Eur. Phys. J. C (2015) 75:204
DOI 10.1140/epjc/s10052-015-3397-6

THE EUROPEAN PHYSICAL JOURNAL C

CrossMark

Regular Article - Theoretical Physics

Parton distributions in the LHC era: MMHT 2014 PDFs

L. A. Harland-Lang[1], **A. D. Martin**[2], **P. Motylinski**[1], **R. S. Thorne**[1,a]

[1] Department of Physics and Astronomy, University College London, London WC1E 6BT, UK
[2] Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LE, UK

Received: 19 December 2014 / Accepted: 8 April 2015 / Published online: 9 May 2015

Abstract We present LO, NLO and NNLO sets of parton distribution functions (PDFs) of the proton determined from global analyses of the available hard scattering data. These MMHT2014 PDFs supersede the 'MSTW2008' parton sets, but they are obtained within the same basic framework. We include a variety of new data sets, from the LHC, updated Tevatron data and the HERA combined H1 and ZEUS data on the total and charm structure functions. We also improve the theoretical framework of the previous analysis. These new PDFs are compared to the 'MSTW2008' parton sets. In most cases the PDFs, and the predictions, are within one standard deviation of those of MSTW2008. The major changes are the $u - d$ valence quark difference at small x due to an improved parameterisation and, to a lesser extent, the strange quark PDF due to the effect of certain LHC data and a better treatment of the $D \to \mu$ branching ratio. We compare our MMHT PDF sets with those of other collaborations; in particular with the NNPDF3.0 sets, which are contemporary with the present analysis.

Contents

[a] e-mail: thorne@hep.ucl.ac.uk

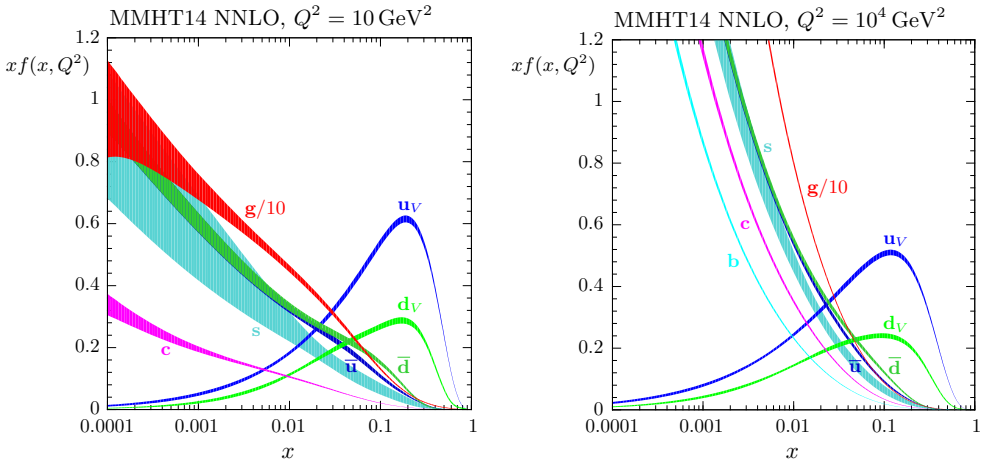

Fig. 1 MMHT2014 NNLO PDFs at $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2, with associated 68 % confidence-level uncertainty bands. The corresponding plot of NLO PDFs is shown in Fig. 20

PDFs and their uncertainties, together with the values of the input parameters. These sets of PDFs are the end products of the analysis – the grids and interpolation code for the PDFs can be found at [14] and will be available at [15] and a new HepForge [16] project site is foreseen. An example is given in Fig. 1, which shows the NNLO PDFs at scales of $Q^2 = 10\,\mathrm{GeV}^2$ and $Q^2 = 10^4\,\mathrm{GeV}^2$, including the associated one-sigma (68 %) confidence-level uncertainty bands.

Section 5 also contains a comparison of the NLO and NNLO PDFs with those of MSTW2008 [1]. The quality of the fit to the data at LO is far worse than that at NLO and NNLO, and is included for completeness, and because of the potential use in LO Monte Carlo generators, though the use of generators with NLO matrix elements is becoming far more standard. In Sect. 6 we make predictions for various benchmark processes at the LHC, and in Sect. 7 we discuss other data sets that are becoming available at the LHC which constrain the PDFs, but that are not included in the present global fit due to failure to satisfy our cut-off date; we refer to dijet and $W + c$ production and to the top quark differential distributions. In Sect. 8 we compare our MMHT PDFs with those of the very recent NNPDF3.0 analysis [17], and also with older sets of PDFs of other collaborations. In Sect. 9 we present our conclusions.

2 Changes in the theoretical procedures

In this section, we list the changes in our theoretical description of the data, from that used in the MSTW analysis [1]. We also glance ahead to mention some of the main effects on the resulting PDFs.

2.1 Input distributions

As is clear from the discussion in the Introduction, one improvement is to use parameterisations for the input distributions based on Chebyshev polynomials. Following the detailed study in [11], we take for most PDFs a parameterisation of the form

$$xf(x, Q_0^2) = A(1-x)^{\eta} x^{\delta} \left(1 + \sum_{i=1}^{n} a_i T_i^{\mathrm{Ch}}(y(x))\right), \quad (1)$$

where $Q_0^2 = 1$ GeV^2 is the input scale and the $T_i^{\mathrm{Ch}}(y)$ are Chebyshev polynomials in y, with $y = 1 - 2x^k$, where we take $k = 0.5$ and $n = 4$. The global fit determines the values of the set of parameters A, δ, η, a_i for each PDF, namely for $f = u_V$, d_V, S, s_+, where S is the light-quark sea distribution

$$S \equiv 2(\bar{u} + \bar{d}) + s + \bar{s}. \quad (2)$$

For $s_+ \equiv s + \bar{s}$ we set $\delta_+ = \delta_S$. As argued in [1] the sea quarks at very low x are governed almost entirely by perturbative evolution, which is flavour independent, and any difference in the shape at very low x is very quickly washed out. Hence, we choose to assume that this universality in the very low x shape is already evident at input. For s_+ we also set the third and fourth Chebyshev polynomials to be the same as for the light sea, as there are not enough data which can constrain the strange quark, while leaving all four parameters in the polynomial free leads to instabilities.

We still have to specify the parameterisations of the gluon and of the differences $\bar{d} - \bar{u}$ and $s - \bar{s}$. For the parameterisation of $\Delta \equiv \bar{d} - \bar{u}$ we set $\eta_\Delta = \eta_S + 2$, and we use the parameterisation

$$x\Delta(x, Q_0^2) = A_\Delta (1-x)^{\eta_\Delta} x^{\delta_\Delta} (1 + \gamma_\Delta x + \epsilon_\Delta x^2). \quad (3)$$

The (poorly determined) strange quark difference is taken to have a simpler input form than that in (1). That is,

$$s_- \equiv x(s - \bar{s}) = A_-(1-x)^{\eta_-} x^{\delta_-} (1 - x/x_0) \quad (4)$$

where A_-, δ_- and η_- are treated as free parameters, and where the final factor in (4) allows us to satisfy the third number sum rule given in (6) below, i.e. x_0 is a crossing point. Finally, it was found long ago [18] that the global fit was considerably improved by allowing the gluon distribution to have a second term with a different small x power

$$xg(x, Q_0^2) = A_g(1-x)^{\eta_g} x^{\delta_g} \left(1 + \sum_{i=1}^{2} a_{g,i} T_i^{\rm Ch}(y(x))\right) + A_{g'}(1-x)^{\eta_{g'}} x^{\delta_{g'}}, \quad (5)$$

where $\eta_{g'}$ is quite large, and concentrates the effect of this term towards small x. This means the gluon has seven free parameters (A_g being constrained by the momentum sum rule), which would be equivalent to using five Chebyshev polynomials if the second term were absent.

