Uncertainties of predictions from parton distributions II: Theoretical errors

A.D. Martin[a], R.G. Roberts[b], W.J. Stirling[a] and R.S. Thorne[c,1]

[a] Institute for Particle Physics Phenomenology, University of Durham, DH1 3LE, UK
[b] Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, UK
[c] Cavendish Laboratory, University of Cambridge,
Madingley Road, Cambridge, CB3 0HE, UK

Abstract

We study the uncertainties in parton distributions, determined in global fits to deep inelastic and related hard scattering data, due to so-called theoretical errors. Amongst these, we include potential errors due to the change of perturbative order (NLO $\to$ NNLO), $\ln(1/x)$ and $\ln(1-x)$ effects, absorptive corrections and higher-twist contributions. We investigate these uncertainties both by including explicit corrections to our standard global analysis and by examining the sensitivity to changes of the x, Q^2, W^2 cuts on the data that are fitted. In this way we expose those kinematic regions where the conventional DGLAP description is inadequate. As a consequence we obtain a set of NLO, and of NNLO, *conservative* partons where the data are fully consistent with DGLAP evolution, but over a restricted kinematic domain. We also examine the potential effects of such issues as the choice of input parameterization, heavy target corrections, assumptions about the strange quark sea and isospin violation. Hence we are able to compare the theoretical errors with those uncertainties due to errors on the experimental measurements, which we studied previously. We use W and Higgs boson production at the Tevatron and the LHC as explicit examples of the uncertainties arising from parton distributions. For many observables the theoretical error is dominant, but for the cross section for W production at the Tevatron both the theoretical and experimental uncertainties are small, and hence the NNLO prediction may serve as a valuable luminosity monitor.

---

[1]Royal Society University Research Fellow.

1 Introduction

A knowledge of the partonic structure of the proton is an essential ingredient in the analysis of hard scattering data from pp or $p\bar{p}$ or ep high energy collisions. However, only the Q^2 (or scale) dependence of the parton distributions can be calculated from perturbative QCD. Perturbative QCD cannot fix their absolute values. Moreover, non-perturbative techniques are still far from being able to predict reliable magnitudes. Rather, to determine the distributions, it is necessary to resort to global analyses of a wide range of deep inelastic and related hard scattering data. The Bjorken x dependence of the distributions is parameterized at some low scale, and a fixed order (either LO, NLO or NNLO) DGLAP evolution performed to specify the distributions at the higher scales where data exist. Much attention is now being devoted to obtaining reliable uncertainties on the parton distributions obtained in this way. One obvious uncertainty is due to the systematic and statistical errors of the data used in the global fit. We will call these the *experimental* errors on the parton distributions and on the physical observables predicted from them. In fact, these so-called experimental uncertainties of the partons have so far been the main focus of attention. They have been estimated by several groups [1]–[10], working within a NLO framework using a variety of different procedures. For instance, in a previous paper [10] we estimated the parton errors using a Hessian approach in which we diagonalised the error matrix and then used the linear propagation of errors to estimate the uncertainty on a variety of typical observables. We confirmed that this simple approach works well in practice by using a more rigorous Lagrange multiplier method to determine the errors on the physical quantities directly. We also compared our results with those obtained in similar analyses performed by the CTEQ collaboration [4, 5, 6].

Besides the *experimental* errors, there are many other sources of uncertainty associated with the global parton analysis. These are the concern of the present paper. They may loosely be called *theoretical* errors. That is, we use theoretical errors to denote all uncertainties on the predicted observables other than those that arise from the systematic and statistical errors of the data that are included in the global fit. The various theoretical errors may be divided into four categories. Uncertainties due to (i) the selection of data fitted, (ii) the truncation of the DGLAP perturbation expansion, (iii) specific theoretical effects[2] (namely $\ln 1/x$, $\ln(1-x)$, absorptive and higher twist corrections) and (iv) input assumptions (such as the choice of parameterization, heavy target corrections, whether isospin violation is allowed and the form of the strange quark sea). These theoretical errors are discussed in turn in the following four sections. Then, in Section 6, we study the implications of the error analysis for determining the value of $\alpha_S(M_Z^2)$ from the global fit, and for the accurate predictions of particularly relevant observables at the Tevatron and the LHC.

