Benchmarking of parton distributions and their uncertainties

R. D. Ball, L. Del Debbio, J. Feltesse, S. Forte, A. Glazov, A. Guffanti, J. I. Latorre, A. Piccione, V. Radescu, J. Rojo, R. S. Thorne, M. Ubiali, G. Watt

1 Introduction

The proper treatment of uncertainties associated to the fit of Parton Distribution Functions (PDF) has become a subject of great interest in the last few years. A simple way of understanding differences between available approaches to parton fits is to fix some hypothesis (say, experimental data, QCD parameters, input parameterizations, error treatment), and check what is the effect of the remaining assumptions. Such studies were previously done in the framework of the first HERA–LHC workshop [1].

In the following we will discuss three benchmark fits. The first one is presented in Sect. 2. It is based on the H12000 parton fit [2], and it compares a new version of this fit, in which uncertainty bands are determined [3, 4] using a Monte Carlo method, to the reference fit, where uncertainty bands are obtained using the standard Hessian method. The main motivation of this benchmark is to study the impact of possible non-Gaussian behaviour of the data and, more generally, the dependence on the error treatment.

The second benchmark is presented in Sect. 3. It is based on the study performed by S. Alekhin and R. Thorne in Ref. [1], which compared the fits by their respective groups to a common reduced set of data with common assumptions, and also to their respective reference (global) fits. This comparison is extended here in two ways. First, the comparison is extended to include an NNPDF fit to the same reduced set of data with the same assumptions, and the NNPDF1.0 reference fit [5]. Second, results are also compared to a fit based on the recent MSTW 2008 [6, 7] analysis. As in the Thorne benchmark fit, this uses slightly different data sets and assumptions; it is furthermore modified to use the same input parameterization and improved treatment of uncertainties as MSTW. The main purpose of these comparisons is to answer the questions (a) to which extent fit results from various groups obtained using different methodologies still differ from each other when common or similar assumptions and a common or similar reduced dataset are used and (b) how the fits to the reduced dataset by each group compare to the fit to the full dataset.

The third benchmark, discussed in Sect. 4, is a further elaboration on the benchmark presented in Sect. 2, extended to include the NNPDF fit, which also uses a Monte Carlo approach. The main purpose of this benchmark is to compare two fits (H1 and NNPDF) which have the same error treatment but different parton parameterizations. The inclusion in this benchmark of the NNPDF fit is also interesting because it allows a comparison of a fit based on a very consistent set of data coming from the H1 collaboration only, to fits which include all DIS data sets, which are less compatible than the H1 sets alone.

1.1 Settings for the H1 benchmark

This analysis is based on all the DIS inclusive data by the H1 collaboration from the HERA-I run. A kinematic cut of $Q^2 > 3.5\ \mathrm{GeV}^2$ is applied to avoid any higher twist effect. The data

points used in the analysis are summarized in Table 1 and Fig. 1.

Data Set	Data points	Observable	Ref.
H197mb	35	$\tilde{\sigma}^{NC,+}$	[8]
H197lowQ2	80	$\tilde{\sigma}^{NC,+}$	[8]
H197NC	130	$\tilde{\sigma}^{NC,+}$	[9]
H197CC	25	$\tilde{\sigma}^{CC,+}$	[9]
H199NC	126	$\tilde{\sigma}^{NC,-}$	[10]
H199CC	28	$\tilde{\sigma}^{CC,-}$	[10]
H199NChy	13	$\tilde{\sigma}^{NC,-}$	[10]
H100NC	147	$\tilde{\sigma}^{NC,+}$	[2]
H100CC	28	$\tilde{\sigma}^{CC,+}$	[2]
Total	612		

Table 1: Data points used in the H1 benchmark after kinematic cuts of $Q^2 > 3.5\ \mathrm{GeV}^2$.

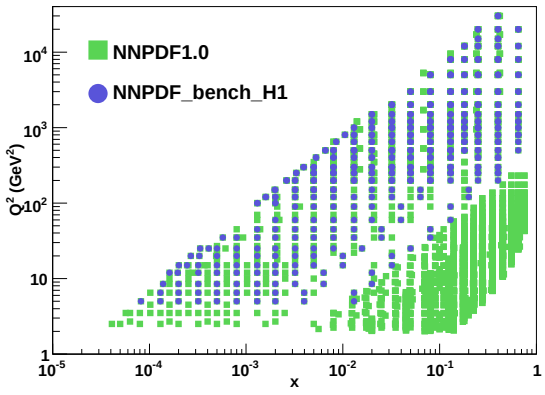

Fig. 1: The data used in the H1 benchmark and in the NNPDF reference fit.

The theoretical assumptions are:

- NLO perturbative QCD in the $\overline{\mathrm{MS}}$ renormalization and factorization scheme;
- zero-mass variable flavour number scheme with quark masses $m_c = 1.4\ \mathrm{GeV}$ and $m_b = 4.5\ \mathrm{GeV}$;
- the strong coupling fixed to $\alpha_s(M_Z) = 0.1185$;
- momentum and valence sum rules enforced;
- starting scale for the evolution at $Q_0^2 = 4\ \mathrm{GeV}^2$;
- strange contribution fixed as

$$s(x,Q_0^2) = \bar{s}(x,Q_0^2) = f_s \bar{D}(x,Q_0^2) = \frac{f_s}{1-f_s}\bar{d}(x,Q_0^2), \tag{1}$$

with $U = u + c$ and $D = d + s + b$ and with $f_s = 0.33$;

- charm contribution fixed as

$$c(x, Q_0^2) = \bar{c}(x, Q_0^2) = f_c \bar{U}(x, Q_0^2) = \frac{f_c}{1 - f_c} \bar{u}(x, Q_0^2), \tag{2}$$

with $f_c = 0.15$;

- five independent PDFs: gluon and U, D, $\bar{U}$, $\bar{D}$ (see definition above);
- iterated solution for evolution (see, e.g. [11], Sect. 1.3).

Both the H1 and NNPDF methodologies are based on

- Monte Carlo method to determine uncertainties. This method will be discussed in detail in Sect. 2.2 below.

They differ in the way PDFs are parameterized:

- H1 parameterizes PDFs as

$$\begin{aligned} xg(x, Q_0^2) &= A_g x^{B_g} (1 - x)^{C_g} [1 + D_g x], \\ xU(x, Q_0^2) &= A_U x^{B_U} (1 - x)^{C_U} [1 + D_U x + F_U x^3], \\ xD(x, Q_0^2) &= A_D x^{B_D} (1 - x)^{C_D} [1 + D_D x], \\ x\bar{U}(x, Q_0^2) &= A_{\bar{U}} x^{B_{\bar{U}}} (1 - x)^{C_{\bar{U}}}, \\ x\bar{D}(x, Q_0^2) &= A_{\bar{U}} x^{B_{\bar{D}}} (1 - x)^{C_{\bar{D}}}, \end{aligned} \tag{3}$$

(4)

which yields 10 free parameters after sum rules are imposed;

- NNPDF parameterizes PDFs with a 2-5-3-1 neural network, which implies 185 free parameters to be fitted.

Because of the large number of parameters, the minimum of the NNPDF fit is determined using the stopping criterion discussed in Sect. 3.2 below, while the minimum of the H1 fit is determined as the standard minimum χ^2 (or maximum likelihood) point of parameter space.

1.2 Settings for the HERA–LHC benchmark

This benchmark was first presented in Ref. [1], where its settings were defined. In order to have a conservative ensemble of experimental data and observables, only structure function DIS data are used. Large kinematic cuts are applied to avoid any higher twist effect. The data points used in the Alekhin analysis are summarized in Table 2 and Fig. 2.