The choice $k = 0.5$, giving $y = 1 - 2\sqrt{x}$ in (1), was found to be preferable in the detailed study presented in [11]. It has the feature that it is equivalent to a polynomial in $\sqrt{x}$, the same as the default MSTW parameterisation. The half-integer separation of terms is consistent with the Regge motivation of the MSTW parameterisation. The optimum order of the Chebyshev polynomials used for the various PDFs is explored in the fit. It generally turns out to be $n = 4$ or 5. The advantage of using a parameterisation based on Chebyshev polynomials is the stability and good convergence of the values found for the coefficients a_i.

The input PDFs are subject to three constraints from the number sum rules

$$\int_0^1 {\rm d}x\, u_V(x, Q_0^2) = 2, \quad \int_0^1 {\rm d}x\, d_V(x, Q_0^2) = 1, \quad \int_0^1 {\rm d}x\, (s(x, Q_0^2) - \bar{s}(x, Q_0^2)) = 0, \quad (6)$$

together with the momentum sum rule

$$\int_0^1 {\rm d}x\, x[u_V(x, Q_0^2) + d_V(x, Q_0^2) + S(x, Q_0^2) + g(x, Q_0^2)] = 1. \quad (7)$$

We use these four constraints to fix A_g, A_u, A_d and x_0 in terms of the other parameters. In total there are 37 free (PDF) parameters in the optimum global fit, and there is also the strong coupling defined at the scale of the Z boson mass M_Z, i.e. $\alpha_s(M_Z^2)$, which we allow to be free when determining the best fit. Checks have been performed on our procedure which show that there is extremely little sensitivity to variation in Q_0^2 for either the fit quality or the PDFs extracted.

2.2 Deuteron corrections

It is still the case that we need deep inelastic data on deuteron targets [19–24] in order to fully separate the u and d distributions at moderate and large values of x. Thus we should consider the correction factor $c(x)$ to be applied to the deuteron data

$$F^d(x, Q^2) = c(x)[F^p(x, Q^2) + F^n(x, Q^2)]/2, \quad (8)$$

where we assume c is independent of Q^2 and where F^n is obtained from F^p by swapping up and down quarks, and anti-quarks; that is, isospin asymmetry is assumed. In the MSTW analysis, motivated by [25], despite the fact that the fit included all the deuteron data present in this analysis, the theory was only corrected for shadowing for small values of x, with a linear form for c with $c = 0.985$ at $x = 0.01$ and $c = 1$ just above $x = 0.1$; above this point it was assumed that $c = 1$.

In Ref. [11] we studied the deuteron correction factor in detail. We introduced the following flexible parameterisation of $c(x)$, which allowed for the theoretical expectations of shadowing (but which also allowed the deuteron correction factor to be determined by the data):

$$c(x) = (1 + 0.01N)\,[1 + 0.01c_1 \ln^2(x_p/x)], \quad x < x_p, \quad (9)$$

$$c(x) = (1 + 0.01N)\,[1 + 0.01c_2 \ln^2(x/x_p) + 0.01c_3 \ln^{20}(x/x_p)], \quad x > x_p, \quad (10)$$

where x_p is a 'pivot point' at which the normalisation is $(1+0.01N)$. For $x < x_p$ there is freedom for $c(x)$ to increase or decrease smoothly depending on the sign of the parameter c_1. The same is true above $x = x_p$, but the very large power in the c_3 term is added to allow for the expected rapid increase of $c(x)$ as $x \to 1$ due to Fermi motion. If, as expected, there is shadowing at low x and also a dip for high, but not too high, x (that is, if both c_1 and c_2 are found to be negative), then x_p is where $c(x)$ will be a maximum, as expected from antishadowing (provided $N > 0$). If we fix the value of x_p, then the deuteron correction factor $c(x)$ is specified by the values of four parameters: the c_i and N. In practice x_p is chosen to be equal to 0.05 at NLO, but a slightly smaller value of $x_p = 0.03$ is marginally preferred at NNLO.

As already emphasised, the introduction of a flexible parameterisation of the deuteron correction, $c(x)$, coupled with the extended Chebyshev parameterisation of the input PDFs was found [11], unlike MSTW [1], to describe the data for lepton charge asymmetry from $W^\pm$ decays well, and, moreover, to give a much better description of the same set of global data as used in the MSTW analysis. The only blemish was that for the best possible fit the four-parameter version of $c(x)$ had an unphysical form (with c_1 positive), so the preferred fit, even though it was of slightly lower quality, was taken to be the three-parameter form with $c_1 = 0$. In the present analysis (which includes the post-MSTW data) this blemish does not occur, and the four-parameter form of the

Table 1 The values of the parameters for the deuteron correction factor found in the MMSTWW [11] and the present (MMHT) global fits

PDF fit	N	c_1	c_2	$c_3 \times 10^8$
MMSTWW, 3 pars.	0.070	0	−0.608	3.36
MMSTWW, 4 pars.	−0.490	0.349	−0.444	3.40
MMHT2014 NLO	0.630 ± 0.831	-0.116 ± 0.507	-0.758 ± 0.324	3.44 ± 1.89
MMHT2014 NNLO	0.589 ± 0.738	-0.116 ± 0.996	-0.384 ± 0.182	0.0489 ± 0.0056

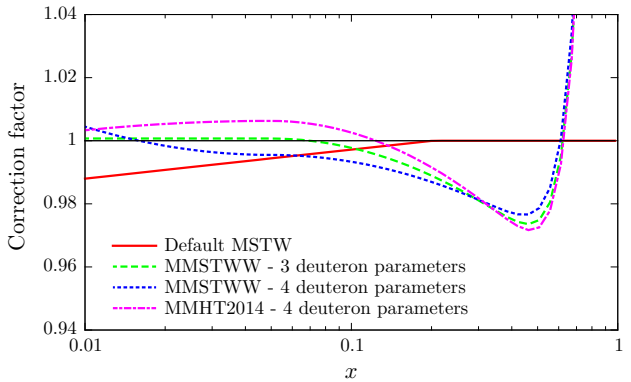

Fig. 2 The deuteron correction factors versus x at NLO shown for the fits listed in Table 1. The error corridor for the MMHT2014 curve is shown in Fig. 3, together with the result at NNLO

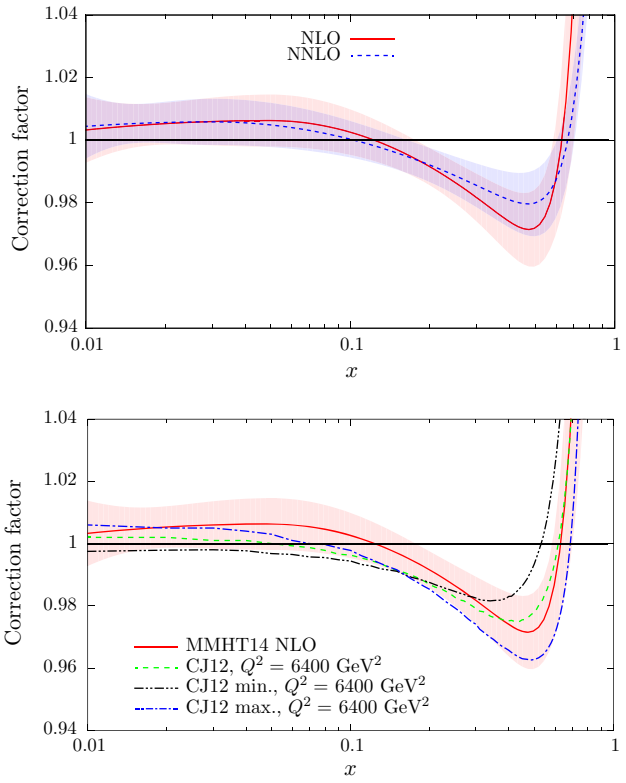

Fig. 3 The deuteron correction factors versus x at NLO and NNLO with uncertainties (*top*) and at NLO compared to the CJ12 corrections (*bottom*)

deuteron correction factor turns out to be much as expected theoretically. The parameters are listed in Table 1 and the corresponding deuteron correction factors shown in Fig. 2. The fit quality for the deuteron structure function data for MMSTWW at NLO with three parameters was 477/513, and it was just a couple lower when four parameters were used. For MMHT2014 at NLO the value is 471/513 and at NNLO is slightly better at 464/513. Hence, the new constraints on the flavour decomposition from the Tevatron and LHC are, if anything, slightly improving the fit to deuteron data, though part of the slight improvement is due to a small change in the way in which NMC data is used – see Sect. 2.7.

