO (solid) and NNLO (dotted); Right: mrst2004 NLO (solid) and NNLO (dotted); Shaded region is uncertainty according to the 40 eigenvector sets of CTEQ6.1.

is demonstrated in Fig. 25, which shows the NLO and NNLO gluon distributions extracted by the MRST group. The difference between the two curves is much smaller than the other uncertainties measured by the 40 eigenvector uncertainty sets of CTEQ6.1, which is shown by the shaded region. The difference is also much smaller than the difference between CTEQ and MRST best fits. Similar conclusions [104] can be found using the NLO and NNLO fits by Alekhin.

9 The neural network approach to parton distributions [12]

The requirements of precision physics at hadron colliders, as has been emphasized through this workshop, have recently led to a rapid improvement in the techniques for the determination of parton distribution functions (pdfs) of the nucleon. Specifically it is now mandatory to determine accurately the uncertainty on these quantities, and the different collaborations performing global pdf analysis [51, 69, 105] have performed estimations of these uncertainties using a variety of techniques. The main difficulty is that one is trying to determine the uncertainty on a function, that is, a probability measure in a space of functions, and to extract it from a finite set of experimental data, a problem which is mathematically ill-posed. It is also known that the standard approach to global parton fits have several shortcomings: the bias introduced by choosing fixed functional forms to parametrize the parton distributions (also known as *model dependence*), the problems to assess faithfully the pdf uncertainties, the combination of inconsistent experiments, and the lack of general, process-independent error propagation techniques. Although the problem of quantifying the uncertainties in pdfs has seen a huge progress since its paramount importance was raised some years ago, until now no unambiguous conclusions have been obtained.

In this contribution we present a novel strategy to address the problem of constructing unbiased parametrizations of parton distributions with a faithful estimation of their uncertainties, based on a combination of two techniques: Monte Carlo methods and neural networks. This strategy, introduced in [106, 107], has been first implemented to address the marginally simpler problem of parametrizing deep-inelastic structure functions $F(x, Q^2)$, which we briefly summarize now. In a first step we construct a Monte Carlo sampling of the experimental data (generating artificial data replicas), and then we

[12] Contributing authors: L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, J. Rojo

train neural networks to each data replica, to construct a probability measure in the space of structure functions $\mathcal{P}\left[F(x,Q^2)\right]$. The probability measure constructed in this way contains all information from experimental data, including correlations, with the only assumption of smoothness. Expectation values and moments over this probability measure are then evaluated as averages over the trained network sample,

$$\left\langle \mathcal{F}\left[F(x,Q^2)\right]\right\rangle = \int \mathcal{D}F\mathcal{P}\left[F(x,Q^2)\right]\mathcal{F}\left[F(x,Q^2)\right] = \frac{1}{N_{\rm rep}}\sum_{k=1}^{N_{\rm rep}}\mathcal{F}\left(F^{({\rm net})(k)}(x,Q^2)\right) . \quad (9.25)$$

where $\mathcal{F}[F]$ is an arbitrary function of $F(x,Q^2)$.

The first step is the Monte Carlo sampling of experimental data, generating $N_{\rm rep}$ replicas of the original $N_{\rm dat}$ experimental data,

$$F_i^{({\rm art})(k)} = \left(1+r_N^{(k)}\sigma_N\right)\left[F_i^{({\rm exp})} + r_i^{s,(k)}\sigma_i^{stat} + \sum_{l=1}^{N_{sys}} r^{l,(k)}\sigma_i^{sys,l}\right], \qquad i=1,\ldots,N_{\rm dat} , \quad (9.26)$$

where r are gaussian random numbers with the same correlation as the respective uncertainties, and $\sigma^{stat},\sigma^{sys},\sigma_N$ are the statistical, systematic and normalization errors. The number of replicas $N_{\rm rep}$ has to be large enough so that the replica sample reproduces central values, errors and correlations of the experimental data.

The second step consists on training a neural network[13] on each of the data replicas. Neural networks are specially suitable to parametrize parton distributions since they are unbiased, robust approximants and interpolate between data points with the only assumption of smoothness. The neural network training consist on the minimization for each replica of the χ^2 defined with the inverse of the experimental covariance matrix,

$$\chi^{2(k)} = \frac{1}{N_{\rm dat}}\sum_{i,j=1}^{N_{\rm dat}}\left(F_i^{({\rm art})(k)} - F_i^{({\rm net})(k)}\right){\rm cov}_{ij}^{-1}\left(F_j^{({\rm art})(k)} - F_j^{({\rm net})(k)}\right) . \quad (9.27)$$

Our minimization strategy is based on Genetic Algorithms (introduced in [108]), which are specially suited for finding global minima in highly nonlinear minimization problems.