---

[2]Of course, differences can also arise from alternative methods of treating the behaviour near the heavy flavour thresholds. However, this issue is now well understood. Either fixed-flavour-number schemes or, preferably, one of the selection of variable-flavour-number schemes can be used for a correct description [11]. Different choices will result in slight differences in the extracted partons, but only minimal variation in predictions for physical quantities.

2   Uncertainties due to the selection of data fitted

In principle, if the DGLAP formalism is valid and the various data sets are compatible, then changing the data that are included in the global analysis should not move the predictions outside the error bands. In practice this is not the case. As an extreme example, consider the analysis of the H1 and BCDMS deep inelastic data for F_2 [7]. The absence of the other fixed target data and, in particular, the Tevatron jet data from the global parton fit results in a gluon which is much smaller for $x \gtrsim 0.3$, and hence larger at small x, from those obtained in global fits to a wider range of data. Another example occurs in the analysis of Ref. [2], where fitting to a small subset of the total data (namely H1(94) [12], BCDMS [13] and E665 [14] data for F_2^p) yields a surprisingly low value $\alpha_S = 0.112 \pm 0.001$, and a very hard high x down-quark distribution. Even fits of different groups, at the same perturbative order, to basically the same data lead to unexpectedly sizable differences in partons and in predictions for observables. For example, compare the MRST2002 and CTEQ6 predictions at NLO for the cross sections for W and Higgs boson production at hadron colliders, shown in Fig. 15 of Ref. [10]. We see that the CTEQ6 values at the Tevatron are more than 1.5% smaller for W production and 8% smaller for Higgs production; differences which are larger than expected from an analysis based on the experimental errors of the data used in the fits.

We begin by investigating the stability of the global parton analysis to different choices of the data cuts (W_{cut}, x_{cut}, Q_{cut}^2), defined such that data with values of W, x or Q^2 below the cut are excluded from the global fit, with the implicit assumption that the remaining data can be described by pure leading-twist DGLAP evolution. We find the minimum values of the data cuts for which this stability occurs. The kinematic variable W is the invariant mass of the system X recoiling against the scattered lepton in deep inelastic lepton–proton scattering $lp \to lX$. It follows that

$$W^2 \simeq Q^2(1-x)/x. \tag{1}$$

In the remainder of this section, all fits are performed at next-to-leading order (NLO) unless otherwise stated. In Section 3 we discuss what happens when we use the NNLO formalism.

2.1   Effect of the cut on W^2

In the original MRST global analyses we have fitted to data with $W^2 > 10$ GeV^2, assuming that this was sufficiently high to avoid resonance structure, large $\ln(1-x)$ effects and associated higher-twist corrections. However, we had no quantitative justification for this precise choice, and, noting that it resulted in a systematically poor fit to SLAC data with $W^2 \gtrsim 10$ GeV^2, we subsequently raised the cut to $W_{\text{cut}}^2 = 12.5$ GeV^2 [15]. This provided an acceptable description of the SLAC data. In order to make a more systematic investigation of the stability of the fit to varying this cut, we performed a series of global analyses with W_{cut}^2 ranging from 12.5 to 25 GeV^2. When raising the cut from 12.5 to 15 GeV^2 we find χ^2 to the remaining data improved by only 4, while an increase of W_{cut}^2 from 15 to 20 or 25 GeV^2 resulted in no significant

improvement. Taking these results at face value, we conclude that $W_{\mathrm{cut}}^2 = 15$ GeV^2 is a conservative choice and that there is no reason to eliminate even more data.

However, inspection of the description of the SLAC and BCDMS data in this low W^2 region shows a lack of compatibility of the two data sets in the region where they overlap; specifically for W^2 in the interval $(6, 15)$ GeV^2 for the higher values of x. Thus the stability achieved at $W_{\mathrm{cut}}^2 = 15$ GeV^2 corresponds to the SLAC data completely disappearing from the fit. Hence, while the resulting fit describes the BCDMS data well, an extrapolation to only slightly lower W^2 values gives a very poor description of the SLAC data. This implies that the achieved stability is an artifact of the incompatibility of the two data sets for $F_2^{p,n}$. Indeed, when phenomenological higher-twist contributions are introduced into the analysis they still have significant impact for $W^2 > 15$ GeV^2, see Fig. 2 of Ref. [16]. This implies that a genuinely conservative choice of cut would be $W_{\mathrm{cut}} \sim 20$ GeV^2. However, because a good description is obtained of the only available data in the $(15, 20)$ GeV^2 interval we may set $W_{\mathrm{cut}} = 15$ GeV^2 without prejudicing the analyses. Future measurements of F_2^p at HERA in this kinematic domain would clearly be valuable.