The uncertainties for the parameters in the MMHT2014 PDF fits are also shown in Table 1. These values are quoted as three times the uncertainty obtained using the standard $\Delta\chi^2 = 1$ rule. In practice we use the so-called "dynamic tolerance" procedure to determine $\Delta\chi^2$ for each of our eigenvectors, as explained in Section 6 of [1], and also discussed in Sect. 5 of this article, and a precise determination of the deuteron correction uncertainty is only obtained from the similar scan over χ^2 as used to determine eigenvector uncertainties. However, a typical value is three times the $\Delta\chi^2 = 1$ uncertainty, and this should give a fairly accurate representation of the deuterium correction uncertainty.[3] The correlation matrices for the deuteron parameters for the NLO and NNLO analyses are, respectively,

$$c_{ij}^{\mathrm{NLO}} = \begin{pmatrix} 1.000 & -0.604 & -0.693 & 0.177 \\ -0.604 & 1.000 & 0.426 & -0.116 \\ -0.693 & 0.426 & 1.000 & -0.360 \\ 0.177 & -0.116 & -0.360 & 1.000 \end{pmatrix}, \quad (11)$$

$$c_{ij}^{\mathrm{NNLO}} = \begin{pmatrix} 1.000 & -0.540 & -0.692 & 0.179 \\ -0.540 & 1.000 & 0.371 & -0.118 \\ -0.692 & 0.371 & 1.000 & -0.341 \\ 0.179 & -0.118 & -0.341 & 1.000 \end{pmatrix}. \quad (12)$$

We plot the central values and uncertainties of the deuteron corrections at NLO and at NNLO in the higher plot of Fig. 3. One can see that the uncertainty is of order 1 % in the region $0.01 \lesssim x \lesssim 0.4$ well constrained by deuteron data. Although the best fits now correspond to a decrease as x becomes very small this is not determined within even a one standard deviation uncertainty band. The lack of deuteron data at high x,

[3] This choice works well for PDF uncertainties, as discussed in [26].

$x \gtrsim 0.75$, mean that the correction factor is not really well determined in this region, and the uncertainty is limited by the form of the parameterisation. However, the sharp upturn at $x \sim 0.6$ is driven by the data.

Until recently, most of the other groups that have performed global PDF analyses do not include deuteron corrections. An exception is the analysis of Ref. [27]. In the present work, and in MMSTWW [11], we have allowed the data to determine what the deuteron correction should be, with an uncertainty determined by the quality of the fit. The CTEQ-Jefferson Lab collaboration [27] have performed three NLO global analyses which differ in the size of the deuteron corrections. They are denoted CJ12min, CJ12med and CJ12max, depending on whether they have mild, medium or strong deuteron corrections. We plot the comparison of these to our NLO deuteron corrections in the lower plot of Fig. 3. The CJ12 corrections are Q^2-dependent due to target mass and higher-twist contributions, as discussed in [28]. These contributions die away asymptotically, so we compare to the CJ12 deuteron corrections quoted at a very high Q^2 value of 6400 GeV^2. In the present analysis it turns out that the data select deuteron corrections that are in very good agreement for $x > 0.2$ with those given by the central CJ set, CJ12med. The behaviour at smaller values of x is sensitive to the lepton charge asymmetry data from $W^{\pm}$ decays at the Tevatron and LHC, the latter of which are not included in the CJ12 fits.

2.3 Nuclear corrections for neutrino data

The neutrino structure function data are obtained by scattering on a heavy-nuclear target. The NuTeV experiment [29] uses an iron target, and the CHORUS experiment [30] scatters on lead. Additionally the dimuon data from CCFR/NuTeV [31] is also obtained from (anti)neutrino scattering from an iron target. In the MSTW analysis [1] we applied the nuclear corrections R_f, defined as

$$f^A(x, Q^2) = R_f(x, Q^2, A)\, f(x, Q^2), \tag{13}$$

separately for each parton flavour f using the results of a NLO fit by de Florian and Sassot [32]. The f^A are defined to be the PDFs of a *proton* bound in a nucleus of mass number A. In the present analysis we use the updated results of de Florian et al., which are shown in Fig. 14 of [33]. The nuclear corrections for the heavy flavour quarks are assumed to be the same as that found for strange quarks, though the contribution from heavy quarks is very small. The updated nuclear corrections are quite similar, except for the strange quark for $x < 0.1$, though this does not significantly affect the extracted values of the strange quark. The new corrections improve the quality of the fit by ~25 units in χ^2, spread over a variety of data sets, including obvious candidates such as NuTeV $F_2(x, Q^2)$, but also HERA structure function data and CDF jet data which are only indirectly affected by nuclear corrections.

As in [1] we multiply the nuclear corrections by a three-parameter modification function, Eq. (73) in [1], which allows a penalty-free change in the details of the normalisation and shape. As in [1] the free parameters choose values $\lesssim 1$, i.e. they chose modification of only a couple of percent at most away from the default values. Hence, for both deuteron and heavy-nuclear corrections, we allow the fit to choose the final corrections with no penalty; but in both cases the corrections are fully consistent with expectation, i.e. any penalty applied would have very little effect.

2.4 General mass – variable flavour number scheme (GM-VFNS)

The treatment of heavy flavours – charm, bottom – has an important impact on the PDFs extracted from the global analysis due to the data available for $F_2^h(x, Q^2)$ with $h = c,\ b$, and also on the heavy flavour contribution to the total structure function at small x. Recall that there are two distinct regions where heavy quark production can be readily described. For $Q^2 \sim m_h^2$ the massive quark may be regarded as being only produced in the final state, while for $Q^2 \gg m_h^2$ the quark can be treated as massless, with the $\ln(Q^2/m_h^2)$ contributions being summed via the evolution equations. The GM-VFNS is the appropriate way to interpolate between these two regions, and as shown recently [34–36], the use of the fixed-flavour number scheme (FFNS) leads to significantly different results in a PDF fit to the GM-VFNS, even at NNLO. However, there is freedom to define different definitions of a GM-VFNS, which has resulted in the existence of various prescriptions, each with a particular reason for its choice. Well-known examples are the original Aivazis–Collins–Olness–Tung (ACOT) [37] and Thorne–Roberts (TR) [38] schemes, and their more recent refinements [39–41]. The MSTW analysis [1] adopted the more recent TR' prescription in [41].

Ideally one would like any GM-VFNS to reduce exactly to the correct fixed-flavour number scheme at low Q^2 and to the correct zero-mass VFNS as $Q^2 \to \infty$. This has been accomplished in [34], by introducing a new 'optimal' scheme which improves the smoothness of the transition region where the number of active flavours is increased by one. The optimal scheme is adopted in the present global analysis.[4]

In general, at NLO, the PDFs, and the predictions using them can vary by as much as 2 % from the mean value due

[4] We do not treat the top quark as a parton, i.e. even at high scale we remain in a five flavour scheme. Even at LHC energies the mass of the top quark is quite large compared to any other scale in the process, and the expressions for the cross sections for top production are all available in the scheme where the top appears in the final state.