The set of trained nets, once is validated through suitable statistical estimators, becomes the sought-for probability measure $\mathcal{P}\left[F(x,Q^2)\right]$ in the space of structure functions. Now observables with errors and correlations can be computed from averages over this probability measure, using eq. (9.25). For example, the average and error of a structure function $F(x,Q^2)$ at arbitrary (x,Q^2) can be computed as

$$\left\langle F(x,Q^2)\right\rangle = \frac{1}{N_{\rm rep}}\sum_{k=1}^{N_{\rm rep}}F^{({\rm net})(k)}(x,Q^2), \quad \sigma(x,Q^2) = \sqrt{\langle F(x,Q^2)^2\rangle - \langle F(x,Q^2)\rangle^2} . \quad (9.28)$$

A more detailed account of the application of the neural network approach to structure functions can be found in [107], which describes the most recent NNPDF parametrization of the proton structure function[14].

Hence this strategy can be used also to parametrize parton distributions, provided one now takes into account perturbative QCD evolution. Therefore we need to define a suitable evolution formalism.

[13] For a more throughly description of neural network, see [106] and references therein

[14] The source code, driver program and graphical web interface for our structure function fits is available at `http://sophia.ecm.ub.es/f2neural`.

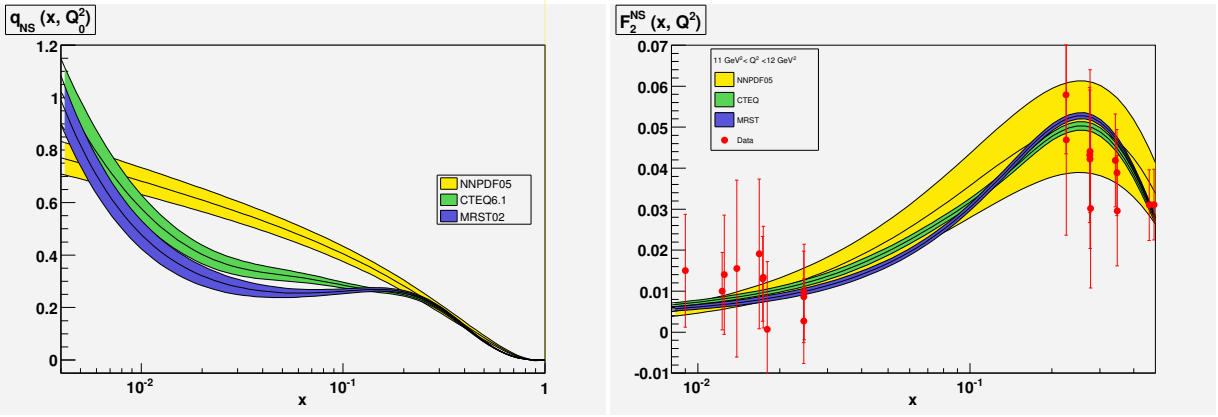

Fig. 26: Preliminary results for the NNPDF q_{NS} fit at $Q_0^2 = 2\,\mathrm{GeV}^2$, and the prediction for $F_2^{NS}(x, Q^2)$ compared with the CTEQ and MRST results.

Since complex neural networks are not allowed, we must use the convolution theorem to evolve parton distributions in $x-$space using the inverse $\Gamma(x)$ of the Mellin space evolution factor $\Gamma(N)$, defined as

$$q(N, Q^2) = q(N, Q_0^2)\Gamma\left(N, \alpha_s\left(Q^2\right), \alpha_s\left(Q_0^2\right)\right) \ , \tag{9.29}$$

The only subtlety is that the x-space evolution factor $\Gamma(x)$ is a distribution, which must therefore be regulated at $x = 1$, yielding the final evolution equation,

$$q(x, Q^2) = q(x, Q_0^2)\int_x^1 dy\,\Gamma(y) + \int_x^1 \frac{dy}{y}\Gamma(y)\left(q\left(\frac{x}{y}, Q_0^2\right) - yq(x, Q_0^2)\right) \ , \tag{9.30}$$

where in the above equation $q(x, Q_0^2)$ is parametrized using a neural network. At higher orders in perturbation theory coefficient functions $C(N)$ are introduced through a modified evolution factor, $\tilde{\Gamma}(N) \equiv \Gamma(N)C(N)$. We have benchmarked our evolution code with the Les Houches benchmark tables [59] at NNLO up to an accuracy of 10^{-5}. The evolution factor $\Gamma(x)$ and its integral are computed and interpolated before the neural network training in order to have a faster fitting procedure.