2.2   Effect of the choice of the cut on x

In Table 1 we show the values of χ^2 for global analyses performed for different values of x_{cut}, together with the number of data points fitted. Each column represents the χ^2 values corresponding to a fit performed with a different choice of the cut in x. For example, if only data above $x_{\mathrm{cut}} = 0.001$ are fitted, then the total $\chi^2 = 2119$, with a contribution of $\chi^2 = 2055$ coming from the subset of data with $x > 0.0025$, and $\chi^2 = 2012$ from the subset with $x > 0.005$, etc.

To obtain a measure of the stability of the analysis to changes in the choice of x_{cut}, we compare fits in adjacent columns, that is with $(x_{\mathrm{cut}})_{i+1}$ and $(x_{\mathrm{cut}})_i$. In particular, it is informative to compare the contributions to their respective χ^2 values from the subset of data with $x > (x_{\mathrm{cut}})_{i+1}$. If stability were achieved, then we would expect the difference $\Delta\chi^2$ between these two χ^2 contributions to be very small. We stress that these two χ^2 contributions describe the quality of the two fits to the *same* subset of the data. Thus, as we shall explain below, a measure, Δ_i^{i+1}, of the stability of the analysis is $\Delta\chi^2$ divided by the number of data points omitted when going from the fit with $(x_{\mathrm{cut}})_i$ to the fit with $(x_{\mathrm{cut}})_{i+1}$. For example, if we raise the x_{cut} from 0.001 to 0.0025 then $\Delta\chi^2 = 2055 - 2040$ for the data with $x > 0.0025$, and the number of data points omitted is $1961 - 1898 = 63$. Thus the measure $\Delta_{0.001}^{0.0025} = 15/63 = 0.24$, as shown in the last row of Table 1. If we were to start from a fit with x_{cut} below the value at which the theoretical framework is expected to be valid (due to neglected $\ln 1/x$ effects, etc.), then we would expect Δ_i^{i+1} to be significantly non-zero. As x_{cut} is increased, then Δ_i^{i+1} should decrease and, in the ideal case, approach and remain near zero; indicating that we are in the stable domain where the theoretical framework is appropriate. In this way the behaviour of Δ_i^{i+1}, as x_{cut} is changed, acts as an indicator of the stability of the analysis.

3

$x_{\rm cut}$:	0	0.0002	0.001	0.0025	0.005	0.01
# data points	2097	2050	1961	1898	1826	1762
$\alpha_S(M_Z^2)$	0.1197	0.1200	0.1196	0.1185	0.1178	0.1180
$\chi^2(x > 0)$	2267					
$\chi^2(x > 0.0002)$	2212	2203				
$\chi^2(x > 0.001)$	2134	2128	2119			
$\chi^2(x > 0.0025)$	2069	2064	2055	2040		
$\chi^2(x > 0.005)$	2024	2019	2012	1993	1973	
$\chi^2(x > 0.01)$	1965	1961	1953	1934	1917	1916
Δ_i^{i+1}		0.19	0.10	0.24	0.28	0.02

Table 1: Each column shows the χ^2 values obtained from a NLO global analysis with a different choice of $x_{\rm cut}$, together with the number of data points fitted and the value of $\alpha_S(M_Z^2)$ obtained. The first χ^2 entry in a given column is the total χ^2, and the subsequent entries show the contributions to χ^2 from subsets of the data that are fitted. The quantity Δ_i^{i+1}, shown in the final row, is a measure of stability to changing the choice of $x_{\rm cut}$, as explained in the text. In these analyses we take the default cut in Q^2, that is $Q_{\rm cut}^2 = 2$ GeV^2.