Table 5 The values of $\chi^2/N_{\rm pts}$ for the data sets included in the global fit. For the NuTeV $\nu N \to \mu\mu X$ data, the number of degrees of freedom is quoted instead of $N_{\rm pts}$ since smearing effects mean nearby points are highly correlated. The details of corrections to data, kinematic cuts applied and definitions of χ^2 are contained in the text

Data set	LO	NLO	NNLO
BCDMS μp F_2 [125]	162/153	176/163	173/163
BCDMS μd F_2 [19]	140/142	143/151	143/151
NMC μp F_2 [20]	141/115	132/123	123/123
NMC μd F_2 [20]	134/115	115/123	108/123
NMC $\mu n/\mu p$ [21]	122/137	131/148	127/148
E665 μp F_2 [22]	59/53	60/53	65/53
E665 μd F_2 [22]	52/53	52/53	60/53
SLAC ep F_2 [23,24]	21/18	31/37	31/37
SLAC ed F_2 [23,24]	13/18	30/38	26/38
NMC/BCDMS/SLAC/HERA F_L [20,24,63–65,125]	113/53	68/57	63/57
E866/NuSea pp DY [88]	229/184	221/184	227/184
E866/NuSea pd/pp DY [89]	29/15	11/15	11/15
NuTeV νN F_2 [29]	35/49	39/53	38/53
CHORUS νN F_2 [30]	25/37	26/42	28/42
NuTeV νN xF_3 [29]	49/42	37/42	31/42
CHORUS νN xF_3 [30]	35/28	22/28	19/28
CCFR $\nu N \to \mu\mu X$ [31]	65/86	71/86	76/86
NuTeV $\nu N \to \mu\mu X$ [31]	53/40	38/40	43/40
HERA e^+p NC 820 GeV [61]	125/78	93/78	89/78
HERA e^+p NC 920 GeV [61]	479/330	402/330	373/330
HERA e^-p NC 920 GeV [61]	158/145	129/145	125 /145
HERA e^+p CC [61]	41/34	34/34	32/34
HERA e^-p CC [61]	29/34	23/34	21/34
HERA ep $F_2^{\rm charm}$ [62]	105 /52	72/52	82/52
H1 99–00 e^+p incl. jets [126]	77/24	14/24	–
ZEUS incl. jets [127,128]	140/60	45/60	–
DØ II $p\bar{p}$ incl. jets [119]	125/110	116/110	119/110
CDF II $p\bar{p}$ incl. jets [118]	78/76	63/76	59/76
CDF II W asym. [66]	55/13	32/13	30/13
DØ II $W \to \nu e$ asym. [67]	47/12	28/12	27/12
DØ II $W \to \nu\mu$ asym. [68]	16/10	19/10	21/10
DØ II Z rap. [90]	34/28	16/28	16/28
CDF II Z rap. [70]	95/28	36/28	40/28
ATLAS W^+, W^-, Z [10]	94/30	38/30	39/30
CMS W asymm $p_T > 35$ GeV [9]	10/11	7/11	9/11
CMS asymm $p_T > 25$ GeV, 30 GeV [77]	7/24	8/24	10/24
LHCb $Z \to e^+e^-$ [79]	76/9	13/9	20/9

Table 5 continued

Data set	LO	NLO	NNLO
LHCb W asymm $p_T > 20$ GeV [78]	27/10	12/10	16/10
CMS $Z \to e^+e^-$ [84]	46/35	19/35	22/35
ATLAS high-mass Drell–Yan [83]	42/13	21/13	17/13
CMS double-diff. Drell–Yan [86]	–	372/132	149/132
Tevatron, ATLAS, CMS $\sigma_{t\bar{t}}$ [91–97]	53/13	7/13	8/13
ATLAS jets (2.76 +7 TeV) [107,108]	162/116	106/116	–
CMS jets (7 TeV) [106]	150/133	138/133	–
All data sets	**3706/2763**	**3267/2996**	**2717/2663**

$$\Delta F = T\sqrt{\sum_{i,j=1}^{n}\frac{\partial F}{\partial a_i}C_{ij}\frac{\partial F}{\partial a_j}}, \tag{31}$$

where $C \equiv H^{-1}$ is the covariance matrix, and $T = \sqrt{\Delta\chi^2_{\rm global}}$ is the "tolerance" for the required confidence interval, usually defined to be $T = 1$ for 68 % confidence level.

It is very useful to diagonalise the covariance (or Hessian) matrix [133], and work in terms of the eigenvectors. The covariance matrix has a set of normalised *orthonormal* eigenvectors v_k defined by

$$\sum_{j=1}^{n} C_{ij}v_{jk} = \lambda_k v_{ik}, \tag{32}$$

where λ_k is the kth eigenvalue and v_{ik} is the ith component of the kth orthonormal eigenvector ($k = 1, \ldots, n$). The parameter displacements from the global minimum can be expanded in terms of rescaled eigenvectors $e_{ik} \equiv \sqrt{\lambda_k}v_{ik}$:

$$\Delta a_i \equiv a_i - a_i^0 = \sum_k e_{ik}z_k, \tag{33}$$

i.e. the z_k are the coefficients when we express a change in parameters away from their best-fit values in terms of the rescaled eigenvectors, and a change in parameters corresponding to $\Delta\chi^2_{\rm global} = 1$ corresponds to $z_k = 1$. This results in the simplification

$$\chi^2_{\rm global} = \chi^2_{\rm min} + \sum_k z_k^2. \tag{34}$$

Eigenvector PDF sets $S_k^{\pm}$ can then be produced with parameters given by

$$a_i(S_k^{\pm}) = a_i^0 \pm t\, e_{ik}, \tag{35}$$

with t adjusted to give the desired $T = \sqrt{\Delta\chi^2_{\rm global}}$. In the limit that Eq. (29) is exact, i.e. there are no significant corrections to quadratic behaviour, $t \equiv T$. We limit our number of eigenvectors so that this is true to a reasonable approximation. This results in the PDF eigenvector sets being obtained by fixing some of the parameters at their best-fit values, otherwise the large degree of correlation between some parameters would lead to significant violations in $t \approx T$.

As in [1] we do not determine the size of the eigenvectors using the standard $\Delta\chi^2 = 1$ or $T = 1$ rule. Rather, we allow $T \neq 1$ to account, primarily, for the tensions in fitting the *different* data sets within fixed-order perturbative QCD. Neither do we use a fixed value of T. Instead we use the "dynamical tolerance" procedure devised in [1]. In brief, we define the 68 % confidence-level region for each data set n (comprising N data points) by the condition that

$$\chi^2_n < \left(\frac{\chi^2_{n,0}}{\xi_{50}}\right)\xi_{68}, \tag{36}$$

where ξ_{68} is the 68th percentile of the χ^2-distribution with N degrees of freedom, and $\xi_{50} \simeq N$ is the most probable value. For each eigenvector (in each of the two directions) we then determine the values of t and T for which the χ^2_n for each data set n are minimised, together with 68 % confidence level limits defined by values at which Eq. (36) ceases to be satisfied. For a perfect data set we would only need the value of ξ_{68}, but for a number of data sets $\chi^2_{n,0}$ is not very close to ξ_{50} ($\xi_{50} \sim n_{\rm pts}$), being potentially both higher and lower, as seen in Table 5. For more details of the "dynamical tolerance" procedure see Section 6.2 of [1].

5.3.2 *Uncertainties of the MMHT2014 PDFs*

The increase in the parameterisation flexibility in the present MMHT analysis leads to an increase in the number of parameters left free in the determination of the PDF uncertainties,

Table 6 continued

Parameter	LO	NLO	NNLO
$a_{+,2}$	–	0.86501	0.88792
A_-	−0.53737	−0.01614	−0.011373
η_-	14.402	7.1599	6.4376
δ_-	0.91595	−0.26403	−0.26403
x_0	0.056131	0.026495	0.028993

as compared to the MSTW2008 analysis. Indeed, we now have 25 eigenvector pairs, rather than the 20 in [1] or even the 23 in [11]. The 25 parameters[13] left free for the determination of the eigenvectors consist of: η, δ, a_2 and a_3 for each of the valence quarks, A, η, δ, a_2 and a_3 for the light sea; $\int_0^1 \mathrm{d}x\ \Delta(x, Q_0^2)$, η and γ for $\bar{d} - \bar{u}$; η, δ, η_- and δ_- for the gluon (or η, δ, a_2 and a_3 at LO); A, η and a_2 for s_+ (or A, η and a_1 at LO); and A and η for s_-. During the determination of the eigenvectors all deuteron parameters, free coefficients for nuclear corrections and all parameters associated with correlated uncertainties, including normalisations, are allowed to vary (some with appropriate χ^2 penalty).