The theoretical assumptions are:

- NLO perturbative QCD in the $\overline{\rm MS}$ renormalization and factorization scheme;
- zero-mass variable flavour number scheme with quark masses $m_c = 1.5$ GeV and $m_b = 4.5$ GeV;
- $\alpha_s(M_Z)$ fitted: the best-fit values are 0.1110 ± 0.0012 (Alekhin) and 0.1132 ± 0.0015 (Thorne);
- momentum and valence sum rules imposed;

Data Set	Data points	Observable	Ref.
ZEUS97	206	F_2^p	[12]
H1lowx97	77	F_2^p	[8]
NMC	95	F_2^p	[13]
NMC_pd	73	F_2^d/F_2^p	[14]
BCDMS	322	F_2^p	[15]
Total	773		

Table 2: Data points used in the HERA–LHC benchmark after kinematic cuts of $Q^2 > 9\ \mathrm{GeV}^2$ and $W^2 > 15\ \mathrm{GeV}^2$ are applied.

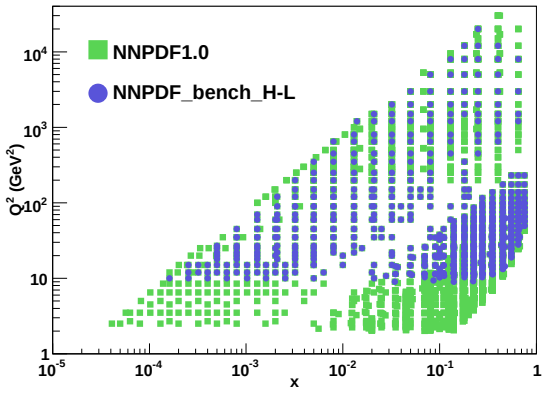

Fig. 2: The data used in the HERA–LHC benchmark and in the NNPDF reference fit.

- starting scale for evolution $Q_0^2 = 1\ \mathrm{GeV}^2$;
- four independent input PDFs (u and d valence, the sea and the gluon);
- no light sea asymmetry: $\bar{u} = \bar{d}$;
- no independent strange PDF:

$$s(x,Q_0^2) + \bar{s}(x,Q_0^2) = 0.5(\bar{u}(x,Q_0^2) + \bar{d}(x,Q_0^2))\,; \tag{5}$$

- iterated solution of evolution equations;

The NNPDF analysis presented here is based on the same data set and theoretical assumptions, the only difference being that the strong coupling is fixed to $\alpha_s(M_Z) = 0.112$, i.e. the average of the fitted values of S. Alekhin and R. Thorne.

The Thorne benchmark used somewhat different data sets and assumptions. Namely:

- A somewhat different dataset is used, as displayed in Table 3. This differs from the dataset of Table 2 and Figure 2 because the NMC and BCDMS fixed-target data on F_2^p used are averaged over different beam energies, and also, HERA reduced cross sections rather than structure function data are used, resulting in an additional nine H1 points. Note that the Thorne benchmark in Ref. [1] also included the F_2^d BCDMS deuterium data.

Data Set	Data points	Observable	Ref.
ZEUS97	206	$\tilde{\sigma}^{\mathrm{NC},+}$	[12]
H1lowx97	86	$\tilde{\sigma}^{\mathrm{NC},+}$	[8]
NMC	67	F_2^p	[13]
NMC_pd	73	F_2^d/F_2^p	[14]
BCDMS	157	F_2^p	[15]
Total	589		

Table 3: Data points used in the MSTW benchmark fit after kinematic cuts of $Q^2 > 9\ \mathrm{GeV}^2$ and $W^2 > 15\ \mathrm{GeV}^2$ are applied.

- All correlations between systematics are neglected, and statistical and systematic errors are added in quadrature.
- Normalizations of individual data sets are fitted with a rescaling of uncertainties to avoid systematic bias.
- The F_2^d/F_2^p data are corrected for nuclear shadowing effects [16].

The MSTW analysis presented here makes the same choices as the Thorne benchmark, but with $\alpha_s(M_Z) = 0.112$, and additionally

- a global correction of -3.4% is applied to the luminosity of the published H1 MB 97 data [8] following a luminosity reanalysis [17].
- a quartic penalty term in the χ^2 definition is given to normalizations which deviate from the central value.

2 Experimental Error Propagation[1]

2.1 Introduction

Standard error estimation of proton parton distribution functions (PDFs) relies on the assumption that all errors follow Gaussian (or normal) statistics. However, this assumption may not always be correct. Some systematic uncertainties such as luminosity and detector acceptance follow rather a log-normal distribution (see Section [18]). Compared to the Gaussian case, the lognormal distribution which has the same mean and root mean square (RMS), is asymmetric and has a shifted peak, as shown illustratively in Figure 3. Therefore, the non-Gaussian behaviour of the experimental uncertainties could lead to an additional uncertainty of the resulting PDFs. An alternative to the standard error propagation is a toy Monte Carlo (MC) method. Here, an implementation of the MC method is presented for estimation of the PDF uncertainties with various assumptions for the error distribution. In addition, this MC method provides an independent cross check of the standard error propagation when assuming the Gaussian error distributions.

2.2 Method

The Monte Carlo technique consists firstly in preparing replicas of the initial data sets which have the central value of the cross sections, σ_i, fluctuating within its systematic and statistical uncer-

[1]Contributing authors: J. Feltesse, A. Glazov, V. Radescu

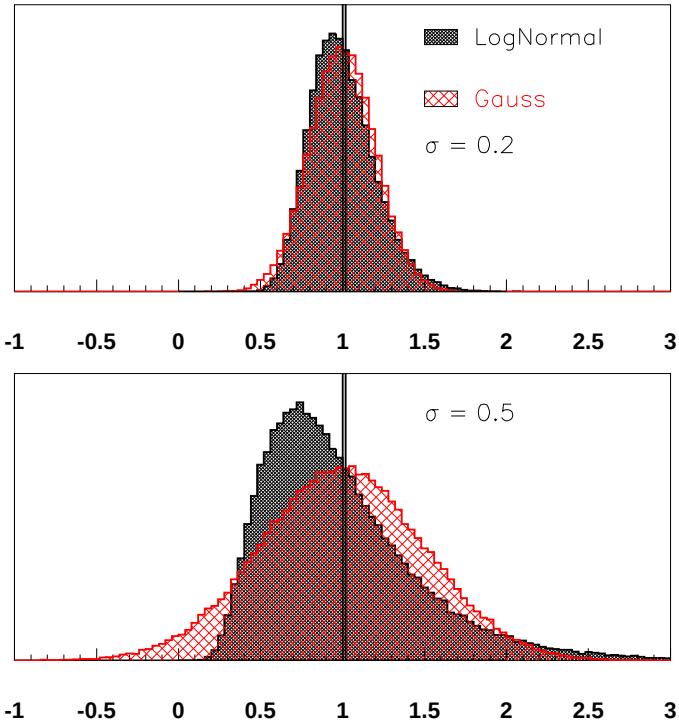

Fig. 3: Comparison of the lognormal (black, darker hatching) and Gaussian (red, lighter hatching) probability distributions. The distributions are shown with mean equal to one, and two different choices for the RMS (for both distribution): $\sigma = 0.2$ (top) and $\sigma = 0.5$.

tainties taking into account all point to point correlations. Various assumptions can be considered for the error distributions. When dealing with the statistical and point to point uncorrelated errors, one could allow each data point to randomly fluctuate within its uncorrelated uncertainty assuming either Gauss, lognormal, or any other desired form of the error distribution. For example, for Gaussian errors

$$\sigma_i \longrightarrow \sigma_i \left(1 + \delta_i^{uncorr} \cdot R_i\right), \tag{6}$$

where δ_i^{uncorr} corresponds to the uncorrelated uncertainties and R_i is a random number chosen from a normal distribution with a mean of 0 and a standard deviation of 1. Hence, the central value of each cross section point i is shifted by $\delta_i^{uncorr} \cdot R_i$.