Inspection of the values of Δ_i^{i+1} in the last row of Table 1 shows a significant improvement in the quality of the fit each time $x_{\rm cut}$ is raised by an amount corresponding to the omission of about a further 70 data points, until the final step when $x_{\rm cut}$ is increased from 0.005 to 0.01, when we see that there is no further improvement at all. In fact, raising $x_{\rm cut}$ from 0.01 to 0.02 confirms this stability. Indeed, Δ_i^{i+1} is greater when raising $x_{\rm cut}$ from 0.0025 to 0.005 than in any of the previous steps. Hence we conclude that $x \simeq 0.005$ is a safe choice of $x_{\rm cut}$. Below this value there is a degree of incompatibility between the data and the theoretical description. At each step the gluon distribution extrapolated below $x_{\rm cut}$ becomes increasingly smaller, allowing the higher x gluon to increase (and to carry more momentum), see Fig. 1. Because almost no momentum is carried by the gluon at very small x, this redistribution of the gluon momentum becomes increasingly possible as $x_{\rm cut}$ is raised, hence explaining why Δ_i^{i+1} has a tendency to increase, until stability is finally reached. In general, we see that there is a slight decrease of $\alpha_S(M_Z^2)$ from our standard value of 0.1197 to 0.1178.

2.3 Effect of the choice of the cut on Q^2

Repeating the above procedure, but now performing fits with different choices of the Q^2 cut, we obtain the results shown in Table 2. Here the behaviour of the values of Δ_i^{i+1} is not so dramatic as for the study of $x_{\rm cut}$. As a consequence it is more difficult to select the value of $Q_{\rm cut}^2$ for which the theoretical description becomes safe. However, there is a general decrease in the value of Δ_i^{i+1} as the value $Q_{\rm cut}^2$ is increased. Certainly the choice $Q_{\rm cut}^2 < 7$ GeV^2 appears inappropriate,
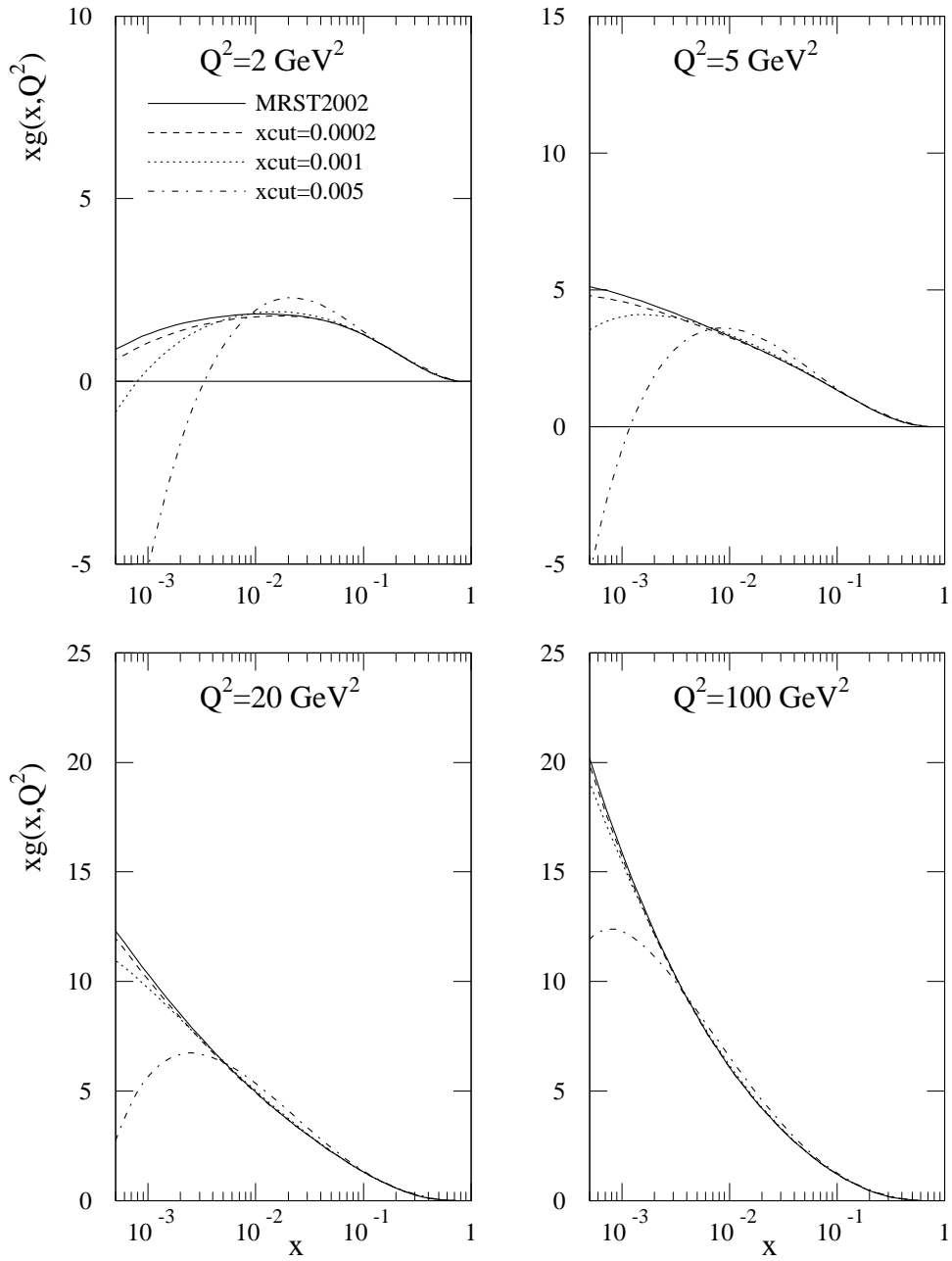
Figure 1: The gluon distribution obtained in NLO global fits with different values of x_{cut}, that is x_{cut} taken to be 0.0002 (dashed curve), 0.001 (dotted) and 0.005 (dot-dashed), compared to the default set with $x_{\mathrm{cut}} = 0$ (continuous curve).