The most constraining data set for each eigenvector direction, and also the values of t and T are shown for the NLO fit in Table 7. The fractional contribution to the total uncertainty of each PDF is then also shown in summary in Table 8. The same information is shown for the NNLO fit in Tables 9 and 10. One can see that for the vast majority of cases there is good agreement between t and T at both NLO and NNLO. Hence, within the region of 68 % uncertainty confidence levels for the PDFs, the χ^2 distribution is quite accurately a quadratic function of the parameters. There is, however, a reasonable degree of asymmetry between the t and T values in the two directions for a single eigenvector, and it is nearly always the case that it is a different data set which is the main constraint in the two directions. In fact, the data set which has the most rapid deterioration in fit quality in one direction is often improving in fit quality until quite a high value of t along the other direction. This is an indication of the tension between data sets, with nearly all eigenvectors having some data sets which pull in opposite directions. The values of t and T for the 68 % confidence levels are on average about $t \approx T \approx 3$, i.e. $\Delta\chi^2_{\mathrm{global}} \approx 10$, though T^2 does vary between about 1 unit and at most $T^2 \approx 40$.

We comment briefly on the manner in which the values of t and T arise for some illustrative cases. For a number of eigenvectors there is one data set which is overwhelmingly most constraining. Examples are eigenvectors 17 and 25 at NLO and 7 and 25 at NNLO. A number of these are where the constraint is from the E866/NuSea Drell–Yan ratio data, since this is one of the few data sets sensitive to the $\bar{d} - \bar{u}$ difference. In these cases the tolerance tends to be low. For the cases where the tolerance is high there are some definite examples where this is due to tension between two data sets. One of the clearest and most interesting examples is eigenvector 13 at NLO. In this case the fit to HERA e^+p NC 820 GeV improves in one direction and deteriorates in the other, while the fit to NMC structure function data for $x < 0.1$ deteriorates in one direction and improves in the other. In this case the NMC data are at low Q^2 and the HERA data at higher Q^2 and the fit does not match either perfectly simultaneously. The effect is smaller at NNLO though is evident in eigenvector 3. Other cases where t is high and data sets are in very significant tension are eigenvector 4 at NLO, where DØ electron and muon asymmetry compete and eigenvector 20 at NLO where CCFR and NuTeV dimuon data prefer a different high-x strange quark. This complete tension is less evident in NNLO eigenvectors. However, there are some cases where one data set has deteriorating fit quality in one direction and improving quality in the other, while another data set deteriorates quickly in one direction, but varies only slowly in the other. Examples of this are eigenvectors 1 and 23 at NLO and eigenvector 1 at NNLO. Often the variation of χ^2 of all data sets is fairly slow except for one data set in one direction and a different data set in another direction. Examples of this are eigenvector 22 at NLO and eigenvectors 10, 22 and 24 at NNLO. A final type of cases is similar, but where one data set deteriorates in both directions but one other deteriorates slightly more quickly in one direction but very slowly in the other. Examples are eigenvector 4 at NNLO, where BCDMS data deteriorates in both directions but SLAC only in one direction and eigenvector 21 at NNLO, where ATLAS W, Z data deteriorates in both directions, but HERA data only in one direction.

We do not show the details of the eigenvectors at LO since we regard this as a much more approximate fit. However, we note that at LO the good agreement between t and T breaks down much more significantly, particularly for eigenvectors with the highest few eigenvalues. This is a feature of even more tension between data sets in the LO fit, and indeed, in the NLO and NNLO fit we would regard these eigenvectors as unstable, and discount them. However, we wish to obtain

[13] The expressions for the input PDFs in terms of the parameters are given in Sect. 2.1.

Table 7 Table of expected $\sqrt{\Delta\chi^2} = t$ and true $\sqrt{\Delta\chi^2} = T$ values for 68 % confidence-level uncertainty for each eigenvector and the most constraining data sets for the MMHT2014 NLO fits

Eigen-vector	$+t$	T	Most constraining data set	$-t$	T	Most constraining data set
1	4.00	3.97	HERA e^+p NC 920 GeV	4.30	4.66	HERA e^+p NC 820 GeV
2	2.50	2.84	HERA e^+p NC 920 GeV	1.80	1.53	NMC μd F_2
3	3.80	4.00	NMC.....HERA F_L	3.70	3.69	NMC μd F_2
4	4.05	4.00	DØ II $W \to \nu e$ asym.	5.00	5.11	DØ II $W \to \nu\mu$ asym.
5	3.40	3.35	DØ II $W \to \nu\mu$ asym.	4.20	4.45	NuTeV $\nu N \to \mu\mu X$
6	1.85	1.88	NuTeV $\nu N \to \mu\mu X$	3.70	3.71	DØ II $W \to \nu\mu$ asym.
7	1.55	1.67	E866/NuSea pd/pp DY	2.15	2.03	E866/NuSea pd/pp DY
8	2.75	2.64	DØ II $W \to \nu\mu$ asym.	1.90	2.01	E866/NuSea pd/pp DY
9	3.40	3.46	E866/NuSea pd/pp DY	3.80	3.78	BCDMS μp F_2
10	3.15	3.47	NuTeV $\nu N \to \mu\mu X$	2.40	2.13	NuTeV νN F_2
11	3.80	3.86	CDF II W asym.	4.00	3.96	E866/NuSea pd/pp DY
12	3.70	3.53	SLAC ed F_2	3.60	3.81	BCDMS μp F_2
13	4.30	5.47	HERA e^+p NC 820 GeV	5.30	4.33	NMC μd F_2
14	3.30	3.36	DØ II $W \to \nu e$ asym.	2.80	3.42	CMS W asym. $p_T > 35$ GeV
15	2.90	3.08	NuTeV νN xF_3	3.30	3.12	E866/NuSea pp DY
16	3.65	3.70	CDF II $p\bar{p}$ incl. jets	2.65	2.64	NuTeV νN xF_3
17	1.80	1.85	E866/NuSea pd/pp DY	2.40	2.16	E866/NuSea pd/pp DY
18	1.15	1.42	CMS asym. $p_T > 25, 30$ GeV	2.60	3.19	BCDMS μp F_2
19	2.60	2.86	CMS asym. $p_T > 25, 30$ GeV	2.10	3.35	DØ II $p\bar{p}$ incl. jets
20	1.60	1.72	CCFR $\nu N \to \mu\mu X$	1.55	1.45	NuTeV $\nu N \to \mu\mu X$
21	2.80	3.45	NuTeV $\nu N \to \mu\mu X$	3.30	3.47	ATLAS W^+, W^-, Z
22	4.70	6.48	NuTeV νN xF_2	4.00	3.67	NuTeV νN xF_3
23	1.90	1.96	NuTeV $\nu N \to \mu\mu X$	4.85	3.50	CCFR $\nu N \to \mu\mu X$
24	2.35	3.13	HERA e^+p NC 920 GeV	3.75	4.27	HERA e^+p NC 920 GeV
25	2.50	2.63	E866/NuSea pd/pp DY	1.30	2.15	E866/NuSea pd/pp DY

a conservative uncertainty on the PDFs at LO, so keep the same number of eigenvectors as at NLO and NNLO.

We see that there is some similarity between the eigenvectors for the NLO and NNLO PDFs, with some, e.g. 1, 5, 7, 19, 20, being constrained by the same data set and corresponding to the same type of PDF uncertainty. In some cases the order of the eigenvectors (determined by size of eigenvalue) is simply modified slightly by the changes between the NLO and NNLO fit e.g. 3 at NLO and 2 at NNLO, 23 at NLO and 24 at NNLO. However, despite the fact that the data fit at NNLO is very similar to that at NLO, and the parameterisation of the input PDFs is identical, the changes in the details of the NLO and NNLO fit are sufficient to remove any very clear mapping between the eigenvectors in the two cases, and some are completely different. We present the details of the eigenvectors at NLO here for the best-fit value of $\alpha_S(M_Z^2) = 0.120$. However, we also make available a NLO PDF set with $\alpha_S(M_Z^2) = 0.118$ with both a central value and a full set of eigenvectors (though the fit quality is 17 units worse for this value of $\alpha_S(M_Z^2)$). It is perhaps comforting to note that there is a practically identical mapping between the NLO eigenvectors for the two values of $\alpha_S(M_Z^2)$, with the main features of PDF uncertainties being the same, without any modification of the order of the eigenvectors. The precise values of t and T are modified a little, and in a couple of cases the most constraining sets changed (always for one which was almost the most constraining set at the other coupling value). The uncertainties (defined by changes in χ^2 relative to the best-fit values in each case) are very similar.