For the systematic errors, the treatment is a bit more complicated than above. This is due to the correlation between data points and that, in general, the data points are sensitive to the systematic sources with a different strength δ_{ij}, where index i (j) runs over all the cross section points (all systematic sources). In order to take this into account, for each systematic source j a uniformly distributed *fluctuation probability* P_j is selected. Then, for each data point i the central value of cross section is shifted such that probability of this shift, which depends on δ_{ij} and the exact form of the probability distribution function, is equal P_j (for positive δ_{ij}) or $(1 - P_j)$ (for negative δ_{ij}). In other words, each central value of the cross section is shifted with

the same probability of the corresponding systematic shift. For example for the Gaussian errors, this procedure is equivalent to

$$\sigma_i \longrightarrow \sigma_i \left(1 + \delta_i^{uncorr} \cdot R_i + \sum_j^{N_{sys}} \delta_{ij}^{corr} \cdot R_j \right), \tag{7}$$

where in addition to the shifts for the uncorrelated errors previously explained, R_j corresponds to another random number chosen from a normal distribution with mean of 0 and standard deviation of 1 as a fluctuation for the systematic source j. Hence, the central values of the cross sections are shifted in addition by $\delta_{ij}^{corr} \cdot R_j$ for each systematic shift.

This preparation of the data is repeated for N times, where high statistics is desirable for more accurate results. For this study we used $N > 100$ which proved to suffice. For each replica, a next to leading order (NLO) QCD fit is performed to extract the PDFs. The errors on the PDFs are estimated from the RMS of the spread of the N lines corresponding to the N individual fits to extract PDF.

A fit to the published H1 HERA-I data of neutral and charged current $e^{\pm}p$ scattering cross sections [2] using the settings discussed in Sect. 1.1 has been performed, using the QCDNUM program [19].

2.3 Validation of the Method

The MC method is tested by comparing the standard error estimation of the PDF uncertainties with the MC techniques by assuming that all the errors (statistical and systematic) follow Gaussian (normal) distribution. Figure 4 shows good agreement between the methods.

2.4 Test of various assumptions for the error distributions

Two cases are considered which may represent most likely the error distributions: (1) the lognormal distribution for the luminosity uncertainty and the rest of the errors are set to follow the Gauss shape, (2) the lognormal distributions for all the systematic errors and the statistical errors are set to follow the Gauss distributions. The results for the first case (1) are shown in Figure 5. The results of the tests for the case when lognormal distributions for all the systematic uncertainties are assumed is shown in Figure 5. We observe that for the precise H1 HERA-1 data the effect of using lognormal distribution, which is considered for some systematic uncertainties more physical, is similar to the pure gauss distribution case.

2.5 Conclusions

A simple method to estimate PDF uncertainties has been built within QCD Fit framework. Assuming only gauss distribution of all errors, the results agree well with the standard error estimation. This method allows to check the effect of non- gauss assumptions for distributions of the experimental uncertainties. For the H1 data, results are similar to the gauss case when using lognormal. The method could be extended for other physical variables (i.e. cross sections) for cross checks with the standard error evaluation.

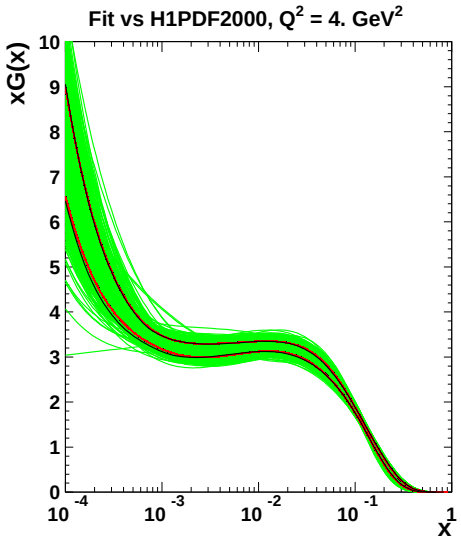

Fig. 4: Comparison between the standard error calculations and the Gauss error distribution is shown for the gluon PDF. Green lines represent the spread of Monte Carlo generated allowances for the errors, and the red lines are the RMS of this spread. The black lines correspond to the standard error calculations of the PDF errors.

3 HERA–LHC Benchmark

This benchmark is based on the Alekhin/Thorne benchmark of Ref. [1], whose settings has been given in Sect. 1.2. Both the Alekhin and Thorne fits had the following features:

- uncertainties determined using the Hessian method with $\Delta\chi^2 = 1$;
- input PDFs are parameterized using the following functional form:

$$x\, f_i(x, Q_0^2) = A_i(1-x)^{b_i}(1 + \epsilon_i x^{0.5} + \gamma_i x)x^{a_i}\,. \tag{8}$$

  with ϵ_i and γ_i set to zero for the sea and gluon distributions. Hence, there were a total of 13 free PDF parameters plus $\alpha_s(M_Z)$ after imposing sum rules.

Here, we reanalyze it within the MSTW and NNPDF approaches. First, we summarize the respective MSTW and NNPDF approaches, and especially their differences when compared to the previous HERALHC benchmark fits of Ref. [1]. Then, results for benchmark fits obtained with the various different approaches are compared to each other. Finally, we compare each benchmark fit to its counterpart based on a wider range of data, i.e. the NNPDF1.0 [5] reference and the MRST01 [20] and MSTW08 [6, 7] PDFs.

3.1 MSTW approach[2]

The benchmark analysis is now much more closely aligned to the global analysis than was the case for the Thorne benchmark compared to the MRST global analysis. It follows the general approach taken by the MRST (or more recently, MSTW) group, and is similar to that described in Ref. [20]. There are some new features which are explained below.

[2]Contributing authors: R. S. Thorne, G. Watt

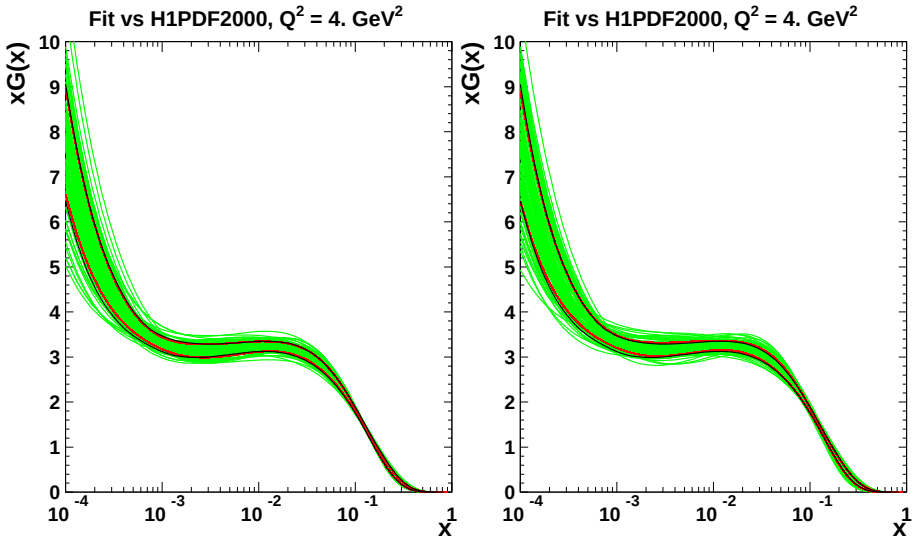

Fig. 5: Comparison between errors on PDFs obtained via standard error calculation (black) where Gauss assumption is used, and errors obtained via Monte Carlo method (red) where luminosity uncertainty is allowed to fluctuate according to lognormal distributions and all the other uncertainties follow the Gaussian distribution (left), and where all the systematic uncertainties are allowed to fluctuate according to lognormal distributions (right). Only the gluon PDF is shown, where the errors are larger. The green lines show the spread of the N individual fits.