As a first application of our method, we extract the nonsinglet parton distribution $q_{NS}(x, Q_0^2) = \frac{1}{6}\left(u + \bar{u} - d - \bar{d}\right)(x, Q_0^2)$ from the nonsinglet structure function $F_2^{NS}(x, Q^2)$ as measured by the NMC [75] and BCDMS [73, 74] collaborations. The preliminary results of a NLO fit with fully correlated uncertainties [109] can be seen in fig. 26 compared to other pdfs sets. Our preliminary results appear to point in the direction that the uncertainties at small x do not allow, provided the current experimental data, to determine if $q_{NS}(x, Q^2)$ grows at small x, as supported by different theoretical arguments as well as by other global parton fits. However, more work is still needed to confirm these results. Only additional nonsinglet structure function data at small x could settle in a definitive way this issue[15].

Summarizing, we have described a general technique to parametrize experimental data in an bias-free way with a faithful estimation of their uncertainties, which has been successfully applied to structure functions and that now is being implemented in the context of parton distribution. The next step will be to construct a full set of parton distributions from all available hard-scattering data using the strategy described in this contribution.

References

[1] Van Neerven, W. L. and Zijlstra, E. B., Phys. Lett. **B272**, 127 (1991).

[15] Like the experimental low x deuteron structure function which would be measured in an hypothetical electron-deuteron run at HERA II, as it was pointed out during the workshop by M. Klein ([110]) and C. Gwenlan

[2] Zijlstra, E. B. and van Neerven, W. L., Phys. Lett. **B273**, 476 (1991).
[3] Zijlstra, E. B. and van Neerven, W. L., Phys. Lett. **B297**, 377 (1992).
[4] Zijlstra, E. B. and van Neerven, W. L., Nucl. Phys. **B383**, 525 (1992).
[5] Moch, S. and Vermaseren, J. A. M., Nucl. Phys. **B573**, 853 (2000).
[6] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. **B688**, 101 (2004).
[7] Vogt, A. and Moch, S. and Vermaseren, J. A. M., Nucl. Phys. **B691**, 129 (2004).
[8] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Phys. Lett. **B606**, 123 (2005).
[9] Vermaseren, J. A. M. and Vogt, A. and Moch, S., Nucl. Phys. **B724**, 3 (2005).
[10] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. **B646**, 181 (2002).
[11] Moch, S. and Vermaseren, J. A. M. and Vogt, A., Nucl. Phys. Proc. Suppl. **135**, 137 (2004).
[12] Catani, S. and Hautmann, F., Nucl. Phys. **B427**, 475 (1994).
[13] Fadin, V. S. and Lipatov, L. N., Phys. Lett. **B429**, 127 (1998).
[14] Kuraev, E. A. and Lipatov, L. N. and Fadin, Victor S., Sov. Phys. JETP **45**, 199 (1977).
[15] Balitsky, I. I. and Lipatov, L. N., Sov. J. Nucl. Phys. **28**, 822 (1978).
[16] Jaroszewicz, T., Phys. Lett. **B116**, 291 (1982).
[17] Kirschner, R. and Lipatov, L. N., Nucl. Phys. **B213**, 122 (1983).
[18] Blümlein, J. and Vogt, A., Phys. Lett. **B370**, 149 (1996).
[19] G. Altarelli et al., *Resummation*. These proceedings.
[20] Van Neerven, W. L. and Vogt, A., Nucl. Phys. **B603**, 42 (2001).
[21] Larin, S. A. and van Ritbergen, T. and Vermaseren, J. A. M., Nucl. Phys. **B427**, 41 (1994).
[22] Larin, S. A. and Nogueira, P. and van Ritbergen, T. and Vermaseren, J. A. M., Nucl. Phys. **B492**, 338 (1997).
[23] Retey, A. and Vermaseren, J. A. M., Nucl. Phys. **B604**, 281 (2001).
[24] Blümlein, J. and Kurth, S., Phys. Rev. **D60**, 014018 (1999).
[25] Blümlein, J., Comput. Phys. Commun. **159**, 19 (2004).
[26] Remiddi, E. and Vermaseren, J. A. M., Int. J. Mod. Phys. **A15**, 725 (2000).
[27] Borwein, J. M. and Bradley, D. M. and Broadhurst, D. J. and Lisonek, P., Trans. Am. Math. Soc. **353**, 907 (2001).
[28] Blümlein, J., *in preparation*.
[29] Blümlein, J. and Moch, S., *in preparation*.
[30] Blümlein, J., Nucl. Phys. Proc. Suppl. **135**, 225 (2004).
[31] Blümlein, J., Comput. Phys. Commun. **133**, 76 (2000).
[32] Blümlein, J. and Moch, S., Phys. Lett. **B614**, 53 (2005).
[33] Nielsen, N., Nova Acta Leopoldina (Halle) **90**, 123 (1909).
[34] Matsuura, T. and van der Marck, S. C. and van Neerven, W. L., Nucl. Phys. **B319**, 570 (1989).
[35] Hamberg, R. and van Neerven, W. L. and Matsuura, T., Nucl. Phys. **B359**, 343 (1991).
[36] Ravindran, V. and Smith, J. and van Neerven, W. L., Nucl. Phys. **B682**, 421 (2004).
[37] Catani, S. and de Florian, D. and Grazzini, M., JHEP **05**, 025 (2001).
[38] Harlander, R. V. and Kilgore, W. B., Phys. Rev. **D64**, 013015 (2001).
[39] Harlander, R. V. and Kilgore, W. B., Phys. Rev. Lett. **88**, 201801 (2002).
[40] Harlander, R. V. and Kilgore, W. B., JHEP **10**, 017 (2002).
[41] Anastasiou, C. and Melnikov, K., Nucl. Phys. **B646**, 220 (2002).
[42] Ravindran, V. and Smith, J. and van Neerven, W. L., Nucl. Phys. **B665**, 325 (2003).
[43] Rijken, P. J. and van Neerven, W. L., Phys. Lett. **B386**, 422 (1996).