Q^2_{cut} (GeV²) :	2	4	7	10	14	20
# data points	2097	1868	1681	1537	1398	1244
$\alpha_S(M_Z^2)$	0.1197	0.1194	0.1185	0.1180	0.1169	0.1174
$\chi^2(Q^2 > 2)$	2267					
$\chi^2(Q^2 > 4)$	2046	2022				
$\chi^2(Q^2 > 7)$	1844	1824	1806			
$\chi^2(Q^2 > 10)$	1716	1694	1670	1656		
$\chi^2(Q^2 > 14)$	1594	1573	1553	1536	1533	
$\chi^2(Q^2 > 20)$	1406	1388	1370	1354	1351	1348
Δ_i^{i+1}	0.11	0.10	0.10	0.02	0.02	

Table 2: Each column shows the χ^2 values obtained from a NLO global analysis with a different choice of the cut in Q^2, together with the number of data points fitted and the value of $\alpha_S(M_Z^2)$ obtained. The first χ^2 entry in a given column is the total χ^2, and the subsequent entries show the contributions to χ^2 from subsets of the data that are fitted. Δ_i^{i+1}, shown in the final row, is a measure of stability to changing the choice of Q^2_{cut}, as explained in the text.

and $Q^2_{\text{cut}} > 14$ GeV² is definitely acceptable. We therefore take the reasonably conservative choice of $Q^2_{\text{cut}} = 10$ GeV². This gradual reduction of Δ_i^{i+1} is an indication of the presence of higher-order corrections, whose relative strength falls off only like $1/\ln Q^2$. If higher-twist corrections were the dominant effect, then we would expect a more dramatic reduction of Δ_i^{i+1} at some low value of Q^2_{cut}.

2.4   Effect of a cut on the product xQ^2

The theoretical uncertainties at small x and small Q^2 may be strongly correlated. That is, the main theoretical uncertainty at small x is due to higher-twist effects rather than higher-order contributions. In order to investigate this possibility, we perform a series of global fits, each with a different choice of cut on the product xQ^2. We begin with $(xQ^2)_{\text{cut}} = 0.001$ GeV², which corresponds to a loss of 42 data points, and proceed in steps of a loss of about a further 50–100 data points until stability is achieved. This occurs when $(xQ^2)_{\text{cut}} = 0.6$ GeV², at which point 589 data points have been removed. The corresponding value of $\alpha_S(M_Z^2) = 0.1183$, and the gluon distribution, are both very similar to those obtained with $x_{\text{cut}} = 0.005$. Hence we arrive at a similar final result to that of applying an x_{cut} alone, but with the loss of about twice as much data. Indeed we have removed all HERA data below $x \sim 0.003$, and have only a handful of high Q^2 points left for $x < 0.005$. Thus to cut on xQ^2 appears to be much more inefficient than to cut on x alone.