- *Input parameterization.* We take the input PDF parameterization at $Q_0^2 = 1$ GeV^2 to be:

$$xu_v(x,Q_0^2) = A_u\, x^{\eta_1}(1-x)^{\eta_2}(1+\epsilon_u\sqrt{x}+\gamma_u\, x)\,, \quad (9)$$

$$xd_v(x,Q_0^2) = A_d\, x^{\eta_3}(1-x)^{\eta_4}(1+\epsilon_d\sqrt{x}+\gamma_d\, x)\,, \quad (10)$$

$$xS(x,Q_0^2) = A_S\, x^{\delta_S}(1-x)^{\eta_S}(1+\epsilon_S\sqrt{x}+\gamma_S\, x)\,, \quad (11)$$

$$xg(x,Q_0^2) = A_g\, x^{\delta_g}(1-x)^{\eta_g}(1+\epsilon_g\sqrt{x}+\gamma_g\, x) + A_{g'}\, x^{\delta_{g'}}(1-x)^{\eta_{g'}}\,, \quad (12)$$

where $S = 2(\bar{u}+\bar{d}+\bar{s})$, $s = \bar{s} = 0.1\,S$ and $\bar{d} = \bar{u}$. The parameters A_u, A_d and A_g are fixed by sum rules, leaving potentially 19 free parameters. In practice, to reduce the number of highly correlated parameters, making linear error propagation unreliable, we determine the central value of the benchmark fit by freeing all 19 parameters, then fix 6 of those at the best-fit values when calculating the Hessian matrix used to determine the PDF uncertainties, giving a total of 13 eigenvectors. This is the same procedure as used in the MSTW 2008 global fit [6, 7], where there are an additional 3 free parameters associated with $\bar{d}-\bar{u}$ and an additional 4 free parameters associated with strangeness, giving a total of 20 eigenvectors. Note that the parameterization used in the previous Alekhin/Thorne benchmark fits was considerably more restrictive, where the ϵ_S, γ_S, ϵ_g and γ_g parameters were set to zero, and the second (negative) gluon term was omitted entirely. In addition, ϵ_u was held fixed for the Thorne benchmark fit, leaving a total of 12 eigenvectors. We find that the more flexible gluon parameterization, allowing it to go negative at very small x, is very highly correlated with the value obtained for α_s, and a value of $\alpha_s(M_Z) = 0.105$ is

obtained if it is allowed to go free at the same time as the other parameters, therefore we instead choose to fix it at $\alpha_s(M_Z) = 0.112$ as in the NNPDF benchmark fit.

- *Error propagation.* Apart from the more flexible input parameterization, the other major difference in the new MSTW version of the HERA–LHC benchmark fit, with respect to the previous Thorne (MRST) version, is the choice of tolerance, $T = \sqrt{\Delta\chi^2}$. The MRST benchmark fit used the standard choice $T = 1$ for one-sigma uncertainties. More precisely, the distance t along each normalized eigenvector direction was taken to be 1, and ideal quadratic behaviour about the minimum was assumed, giving $T \approx t = 1$. The MRST global fit used $T = \sqrt{50}$ for a 90% confidence level (C.L.) uncertainty band; however, this is not appropriate when fitting a smaller number of data sets. Recently, a new procedure has been developed [6, 7] which enables a *dynamic* determination of the tolerance for each eigenvector direction, by demanding that each data set must be described within its one-sigma (or 90%) C.L. limits according to a hypothesis-testing criterion, after rescaling the χ^2 for each data set so that the value at the global minimum corresponds to the most probable value. Application of this procedure to the MSTW benchmark fit gives $T \sim 3$ for one-sigma uncertainties and $T \sim 5$ for 90% C.L. uncertainties. For the MSTW global fit, the typical values of T required are slightly larger, with more variation between different eigenvector directions. The increase in T in the global fit is mainly due to the inclusion of some less compatible data sets, while the greater variation in T between eigenvectors is due to the fact that some parameters, particularly those associated with s and $\bar{s}$, are constrained by far fewer data sets than others. In the MSTW fits, the data set normalizations are allowed to vary, with the aforementioned penalty term, when determining the PDF uncertainties. For global fits this automatically leads to a small increase in uncertainty compared to the MRST determinations, where data set normalisations were held fixed when calculating the Hessian matrix used for error propagation. In the MRST benchmark fit the data set normalizations were allowed to vary. To calculate the uncertainty bands from the eigenvector PDF sets, we use the formula for asymmetric errors given, for example, in Eq. (13) of Ref. [20].

3.2 NNPDF approach[3]

The NNPDF approach was proposed in Ref. [21], and it was applied there and in Ref. [22] to the parameterization of the structure function $F_2(x, Q^2)$ with only two or more experimental data sets respectively. In Ref. [23] it was first used for the determination of a single PDF (the isotriplet quark distribution), and in Ref. [5] a full set of PDFs fit based on DIS data (NNPDF1.0) was presented. Because the method has been discussed extensively in these references, here we only summarize briefly its main features.

- *Error propagation.* We make a Monte Carlo sample of the probability distribution of the experimental data by generating an ensemble of N replicas of artificial data following a multi-gaussian distribution centered on each data point with full inclusion of the experimental covariance matrix. Each replica is used to construct a set of PDFs, thereby propagating the statistical properties of the data Monte Carlo sample to a final Monte Carlo

[3]Contributing authors: R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, A. Piccione, J. Rojo, M. Ubiali

sample of PDFs. Here we shall take $N = 100$. The method is the same as discussed in Sect. 2.2, the only difference being the treatment of normalization errors: relative normalizations are fitted in the H1 approach, while they are included among the systematic errors in the Monte Carlo data generation in the NNPDF approach (see Refs. [2, 5] for details of the respective procedures) .

- *Input parameterization*. Each PDF is parameterized with a functional form provided by a neural network. The architecture for the neural network is the same for all PDFs, and yields a parameterization with 37 free parameters for each PDF. This is a very redundant parameterization, it is chosen in order to avoid parameterization bias; neural networks are a particularly convenient way of dealing with redundant parameterizations. Note that sum rules are also imposed.
- *Minimization*. A redundant parameterization allows for fitting not only the underlying physical behaviour, but also statistical noise. Therefore, the minimization is stopped not at the absolute minimum of the χ^2, but rather before one starts fitting noise. This optimal stopping point is determined as follows: the data in each replica are randomly assigned either to a training or to a validation set. The fit is performed on data of the training set only, while the validation set is used as a monitor. The fit is stopped when the quality of the fit to the training set keeps improving, but the quality of the fit to the validation set deteriorates.

3.3 Comparison between the Benchmark Parton Distributions

Data Set	$\chi^2_{\rm bench}/N_{\rm data}$	$\chi^2_{\rm global}/N_{\rm data}$
ZEUS97	1.09	1.18
H1lowx97	1.03	1.00
NMC	1.40	1.45
NMC_pd	1.24	1.32
BCDMS	1.21	1.98
Total	1.19	1.53

Table 4: NNPDF χ^2 for the total and each single data set, both for the benchmark and global fit.