Table 8 The three numbers in each entry are the fractional contribution to the total uncertainty for the g, u_v, ... input distributions in the small x ($x < 0.01$), medium x ($0.01 < x < 0.1$) and large x ($x > 0.1$) regions, respectively, arising from eigenvector k in the NLO global fit. Each number has been multiplied by ten; for example, 4 denotes 0.4. For a precise value of x the sum of each column should be 10. However, the entries shown are the maximum fraction in each interval of x, so often do not satisfy this condition. In general we do not show contributions below 5 %, but for the first two eigenvectors at NLO no uncertainty contribution is this large, so we show the largest contributions

Eigen vector	g	u_v	d_v	S(ea)	$\bar{d}-\bar{u}$	$s+\bar{s}$	$s-\bar{s}$
1	–	–	–	0 0.3 0	–	–	–
2	–	–	–	0 0.4 0	–	–	–
3	4 0 0	–	–	–	–	–	–
4	2 0 0	0 0 2	–	–	–	–	–
5	1 0 0	–	–	1 0 0	–	–	1 0 0
6	–	–	–	–	–	–	2 1 2
7	–	–	–	–	0 2 2	–	–
8	–	–	0 0 2	–	0 1 2	–	–
9	–	1 2 3	–	–	0 1 2	–	–
10	–	–	–	2 1 0	–	2 3 1	–
11	–	0 1 2	2 3 4	–	0 1 1	–	–
12	–	4 3 5	1 2 2	0 1 0	–	–	–
13	8 5 2	1 1 1	0 0 1	1 1 0	–	–	–
14	–	–	2 3 7	–	–	–	–
15	1 2 2	1 1 2	2 1 2	0 0 1	1 1 0	–	–
16	0 1 5	1 2 2	0 1 2	0 3 3	1 2 0	–	–
17	–	–	–	0 0 1	2 3 4	–	–
18	–	4 4 0	0 1 0	–	–	–	–
19	–	–	2 3 2	–	–	–	–
20	–	–	–	0 0 1	1 0 0	0 0 6	1 0 0
21	0 0 1	1 2 0	2 1 2	4 4 4	0 1 0	5 6 6	4 3 3
22	1 2 0	1 0 1	2 2 2	4 2 4	0 0 1	2 1 2	1 0 0
23	–	0 1 0	0 0 1	1 0 3	1 0 0	1 2 2	2 8 10
24	0 5 6	–	0 1 1	0 1 0	0 0 1	–	–
25	–	–	–	–	7 4 9	–	–

5.3.3 *Data sets which most constrain the MMHT2014 PDFs*

It is very clear from Tables 7 and 9 that a wide variety of different data types are responsible for constraining the PDFs. At NLO 6 of the 50 eigenvector directions are constrained by HERA structure function data, 13 by fixed-target data structure function data, and 4 by the newest LHC data. Three of the LHC driven constraints are on the valence quarks and come from lepton asymmetry data. One is a constraint on the strange quark from the ATLAS W and Z data. There are still nine constraints from Tevatron data, again mainly on the details of the light-quark decomposition. The CCFR and NuTeV dimuon data [31] constrain eight eigenvector directions because they still provide by far the dominant constraint on the strange and antistrange quarks, which have five free parameters in the eigenvector determination. Similarly, the E866 Drell–Yan total cross section asymmetry data constrain 10 eigenvector directions mainly because the asymmetry data are still by far the best constraint on $\bar{d}-\bar{u}$, which has three free parameters.

At NNLO the picture is quite similar, but now HERA data constrain 11 eigenvector directions. Fixed-target data are similar to NLO with 10, but the Tevatron reduces to six. The LHC data now constrain eight eigenvector directions. As at NLO, this is dominantly lepton asymmetry data

constraint from the data in the larger x domain. For the two processes which constrain the high x gluon PDF, that is, jet production and the differential distributions for top-quark-pair production, it will be important to complete the NNLO formalism. There are already some results for the former process in [115–117] and for the latter process in [153]. On the experimental side it will be important to reliably measure the distributions for these processes, particularly for values of p_T^t and rapidity y_t, which are as large as possible.

Acknowledgments We particularly thank W. J. Stirling and G. Watt for numerous discussions on PDFs and for previous work without which this study would not be possible. We would like to thank Richard Ball, Jon Butterworth, Mandy Cooper-Sarkar, Albert de Roeck, Stefano Forte, Jun Gao, Joey Huston, Misha Ryskin, Pavel Nadolsky, Voica Radescu, Juan Rojo and Maria Ubiali for various discussions on PDFs and related issues. We would also like to thank Jon Butterworth and Mandy Cooper-Sarkar for helpful information on ATLAS data, Klaus Rabbertz and Ping Tan for help with CMS data and Ronan McNulty, Tara Shears and David Ward for LHCb data. We would also like to thank Andrey Sapranov, Pavel Starovoitov, Mark Sutton for help with APPLgrid, and Ben Watt for playing an instrumental role in interfacing this to the fitting code. We would also like to thank Alberto Accardi for providing the numbers for the CJ12 deuteron corrections and for discussions about the comparison. This work is supported partly by the London Centre for Terauniverse Studies (LCTS), using funding from the European Research Council via the Advanced Investigator Grant 267352. RST would also like to thank the IPPP, Durham, for the award of a Research Associateship held while most of this work was performed. We thank the Science and Technology Facilities Council (STFC) for support via grant awards ST/J000515/1 and ST/L000377/1.