2.5 Combined cut on x and Q^2

Finally we check the effect of imposing a range of cuts on x having already chosen the conservative Q^2 cut of $Q^2_{\text{cut}} = 10$ GeV^2, and alternatively choosing different cuts in Q^2 having already selected the safe x_{cut} of 0.005. In each case the additional cut leads to further improvements in the quality of the fit, and complete stability is only achieved when the combined cuts $x_{\text{cut}} = 0.005$ and $Q^2_{\text{cut}} = 10$ GeV^2 are imposed. ($\Delta_i^{i+1} = 0.13$ when going from $x_{\text{cut}} = 0.0025$ to 0.005 at $Q^2_{\text{cut}} = 10$ GeV^2, and $\Delta_i^{i+1} = 0.05$ when going from $Q^2_{\text{cut}} = 7$ GeV^2 to 10 GeV^2 at $x_{\text{cut}} = 0.005$). This combined cut yields, in our opinion, parton distributions largely free from theoretical uncertainties, but only within this restricted kinematic domain. We denote this conservative parton set by MRST(cons). The corresponding value of $\alpha_S(M_Z^2)$ is 0.1162, and the gluon distribution is similar in shape to that obtained for the fit with $x_{\text{cut}} = 0.005$ (and $Q^2_{\text{cut}} = 2$ GeV^2). To estimate the error of this MRST(cons) prediction of α_S we use a tolerance of $\Delta\chi^2 = 5$, instead of our previous choice of $\Delta\chi^2 = 20$ [10], since we now have a more compatible collection of data, without so many accompanying 'tensions' between different data sets. However, the reduced amount of fitted data with, in particular, a more limited range of Q^2, yields an intrinsically larger error so that our final uncertainty on α_S is again approximately ± 0.002 (despite the smaller tolerance).

Fig. 2 compares the MRST(cons) partons distributions with those of our default set. We stress that these partons have a limited range of applicability. In order to investigate the theoretical uncertainty outside the kinematic domain, $x > 0.005$ and $Q^2 > 10$ GeV^2, of the fitted data, we next study a variety of possible theoretical corrections to the NLO leading-twist framework.

3 Uncertainty associated with change from NLO to NNLO fit

Clearly there will be uncertainties associated with the truncation of the DGLAP evolution at a given perturbative order. In the past, these have usually been estimated by seeing the effect of changes of scale on physical observables in the NLO analysis. However, this approach is flawed since it provides no information about higher logs in $1/x$ and $(1-x)$ at higher orders, and now we can do better. The deep inelastic coefficient functions are known at NNLO [17]. Moreover, valuable partial information has been obtained about the NNLO splitting functions [18, 19]. This greatly limits the possible behaviour of the splitting functions down to quite small values of x. Indeed, van Neerven and Vogt [20] have constructed a range of compact analytic functions that are all compatible with the available information.