The χ^2 per data point for the NNPDF and MSTW fits are shown in Table 4 and 5 respectively. Note that in the MSTW fit statistical and systematic errors are added in quadrature, so the quantity shown is the diagonal contribution to the χ^2. The quality of the NNPDF is seen to be uniformly good. The quality of the MSTW is also uniform, though it cannot be compared directly because of the different way systematics are treated. The comparison of each benchmark fit to the corresponding global fit will be discussed in Sect. 3.4 below.

In Fig. 6 the PDFs from the NNPDF and MSTW benchmark fits presented here are compared to those by Thorne from Ref. [1] at the same reference scale of $Q^2 = 20\ {\rm GeV}^2$ used there (denoted as MRST01 in the figure). The benchmark fit by Alekhin [1] is not shown as the PDFs are very close to the those by Thorne displayed in Fig. 6.

Data set	$\chi^{2\,\mathrm{diag}}_{\mathrm{bench}}/N_{\mathrm{data}}$	$\chi^{2\,\mathrm{diag}}_{\mathrm{global}}/N_{\mathrm{data}}$
ZEUS97	0.76	0.79
H1lowx97	0.53	0.54
NMC	1.08	1.11
NMC_pd	0.78	0.89
BCDMS	0.74	1.13
Total	0.76	0.89

Table 5: MSTW χ^2 for the total and each single data set, both for the benchmark and global fit. Notice that statistical and systematic errors are added in quadrature and that relative data set normalizations are fitted.

For PDFs and kinematical regions where data are available, namely the small-x gluon and sea quark and the large-x u_v distributions, the central values of the NNPDF fit are quite close to those of the MRST and MSTW fits, despite the differences in methodology. The central values of the PDFs are slightly different for the MRST and MSTW benchmark fits due to the use of BCDMS F_2^d data in the former, which affects mainly valence quarks. Where extrapolation is needed, such as for the d_v distribution, which is constrained only by the small amount of data on the ratio F_2^d/F_2^p, or the large-x sea quark, central values are rather more different (though the Alekhin/MRST/MSTW benchmark central values are within the NNPDF error band). The exception is the smallest-x gluon, where the form of the MSTW parameterization results in a very sharp turn-over. However, even here the uncertainty bands are close to overlapping.

Differences are sizeable in the estimation of uncertainties. Firstly, uncertainty bands for NNPDF benchmark are significantly larger than for the MSTW benchmark, which in turn are in general somewhat larger than those for the MRST benchmark. The difference between MRST and MSTW, which are based on similar methodology, is due to use of a dynamic tolerance and a more flexible gluon parameterization in MSTW (see Sect. 3.1). Secondly, the width of the uncertainty band for NNPDF benchmark varies rather more than that of the MRST benchmark according to the PDF and the kinematic region, though this is not quite so much the case comparing to MSTW benchmark. Indeed, the NNPDF uncertainties are quite small in the region between $x = 0.01$ and $x = 0.1$ (where there is the bulk of HERA and fixed-target data), while they blow up in the large-x region for the sea quark or the small-x gluon, where there is less or no experimental information. The smallness of the uncertainty band for MSTW for the small-x valence quarks may be partially due to the lack of flexibility in the parameterization: note that because of sum rules, the size of uncertainties in the data and extrapolation region are correlated.

Finally, the MRST/MSTW central value generally falls within the NNPDF uncertainty band, but the NNPDF central value tends to fall outside the MRST/MSTW uncertainty band whenever the central values differ significantly.

3.4 Comparison of the Benchmark Parton Distributions and Global Fits

In Fig. 7 we compare the NNPDF benchmark fit to the NNPDF1.0 reference fit of Ref. [5] (NNPDF global, henceforth), while in Fig. 8 we compare the MSTW benchmark fit to the

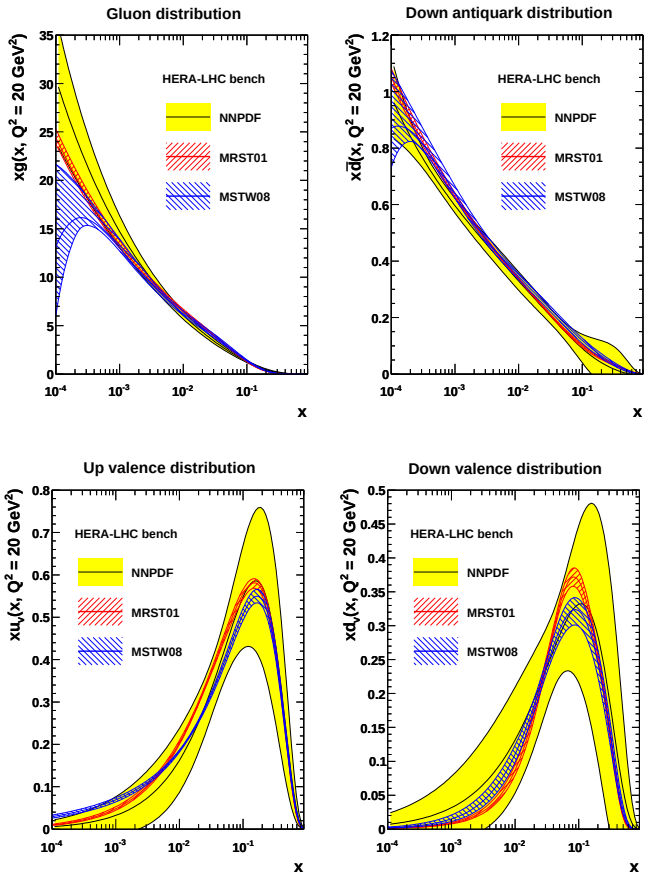

Fig. 6: Comparison of the NNPDF, MRST and MSTW benchmark fits for the gluon, d-sea, u-valence and d-valence at $Q^2 = 20$ GeV^2. All uncertainties shown correspond to one–σ bands.

MRST01 [20] (MRST global, henceforth) and MSTW08 [6, 7] global fits (MSTW global, henceforth).

The χ^2 of the NNPDF benchmark and global fits are compared in Table 4, while those of the MSTW benchmark and global fits are compared in Table 5. Note that for the NNPDF fits the χ^2 is computed using the full covariance matrix, while for the MSTW fits systematic and statistical uncertainties are added in quadrature. Note also that the MRST and MSTW global fits are carried out in a general-mass variable flavour number scheme rather than the zero-mass variable flavour number scheme used in the corresponding benchmark fits, whereas for NNPDF both global and benchmark fits are done with a zero-mass variable flavour number scheme. Comparison of the quality of each benchmark to the corresponding global fit to the same points in Table 5 shows a significant deterioration in the quality of the fit (total $\Delta\chi^2 \gg 1$), especially for the BCDMS F_2^p data. All fits appear to be acceptable for all data sets: for instance, even though the χ^2 of the NNPDF global fit for the benchmark subset of data is 1.98, it is equal to 1.59 [5] for the full BCDMS set of data. However, the increase in χ^2 suggests that there might be data inconsistencies.