References

1. A.D. Martin, W.J. Stirling, R.S. Thorne, G. Watt, Eur. Phys. J. C **63**, 189 (2009). arXiv:0901.0002
2. H.-L. Lai et al., Phys. Rev. D **82**, 074024 (2010). arXiv:1007.2241
3. R.D. Ball et al., Nucl. Phys. B **867**, 244 (2013). arXiv:1207.1303
4. ZEUS Collaboration, H1 Collaboration, A. Cooper-Sarkar, PoS EPS-HEP2011, 320 (2011). arXiv:1112.2107
5. S. Alekhin, J. Bluemlein, S. Moch, Phys. Rev. D **89**, 054028 (2014). arXiv:1310.3059
6. P. Jimenez-Delgado, E. Reya, Phys. Rev. D **89**, 074049 (2014). arXiv:1403.1852
7. R. Thorne, L. Harland-Lang, A. Martin, P. Motylinski, PoS DIS2014, 046 (2014). arXiv:1407.4045
8. P. Motylinski, L. Harland-Lang, A.D. Martin, R.S. Thorne (2014). arXiv:1411.2560
9. CMS Collaboration, S. Chatrchyan et al., Phys. Rev. Lett. **109**, 111806 (2012). arXiv:1206.2598
10. ATLAS Collaboration, G. Aad et al., Phys. Rev. D **85**, 072004 (2012). arXiv:1109.5141
11. A.D. Martin et al., Eur. Phys. J. C **73**, 2318 (2013). arXiv:1211.1215
12. CMS, S. Chatrchyan et al., Phys. Rev. D **90**, 032004 (2014). arXiv:1312.6283
13. ATLAS, G. Aad et al. (2014). arXiv:1407.0573
14. http://www.hep.ucl.ac.uk/mmht/. Accessed 24 Apr 2015
15. http://lhapdf.hepforge.org. Accessed 24 Apr 2015
16. http://www.hepforge.org/. Accessed 24 Apr 2015
17. The NNPDF Collaboration, R. D. Ball et al. (2014). arXiv:1410.8849
18. A.D. Martin, R. Roberts, W. Stirling, R. Thorne, Eur. Phys. J. C **23**, 73 (2002). arXiv:hep-ph/0110215
19. BCDMS Collaboration, A. Benvenuti et al., Phys. Lett. B **237**, 592 (1990)
20. New Muon Collaboration,M. Arneodo et al., Nucl. Phys. B **483**, 3 (1997). arXiv:hep-ph/9610231
21. New Muon Collaboration, M. Arneodo et al., Nucl. Phys. B **487**, 3 (1997). arXiv:hep-ex/9611022
22. E665 Collaboration,M. Adams et al., Phys. Rev. D **54**, 3006 (1996)
23. L. Whitlow, E. Riordan, S. Dasu, S. Rock, A. Bodek, Phys. Lett. B **282**, 475 (1992)
24. L. Whitlow, S. Rock, A. Bodek, E. Riordan, S. Dasu, Phys. Lett. B **250**, 193 (1990)
25. B. Badelek, J. Kwiecinski, Phys. Rev. D **50**, 4 (1994). arXiv:hep-ph/9401314
26. G. Watt, R. Thorne, JHEP **1208**, 052 (2012). arXiv:1205.4024
27. J.F. Owens, A. Accardi, W. Melnitchouk, Phys. Rev. D **87**, 094012 (2013). arXiv:1212.1702
28. A. Accardi, A.I.P. Conf. Proc. **1369**, 210 (2011). arXiv:1101.5148
29. NuTeV Collaboration, M. Tzanov et al., Phys. Rev. D **74**, 012008 (2006). arXiv:hep-ex/0509010
30. CHORUS Collaboration, G. Onengut et al., Phys. Lett. B **632**, 65 (2006)
31. NuTeV Collaboration, M. Goncharov et al., Phys. Rev. D **64**, 112006 (2001). arXiv:hep-ex/0102049
32. D. de Florian, R. Sassot, Phys. Rev. D **69**, 074028 (2004). arXiv:hep-ph/0311227
33. D. de Florian, R. Sassot, P. Zurita, M. Stratmann, Phys. Rev. D **85**, 074028 (2012). arXiv:1112.6324
34. R.S. Thorne, Phys. Rev. D **86**, 074017 (2012). arXiv:1201.6180
35. The NNPDF Collaboration, R.D. Ball et al., Phys. Lett. B **723**, 330 (2013). arXiv:1303.1189
36. R. Thorne, Eur. Phys. J. C **74**, 2958 (2014). arXiv:1402.3536
37. M. Aivazis, F.I. Olness, W.-K. Tung, Phys. Rev. D **50**, 3085 (1994). arXiv:hep-ph/9312318
38. R.S. Thorne, R.G. Roberts, Phys. Rev. D **57**, 6871 (1998). arXiv:hep-ph/9709442
39. A. Chuvakin, J. Smith, W. van Neerven, Phys. Rev. D **61**, 096004 (2000). arXiv:hep-ph/9910250
40. W.-K. Tung, S. Kretzer, C. Schmidt, J. Phys. G **28**, 983 (2002). arXiv:hep-ph/0110247
41. R.S. Thorne, Phys. Rev. D **73**, 054019 (2006). arXiv:hep-ph/0601245
42. G. D'Agostini, Nucl. Instrum. Methods **A346**, 306 (1994)
43. NNPDF Collaboration, R.D. Ball et al., JHEP **1005**, 075 (2010). arXiv:0912.2276
44. NuTeV Collaboration,D. Mason et al., Phys. Rev. Lett. **99**, 192001 (2007)
45. T. Bolton (1997). arXiv:hep-ex/9708014
46. S. Alekhin, J. Blumlein, S. Moch, Eur. Phys. J. C **71**, 1723 (2011). arXiv:1101.5261
47. R. Thorne, G. Watt, JHEP **1108**, 100 (2011). arXiv:1106.5789
48. NNPDF Collaboration, R.D. Ball et al., Phys. Lett. B **704**, 36 (2011). arXiv:1102.3182
49. M. Dasgupta, B. Webber, Phys. Lett. B **382**, 273 (1996). arXiv:hep-ph/9604388

50. LHCb collaboration, R. Aaij et al., J. Phys. G **41**, 055002 (2014). arXiv:1401.3288
51. LHCb Collaboration, LHCb-CONF-2012-013 (2012)
52. R. Thorne, A. Martin, W. Stirling, G. Watt (2008). arXiv:0808.1847
53. E. de Oliveira, A. Martin, M. Ryskin, Eur. Phys. J. C **72**, 2069 (2012). arXiv:1205.6108
54. E. de Oliveira, A. Martin, M. Ryskin, Eur. Phys. J. C **73**, 2361 (2013). arXiv:1212.3135
55. S. Jones, A. Martin, M. Ryskin, T. Teubner, JHEP **1311**, 085 (2013). arXiv:1307.7099
56. D.Y. Ivanov, B. Pire, L. Szymanowski, J. Wagner (2014). arXiv:1411.3750
57. C. White, R. Thorne, Phys. Rev. D **75**, 034005 (2007). arXiv:hep-ph/0611204
58. M. Ciafaloni, D. Colferai, G. Salam, A. Stasto, JHEP **0708**, 046 (2007). arXiv:0707.1453
59. G. Altarelli, R.D. Ball, S. Forte, Nucl. Phys. B **799**, 199 (2008). arXiv:0802.0032
60. E. de Oliveira, A. Martin, M. Ryskin, Eur. Phys. J. C **74**, 3118 (2014). arXiv:1404.7670
61. H1 and ZEUS Collaboration, F. Aaron et al., JHEP **1001**, 109 (2010). arXiv:0911.0884
62. H1 Collaboration, ZEUS Collaboration, H. Abramowicz et al., Eur. Phys. J. C **73**, 2311 (2013). arXiv:1211.1182
63. H1 Collaboration, F. Aaron et al., Phys. Lett. B **665**, 139 (2008). arXiv:0805.2809
64. H1 Collaboration, F. Aaron et al., Eur. Phys. J. C **71**, 1579 (2011). arXiv:1012.4355
65. ZEUS Collaboration, S. Chekanov et al., Phys. Lett. B **682**, 8 (2009). arXiv:0904.1092
66. CDF Collaboration, T. Aaltonen et al., Phys. Rev. Lett. **102**, 181801 (2009). arXiv:0901.2169
67. D0 Collaboration, V. Abazov et al., Phys. Rev. Lett. **101**, 211801 (2008). arXiv:0807.3367
68. D0 Collaboration, V.M. Abazov et al., Phys. Rev. D **88**, 091102 (2013). arXiv:1309.2591
69. Y. Li, F. Petriello, Phys. Rev. D **86**, 094034 (2012). arXiv:1208.5967
70. CDF Collaboration, T. A. Aaltonen et al., Phys. Lett. B **692**, 232 (2010). arXiv:0908.3914
71. T. Carli et al., Eur. Phys. J. C **66**, 503 (2010). arXiv:0911.2985
72. J.M. Campbell, R.K. Ellis, Phys. Rev. D **65**, 113007 (2002). arXiv:hep-ph/0202176
73. J.M. Campbell, R.K. Ellis, F. Tramontano, Phys. Rev. D **70**, 094012 (2004). arXiv:hep-ph/0408158
74. S. Catani, G. Ferrera, M. Grazzini, JHEP **1005**, 006 (2010). arXiv:1002.3115
75. R.D. Ball et al., JHEP **1304**, 125 (2013). arXiv:1211.5142
76. ATLAS Collaboration, G. Aad et al., Phys. Rev. Lett. **109**, 012001 (2012). arXiv:1203.4051
77. CMS Collaboration, S. Chatrchyan et al., JHEP **1104**, 050 (2011). arXiv:1103.3470
78. LHCb Collaboration, R. Aaij et al., JHEP **1206**, 058 (2012). arXiv:1204.1620
79. LHCb collaboration, R. Aaij et al., JHEP **1302**, 106 (2013). arXiv:1212.4620
80. LHCb Collaboration, LHCb-CONF-2013-007, CERN-LHCb-CONF-2013-007 (2013)
81. A. Martin, R. Roberts, W. Stirling, R. Thorne, Eur. Phys. J. C **39**, 155 (2005). arXiv:hep-ph/0411040
82. A. Martin, M. Ryskin, Eur. Phys. J. C **74**, 3040 (2014). arXiv:1406.2118
83. ATLAS Collaboration, G. Aad et al., Phys. Lett. B **725**, 223 (2013). arXiv:1305.4192
84. CMS Collaboration, S. Chatrchyan et al., Phys. Rev. D **85**, 032002 (2012). arXiv:1110.4973
85. R.D. Nnpdf,Ball et al., Nucl. Phys. B **877**, 290 (2013). arXiv:1308.0598
86. CMS Collaboration, S. Chatrchyan et al., JHEP **1312**, 030 (2013). arXiv:1310.7291
87. ATLAS Collaboration, G. Aad et al., JHEP **1406**, 112 (2014). arXiv:1404.1212
88. J.C. Webb (2003). arXiv:hep-ex/0301031
89. NuSea Collaboration, R. Towell et al., Phys. Rev. D **64**, 052002 (2001). arXiv:hep-ex/0103030
90. D0 Collaboration, V. Abazov et al., Phys. Rev. D **76**, 012003 (2007). arXiv:hep-ex/0702025
91. CDF Collaboration, D0 Collaboration, T.A. Aaltonen et al., Phys. Rev. D **89**, 072001 (2014). arXiv:1309.7570
92. ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C **71**, 1577 (2011). arXiv:1012.1792
93. ATLAS Collaboration, G. Aad et al., Phys. Lett. B **707**, 459 (2012). arXiv:1108.3699
94. ATLAS Collaboration, G. Aad et al., Phys. Lett. B **711**, 244 (2012). arXiv:1201.1889
95. ATLAS Collaboration, G. Aad et al., JHEP **1205**, 059 (2012). arXiv:1202.4892
96. ATLAS Collaboration, G. Aad et al., Phys. Lett. B **717**, 89 (2012). arXiv:1205.2067
97. ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C **73**, 2328 (2013). arXiv:1211.7205
98. CMS Collaboration, S. Chatrchyan et al., Phys. Rev. D **85**, 112007 (2012). arXiv:1203.6810
99. CMS Collaboration, S. Chatrchyan et al., JHEP **1211**, 067 (2012). arXiv:1208.2671
100. CMS Collaboration, S. Chatrchyan et al., Phys. Lett. B **720**, 83 (2013). arXiv:1212.6682
101. CMS Collaboration, S. Chatrchyan et al., Eur. Phys. J. C **73**, 2386 (2013). arXiv:1301.5755
102. CMS Collaboration, S. Chatrchyan et al., JHEP **1305**, 065 (2013). arXiv:1302.0508
103. CMS Collaboration, S. Chatrchyan et al., JHEP **1402**, 024 (2014). arXiv:1312.7582
104. M. Czakon, P. Fiedler, A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013). arXiv:1303.6254
105. ATLAS, CDF, CMS, D0 Collaborations (2014). arXiv:1403.4427
106. CMS Collaboration, S. Chatrchyan et al., Phys. Rev. D **87**, 112002 (2013). arXiv:1212.6660
107. ATLAS Collaboration, G. Aad et al., Phys. Rev. D **86**, 014022 (2012). arXiv:1112.6297
108. ATLAS, G. Aad et al., Eur. Phys. J. C **73**, 2509 (2013). arXiv:1304.4739
109. B. Watt, P. Motylinski, R. Thorne, Eur. Phys. J. C **74**, 2934 (2014). arXiv:1311.5703
110. T. Kluge, K. Rabbertz, M. Wobisch, p. 483 (2006). arXiv:hep-ph/0609285
111. fastNLO Collaboration, D. Britzger, K. Rabbertz, F. Stober, M. Wobisch (2012). arXiv:1208.3641
112. Z. Nagy, Phys. Rev. D **68**, 094002 (2003). arXiv:hep-ph/0307268
113. Z. Nagy, Phys. Rev. Lett. **88**, 122003 (2002). arXiv:hep-ph/0110315
114. CMS Collaboration, CMS-PAS-SMP-12-028
115. A. Gehrmann-De Ridder, T. Gehrmann, E. Glover, J. Pires, Phys. Rev. Lett. **110**, 162003 (2013). arXiv:1301.7310
116. J. Currie, A. Gehrmann-De Ridder, E. Glover, J. Pires, JHEP **1401**, 110 (2014). arXiv:1310.3993
117. J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E.N. Glover, J. Pires, PoS RADCOR2013, 004 (2014). arXiv:1312.5608
118. CDF Collaboration, A. Abulencia et al., Phys. Rev. D **75**, 092006 (2007). arXiv:hep-ex/0701051