[44] Rijken, P. J. and van Neerven, W. L., Nucl. Phys. **B487**, 233 (1997).
[45] Rijken, P. J. and van Neerven, W. L., Phys. Lett. **B392**, 207 (1997).
[46] Blümlein, J. and Ravindran, V., Nucl. Phys. Proc. Suppl. **135**, 24 (2004).
[47] Blümlein, J. and Ravindran, V., Nucl. Phys. **B716**, 128 (2005).
[48] Blümlein, J. and Böttcher, H. and Guffanti, A., Nucl. Phys. Proc. Suppl. **135**, 152 (2004).
[49] Blumlein, J. and Bottcher, H. and Guffanti, A., *in preparation*.
[50] Alekhin, S., Phys. Rev. **D63**, 094022 (2001).
[51] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Phys. Lett. **B531**, 216 (2002).
[52] Botje, M. and Klein, M. and Pascaud, C. (1996).
[53] Blümlein, J. and Riemersma, S. and van Neerven, W. L. and Vogt, A., Nucl. Phys. Proc. Suppl. **51C**, 97 (1996).
[54] Blümlein, J. and Guffanti, A., *Scheme-invariant nnlo evolution for unpolarized dis structure functions*. Preprint hep-ph/0411110, 2004.
[55] Furmanski, W. and Petronzio, R., Zeit. Phys. **C11**, 293 (1982).
[56] Catani, S., Z. Phys. **C75**, 665 (1997).
[57] Blümlein, J. and Ravindran, V. and van Neerven, W. L., Nucl. Phys. **B586**, 349 (2000).
[58] Alekhin, S. and Blümlein, J., Phys. Lett. **B594**, 299 (2004).
[59] Giele, W. et al, *The qcd/sm working group: Summary report*. Preprint hep-ph/0204316, 2002.
[60] Blümlein, J. and Böttcher, H., Nucl. Phys. **B636**, 225 (2002).
[61] Van Neerven, W. L. and Vogt, A., Phys. Lett. **B490**, 111 (2000).
[62] Blümlein, J. et al., *A detailed comparison of nlo qcd evolution codes*. Preprint hep-ph/9609400, 1996.
[63] Vogt, A., Comput. Phys. Commun. **170**, 65 (2005).
[64] Lai, H. L. et al., Eur. Phys. J. **C12**, 375 (2000).
[65] Buza, M. and Matiounine, Y. and Smith, J. and van Neerven, W. L., Eur. Phys. J. **C1**, 301 (1998).
[66] Chuvakin, A. and Smith, J., Comput. Phys. Commun. **143**, 257 (2002).
[67] Larin, S. A. and van Ritbergen, T. and Vermaseren, J. A. M., Nucl. Phys. **B438**, 278 (1995).
[68] Chetyrkin, K. G. and Kniehl, B. A. and Steinhauser, M., Phys. Rev. Lett. **79**, 2184 (1997).
[69] Alekhin, S., Phys. Rev. **D68**, 014002 (2003).
[70] Alwall, J. and Ingelman, G., Phys. Rev. **D71**, 094015 (2005);
Alwall, J. and Ingelman, G., Phys. Rev. **D70**, 111505 (2004).
[71] Alwall, J. and Ingelman, G., Phys. Lett. **B596**, 77 (2004).
[72] Whitlow, L. W. and Riordan, E. M. and Dasu, S. and Rock, Stephen and Bodek, A., Phys. Lett. **B282**, 475 (1992).
[73] Benvenuti, A. C. et al., Phys. Lett. **B223**, 485 (1989).
[74] Benvenuti, A. C. et al., Phys. Lett. **B237**, 592 (1990).
[75] Arneodo, M. et al., Nucl. Phys. **B483**, 3 (1997).
[76] Adloff, C. et al., Eur. Phys. J. **C21**, 33 (2001).
[77] Chekanov, S. et al., Eur. Phys. J. **C21**, 443 (2001).
[78] Kazakov, D. I. and Kotikov, A. V., Phys. Lett. **B291**, 171 (1992).
[79] Alekhin, S., *Parton distribution functions from the precise nnlo qcd fit*. Preprint hep-ph/0508248, 2005.
[80] Huston, J. and Pumplin, J. and Stump, D. and Tung, W. K., JHEP **06**, 080 (2005).