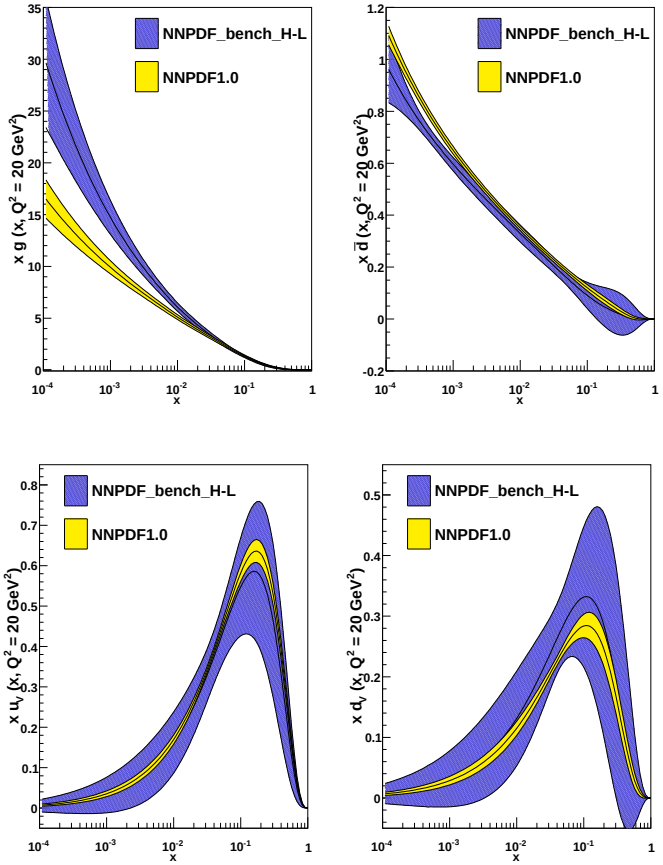

Fig. 7: Comparison of the NNPDF benchmark and reference fits for the gluon, d-sea, u-valence and d-valence at $Q^2 = 20\ \mathrm{GeV}^2$.

Let us now compare each pair of benchmark and global fits. For NNPDF, the difference in central value between benchmark and reference is comparable to that found between the MRST or Alekhin global fits and their benchmark counterparts in Ref. [1]. However, the NNPDF global and benchmark fits remain compatible within their respective error bands. Indeed, the NNPDF benchmark fit has a rather larger error band than the reference, as one would expect from a fit based on a rather smaller set of (compatible) data. Such a behaviour was however not observed in the comparison between global and benchmark MRST and Alekhin fits of Ref. [1].

It is interesting to observe that the gluon shape at low x of the benchmark and global NNPDF disagree at the one σ level (though they agree at two σ). This can be understood as a consequence of the fact that the value of α_s in the two fits is sizably different ($\alpha_s = 0.112$ vs. $\alpha_s = 0.119$). Theoretical uncertainties related to the value of α_s were shown in Ref. [5] to be negligible and thus not included in the NNPDF error band, but of course they become relevant if α_s is varied by several standard deviations (3.5 σ, in this case).

Coming now to MSTW, we first notice that, as discussed in Sect. 3.3, the MSTW bench-

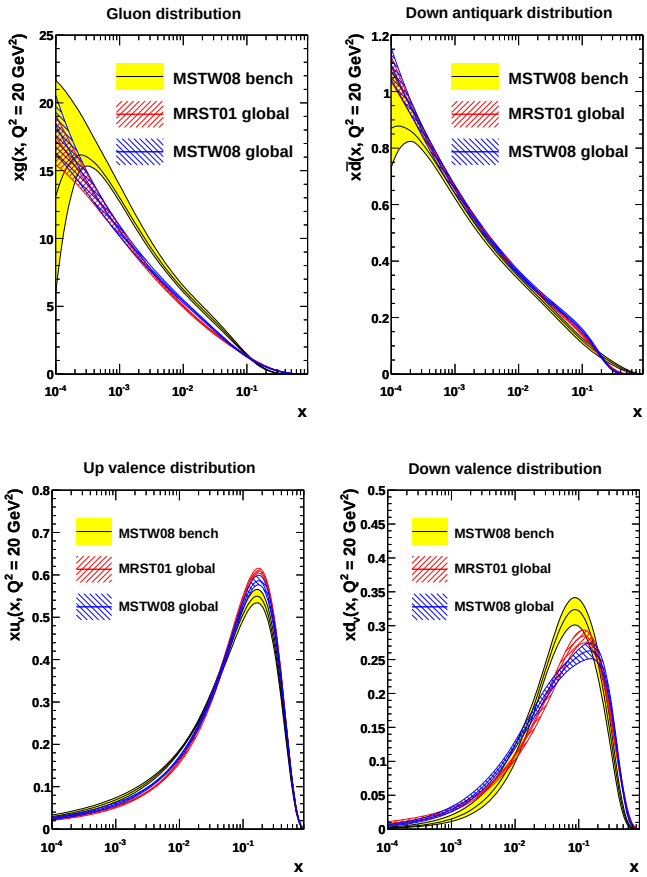

Fig. 8: Comparison of the MSTW benchmark and MRST/MSTW global fits for the gluon, d-sea, u-valence and d-valence at $Q^2 = 20\ \mathrm{GeV}^2$. All uncertainties shown correspond to one–σ bands.

mark set has somewhat larger uncertainty bands than the MRST benchmark set and thus also than each of the sets obtained from global fits. Consequently, the MSTW benchmark PDFs are generally far more consistent with the MSTW global fit sets than the corresponding comparison between MRST benchmark PDFs and global fit PDFs shown in Ref. [1], largely due to the more realistic uncertainties in the MSTW benchmark. Comparing central values we see exactly the same feature in the gluon distribution as the NNPDF group, and the explanation is likewise the same, highlighting possible difficulties in comparing PDFs obtained with different values of $\alpha_s(M_Z)$.

Unlike for the NNPDF group, the MSTW group sees some degree of incompatibility between the benchmark PDFs and the global fit PDFs for the valence quarks, particularly in the case of the down valence. This may be related to the assumption $\bar{u} = \bar{d}$, which constrains valence quarks and sea quarks in an artificial manner since there is less flexibility to alter each independently. Indeed, in the global fits there is an excess of $\bar{d}$ over $\bar{u}$ which maximizes at $x = 0.1$. Forcing equivalence of antiquark distributions might therefore lead to a deficit of down sea quarks

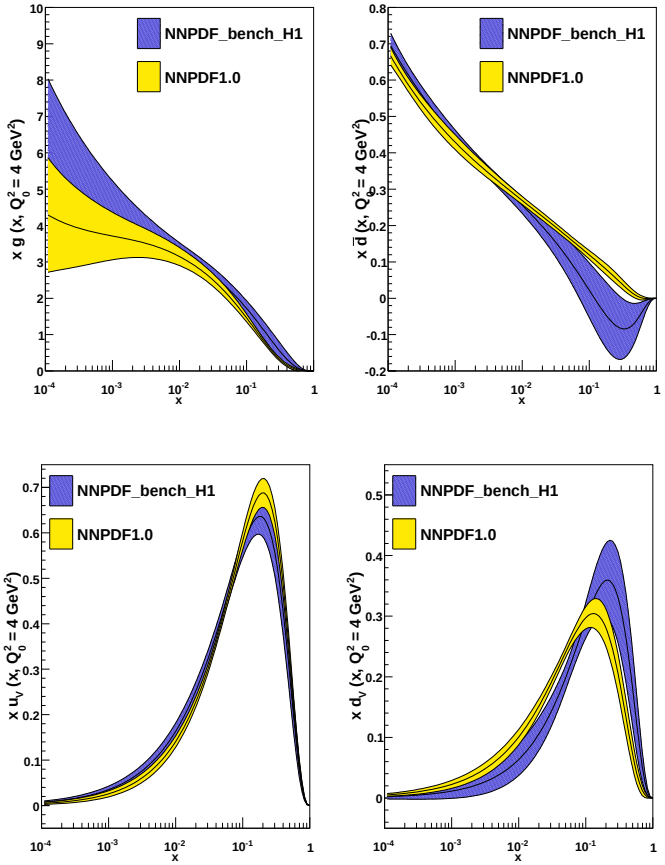

Fig. 9: Comparison of the NNPDF benchmark and reference fits for the gluon, d-sea, u_v and d_v at $Q^2 = 4$ GeV^2.

and a corresponding excess of up sea quarks, and also, for the same reason, to an excess of down valence quarks. These are indeed seen both in the NNPDF and MSTW benchmark fits when compared to the respective global fits. The effect is however well within the uncertainty bands for NNPDF, which indeed do not observe any statistically significant difference between results of a fit to the reduced benchmark data set with the $\bar{u} = \bar{d}$ assumption (as presented in Fig. 7) or without it (as presented in Ref. [5], Fig. 12).