119. D0 Collaboration, V.M. Abazov et al., Phys. Rev. D **85**, 052006 (2012). arXiv:1110.3771
120. N. Kidonakis, J. Owens, Phys. Rev. D **63**, 054019 (2001). arXiv:hep-ph/0007268
121. M.C. Kumar, S.-O. Moch, Phys. Lett. B **730**, 122 (2014). arXiv:1309.5311
122. D. de Florian, P. Hinderer, A. Mukherjee, F. Ringer, W. Vogelsang, Phys. Rev. Lett. **112**, 082001 (2014). arXiv:1310.7192
123. G.P. Salam, G. Soyez, JHEP **0705**, 086 (2007). arXiv:0704.0292
124. S. Carrazza, J. Pires (2014). arXiv:1407.7031
125. BCDMS Collaboration, A. Benvenuti et al., Phys. Lett. B **223**, 485 (1989)
126. H1 Collaboration, A. Aktas et al., Phys. Lett. B **653**, 134 (2007). arXiv:0706.3722
127. ZEUS Collaboration, S. Chekanov et al., Nucl. Phys. B **765**, 1 (2007). arXiv:hep-ex/0608048
128. ZEUS Collaboration, S. Chekanov et al., Phys. Lett. B **547**, 164 (2002). arXiv:hep-ex/0208037
129. Particle Data Group,K. Olive et al., Chin. Phys. C **38**, 090001 (2014)
130. S. Alekhin et al. (2011). arXiv:1101.0536
131. M. Botje et al. (2011). arXiv:1101.0538
132. G. Watt, JHEP **1109**, 069 (2011). arXiv:1106.5788
133. J. Pumplin et al., Phys. Rev. D **65**, 014013 (2001). arXiv:hep-ph/0101032
134. H.-L. Lai et al., Phys. Rev. D **82**, 054021 (2010). arXiv:1004.4624
135. R. Thorne, A. Martin, W. Stirling, G. Watt, PoS DIS2010, 052 (2010). arXiv:1006.2753
136. A. Martin, W. Stirling, R. Thorne, G. Watt, Eur. Phys. J. C **64**, 653 (2009). arXiv:0905.3531
137. R. Hamberg, W. van Neerven, T. Matsuura, Nucl. Phys. B **359**, 343 (1991)
138. R.V. Harlander, W.B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002). arXiv:hep-ph/0201206
139. A. Djouadi, M. Spira, P. Zerwas, Phys. Lett. B **264**, 440 (1991)
140. ATLAS Collaboration, G. Aad et al., JHEP **1405**, 059 (2014). arXiv:1312.3524
141. ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C **73**, 2261 (2013). arXiv:1207.5644
142. CMS Collaboration, S. Chatrchyan et al., Eur. Phys. J. C **73**, 2339 (2013). arXiv:1211.2220
143. ATLAS Collaboration, G. Aad et al., JHEP **1405**, 068 (2014). arXiv:1402.6263
144. CMS Collaboration, S. Chatrchyan et al., JHEP **1402**, 013 (2014). arXiv:1310.1138
145. ATLAS Collaboration, G. Aad et al. (2014). arXiv:1410.8857
146. J.M. Campbell, R.K. Ellis (2012). arXiv:1204.1513
147. M. Guzzi, K. Lipka, S.-O. Moch (2014). arXiv:1406.0386
148. J. Butterworth et al. (2014). arXiv:1405.1067
149. A. Martin, W. Stirling, R. Thorne, G. Watt, Eur. Phys. J. C **70**, 51 (2010). arXiv:1007.2624
150. R.D. Ball et al., Nucl. Phys. B **849**, 296 (2011). arXiv:1101.1300
151. https://indico.cern.ch/event/343303/. Accessed 24 Apr 2015
152. https://indico.cern.ch/event/252045/. Accessed 24 Apr 2015
153. M. Czakon, P. Fiedler, A. Mitov, (2014). arXiv:1411.3007