[81] Altarelli, G. and Ball, R. D. and Forte, S., Nucl. Phys. **B674**, 459 (2003).

[82] Moreno, G. et al., Phys. Rev. **D43**, 2815 (1991).

[83] Towell, R. S. et al., Phys. Rev. **D64**, 052002 (2001).

[84] Anastasiou, C. and Dixon, L. J. and Melnikov, K. and Petriello, F., Phys. Rev. Lett. **91**, 182002 (2003).

[85] Bazarko, A. O. et al., Z. Phys. **C65**, 189 (1995).

[86] Alekhin, S., *Statistical properties of the estimator using covariance matrix*. Preprint hep-ex/0005042, 2000.

[87] Botje, M., Eur. Phys. J. **C14**, 285 (2000).

[88] Giele, W.T. and Keller, S., Phys. Rev. **D58**, 094023 (1998);
Giele, Walter T. and Keller, Stephane A. and Kosower, David A., *Parton distribution function uncertainties*. Preprint hep-ph/0104052, 2001.

[89] Pumplin, J. et al., JHEP **07**, 012 (2002).

[90] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C28**, 455 (2003).

[91] ZEUS Coll., Chekanov, S. et al., Phys. Rev. **D67**, 012007 (2003).

[92] Whalley, M., *talk at this workshop*.

[93] Martin, A. D. et al., Eur. Phys. J **C23**, 73 (2002).

[94] A. Cooper-Sarkar, C. Gwenlan, *Comparison and combination of zeus and h1 pdf analyses*. These proceedings.

[95] Webb, J. C. and others, *Absolute drell-yan dimuon cross sections in 800-gev/c p p and p d collisions*. Preprint hep-ex/0302019, 2003.

[96] Abbott, T. et al., Phys. Rev. Lett. **86**, 1707 (2001).

[97] Affolder, T. et al., Phys. Rev. **D64**, 032001 (2001).

[98] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Eur. Phys. J. **C35**, 2004 (2004).

[99] Pumplin, J. and Stump, D. R. and Tung, W. K., Phys. Rev. **D65**, 014011 (2002).

[100] Stump, D. et al., Phys. Rev. **D65**, 014012 (2002).

[101] Pumplin, J. et al., Phys. Rev. **D65**, 014013 (2002).

[102] Martin, A. D. and Roberts, R. G. and Stirling, W. J. and Thorne, R. S., Phys. Lett. **B604**, 61 (2004).

[103] Alekhin, S., *Nnlo parton distributions from deep-inelastic scattering data*. Preprint hep-ph/0311184, 2003.

[104] Pumplin, J. et al., *Parton distributions*. Preprint hep-ph/0507093, 2005.

[105] Stump, D. et al., JHEP **10**, 046 (2003).

[106] Forte, S. and Garrido, L. and Latorre, J. I. and Piccione, A., JHEP **05**, 062 (2002).

[107] Del Debbio, L. and Forte, S. and Latorre, J. I. and Piccione, A. and Rojo, J., JHEP **03**, 080 (2005).

[108] Rojo, J. and Latorre, J. I., JHEP **01**, 055 (2004).

[109] Rojo, J. and Del Debbio, L. and Forte, S. and Latorre, J. I. and Piccione, A., *The neural network approach to parton fitting*. Preprint hep-ph/0505044, 2005.

[110] M. Klein, B. Reisert, *Determination of the light quark momentum distributions at low x at hera*. These proceedings.