As well as this important effect one sees that the main discrepancy at $x = 0.1$ for down valence quarks is greater when comparing the benchmark fits to the global MSTW fit than to the global MRST fit. This is because recent new Tevatron data on Z rapidity distributions and lepton asymmetry from W decays provide a strong constraint on the down quark, and some of this new data shows considerable tension with other data sets.

4 H1 Benchmark

We now discuss the extension of the fit using the settings of Sect. 1.1 to also include the NNPDF approach. Results are compared both to those of the NNPDF reference fit, and to those obtained by the H1 fit of Sect. 2 to the same data. We then compare the NNPDF benchmark and reference, with the specific aim of addressing the issue of the dependence of the results on the size of the data set (H1 dataset vs. the HERA–LHC dataset of Sect. 3). Finally, the H1 and NNPDF benchmark fits are compared to each other with the purpose of understanding the impact of the respective methodologies.

4.1 NNPDF analysis[4]

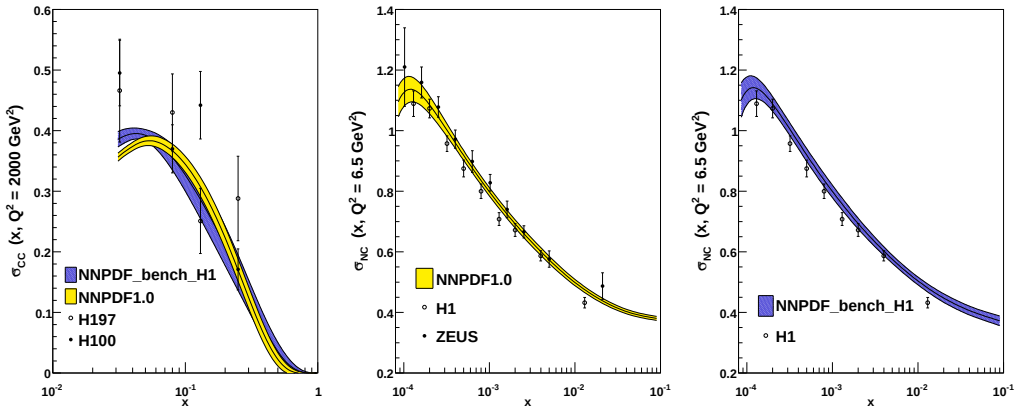

Fig. 10: Left: NNPDF benchmark and reference fits at $\sqrt{s} = 301\text{GeV}$ compared to H1 charged current data. Center: NNPDF reference fit compared to H1 and ZEUS neutral current data. Right: NNPDF benchmark fit compared to H1 neutral current data.

The results of the NNPDF benchmark are compared to the NNPDF reference fit results in Fig. 9. The general features of the benchmark are analogous to those of the HERA–LHC benchmark discussed in Section 3.4, with some effects being more pronounced because the benchmark dataset is now even smaller. Specifically, we observe that uncertainties bands blow up when data are removed: this is very clear for instance in the $\bar{d}$ distribution at large-x, as a consequence of the fact that the benchmark dataset of Table 1 does not include deuterium data. The negative value of this PDF at large x is presumably unphysical and it would disappear if positivity of charged current cross sections were imposed, including also the (anti-)neutrino ones. The only positivity constraint in the NNPDF fit is imposed on the F_L structure function [5], because this is the only DIS observable whose positivity is not constrained by the full data set.

It is interesting to note however that this effect is not observed for the u_v distribution, where instead the benchmark and the reference fit show almost equal uncertainties. In order to

[4]Contributing authors R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, A. Piccione, J. Rojo, M. Ubiali

understand this, in Fig. 10 we compare two situations with or without error shrinking, by examining the predictions obtained using the benchmark and reference fits for some observables to the corresponding data. A first plot (left) shows the shrinking of the uncertainty on the prediction for the charged–current cross section in the reference fit. This is mostly due to the CHORUS neutrino data, which are in the reference and not in the benchmark. These data are clearly consistent with the H1 data shown in the plot. The subsequent pair of plots compares (center) the prediction for the neutral–current cross section from the reference fit compared to H1 and ZEUS data (both of which are used for the reference fit), and (right) from the benchmark fit to the H1 data only (which are the only ones used in the benchmark fit). The uncertainty bands in the two fits are similar size: indeed, the ZEUS and H1 data display a systematic disagreement which is approximately the size of this uncertainty band. Hence, the (small but significant) systematic inconsistency between the ZEUS and H1 data prevents reduction of the uncertainty band when the ZEUS data are added to the fit, beyond the size of this discrepancy. Therefore, the NNPDF methodology leads to combined uncertainties for inconsistent data which are similar to those obtained with the so–called PDG (or scale-factor) method [24].

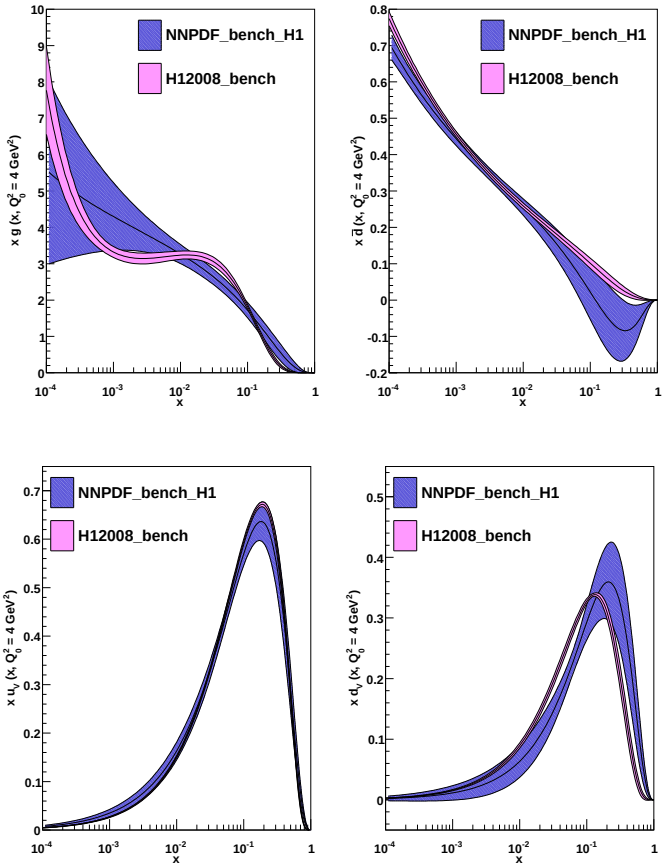

Fig. 11: Comparison of the NNPDF and H1 benchmark fit for the gluon, d-sea, u_v and d_v at $Q^2 = 4\ \mathrm{GeV}^2$.

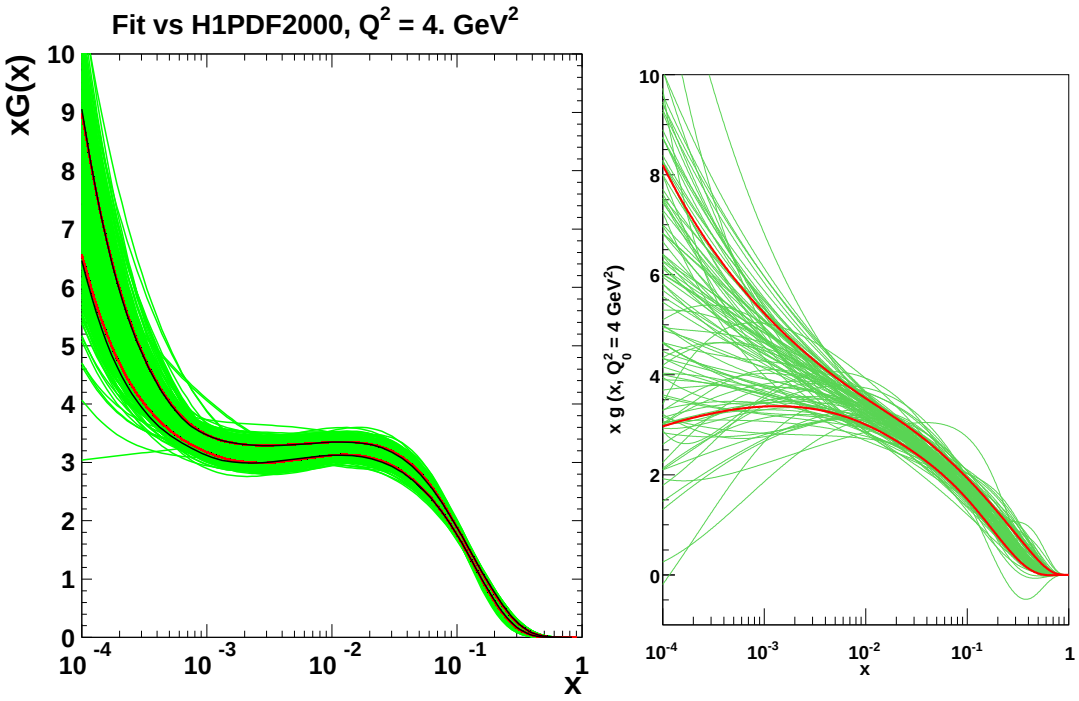

Fig. 12: The Monte Carlo set of gluon PDFs for the H1 benchmark (left, same as Fig. 4) and the NNPDF benchmark. The red lines show the one-sigma contour calculated from the Monte Carlo set, and in the H1 case the black lines show the Hessian one-sigma contour.

Data Set	$\chi^2_{\rm H1}/N_{\rm data}$	$\chi^2_{\rm NNPDF}/N_{\rm data}$
H197mb	0.83	0.82
H197lowQ2	0.90	0.87
H197NC	0.69	0.80
H197CC	0.73	0.97
H199NC	0.88	1.01
H199CC	0.62	0.84
H199NChy	0.35	0.35
H100NC	0.97	1.00
H100CC	1.07	1.38
Total	0.88	0.96

Table 6: H1 and NNPDF χ^2 for the total and each single data set. Cross correlations among data sets are neglected to evaluate the χ^2 of a single data set.

Notice that if relative normalization are fitted (as done by in the H1 approach of Sect. 2) instead of being treated simply as a source of systematics, this systematic inconsistency would be significantly reduced in the best-fit. The associate uncertainty however then appears as an

addition source of systematics. This happens when H1 and ZEUS data are combined in a single dataset (see Section [18] below). In the NNPDF approach, instead, this systematics is produced by the Monte Carlo procedure.

4.2 Comparison between the Benchmark Parton Distributions

The χ^2 of the H1 and NNPDF benchmarks are given in Table 6, while the corresponding PDFs are compared in Fig. 11. Furthermore, in Fig. 12 we show the respective full Monte Carlo PDF sets in the case of the gluon distribution.

The quality of the two fits is comparable, the differences in χ^2 being compatible with statistical fluctuations. In the region where experimental information is mostly concentrated, specifically for the u_v distribution over all the x-range and for the $\bar{d}$ and the d_v distributions in the small-x range, the results of the two fits are in good agreement, though the H1 uncertainty bands are generally somewhat smaller.

In the region where experimental information is scarce or missing, sizable differences are found, similar to those observed when comparing the MRST/MSTW bench and NNPDF bench to the HERA–LHC benchmark of Sect. 3.3. Specifically, in these regions NNPDF uncertainties are generally larger than H1 bands: the width of the uncertainty band for the H1 fit varies much less between the data and extrapolation regions than that of the NNPDF bench. Also, the H1 central value always falls within the NNPDF uncertainty band, but the NNPDF central value tends to fall outside the H1 uncertainty band whenever the central values differ significantly. Figure 12 suggests that this may be due to the greater flexibility of the functional form in the NNPDF fit. Specifically, the $\bar{d}$ quark distribution at large x does not become negative in the H1 fit, because this behaviour is not allowed by the parameterization.

References

[1] M. Dittmar *et al.*, *Parton distributions: Summary report for the HERA - LHC workshop*. Preprint hep-ph/0511119, 2005.

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

[3] W. T. Giele and S. Keller, Phys. Rev. **D58**, 094023 (1998). hep-ph/9803393.

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

[5] NNPDF Collaboration, R. D. Ball *et al.*, Nucl. Phys. **B809**, 1 (2009). 0808.1231.

[6] G. Watt, A. D. Martin, W. J. Stirling, and R. S. Thorne, *Recent Progress in Global PDF Analysis*. Preprint 0806.4890, 2008.

[7] R. S. Thorne, A. D. Martin, W. J. Stirling, and G. Watt, *Parton Distributions for the LHC*. Preprint 0706.0456, 2007.

[8] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C21**, 33 (2001). hep-ex/0012053.

[9] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C13**, 609 (2000). hep-ex/9908059.

[10] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C19**, 269 (2001). hep-ex/0012052.

[11] W. Giele *et al.*, *The QCD/SM working group: Summary report*. Preprint hep-ph/0204316, 2002.

[12] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C21**, 443 (2001). hep-ex/0105090.

[13] New Muon Collaboration Collaboration, M. Arneodo *et al.*, Nucl. Phys. **B483**, 3 (1997). hep-ph/9610231.

[14] New Muon Collaboration Collaboration, M. Arneodo *et al.*, Nucl. Phys. **B487**, 3 (1997). hep-ex/9611022.

[15] BCDMS Collaboration, A. C. Benvenuti *et al.*, Phys. Lett. **B223**, 485 (1989).

[16] B. Badelek and J. Kwiecinski, Phys. Rev. **D50**, 4 (1994). hep-ph/9401314.

[17] A. d. R. Vargas Trevino, *Measurement of the inclusive e p scattering cross section at low Q**2 and x at HERA*. Prepared for 15th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2007), Munich, Germany, 16-20 Apr 2007.

[18] *Extraction of the proton pdfs from a combined fit of h1 and zeus inclusive dis cross sections*. These proceedings.

[19] M. Botje, *Qcdnum16.12*. Available from http://www.nikhef.nl/ h24/qcdnum/.

[20] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. **C28**, 455 (2003). hep-ph/0211080.

[21] S. Forte, L. Garrido, J. I. Latorre, and A. Piccione, JHEP **05**, 062 (2002). hep-ph/0204232.

[22] NNPDF Collaboration, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, JHEP **03**, 080 (2005). hep-ph/0501067.

[23] NNPDF Collaboration, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, JHEP **03**, 039 (2007). hep-ph/0701127.

[24] Particle Data Group Collaboration, C. Amsler *et al.*, Phys. Lett. **B667**, 1 (2008).