3.1 Effect of NNLO corrections

We have performed NNLO global analyses in two previous publications [21, 22]. The dominant effects of the NNLO corrections are indeed the additional $\ln(1-x)$ terms in the nonsinglet
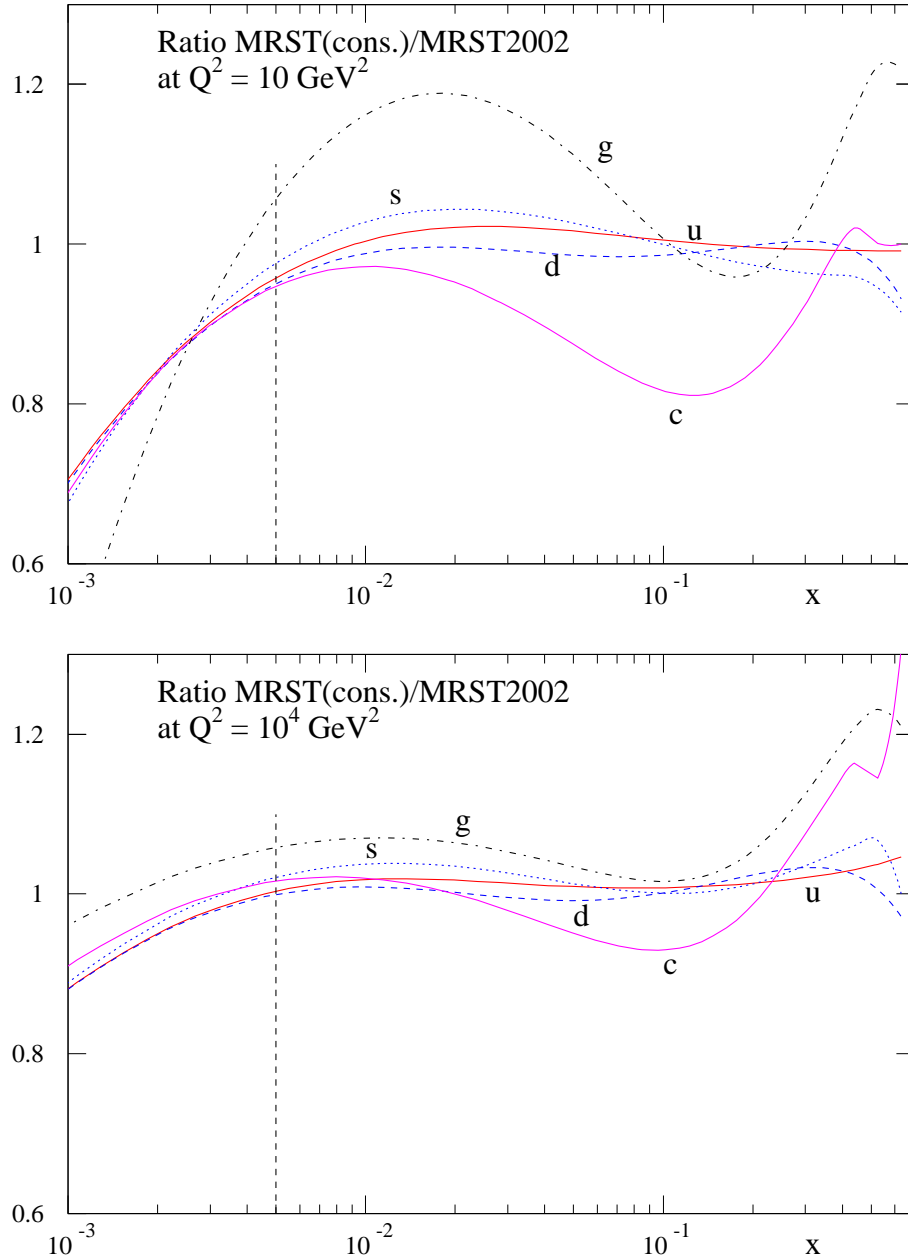
Figure 2: Comparison of MRST(cons) partons with the default NLO set MRST(2002). The former partons are only reliable for $x > 0.005$ and $Q^2 > 10$ GeV^2.

coefficient functions which influence large x, and the behaviour of the coefficient functions and splitting functions at small x, which is heavily determined by the leading $\ln(1/x)$ terms. The additional terms in the nonsinglet coefficient function lead to an enhancement of the large x structure functions, and hence a small, but significant, reduction in the valence quark distributions at large x. However, because this enhancement is proportional to $\alpha_S^2(Q^2)$ it falls quickly with Q^2 and the evolution of the structure functions increases at large x. This enables a slightly better fit to be obtained, but the natural increase in evolution speed results in a lower value of $\alpha_S(M_Z^2)$, i.e. it falls from ~ 0.119 at NLO to ~ 0.116 at NNLO, which is mainly determined by the high-x structure function evolution.

At small x the speed of evolution of $F_2(x, Q^2)$ is also increased, both by the behaviour of the NNLO splitting function $P_{qg}^{(2)}(x)$ and the NNLO coefficient function $C_{2g}^{(2)}(x)$. The evolution of the gluon is decreased, however, due to the leading $\ln(1/x)/x$ term in $P_{gg}^{(2)}(x)$ having a negative coefficient. In our first analysis we found that the small x gluon at NNLO was far below that at NLO, but this was largely due to the initial estimate of $P_{qg}^{(2)}(x)$, which was reduced significantly when some additional moments became available. In our more recent analysis we do indeed find that the NNLO gluon is a little smaller at small x due to the natural increase in evolution of $F_2(x, Q^2)$. However, the large x gluon needs to be larger at NNLO because the decrease in the large x quarks discussed above reduces the high-E_T Tevatron jet cross section, which must be corrected by an increase in the large-x gluon distribution.[3] This redistribution of the gluon between large and small x is qualitatively in agreement with the momentum sum rule, which is a strong constraint on the form of the gluon. Indeed, the complete NNLO fit works nicely, and displays a very slight improvement in quality compared to NLO.

However, it is not clear whether additional large logarithms in $(1 - x)$ and $1/x$ beyond NNLO will lead to further modifications to the partons, and any resulting predictions. We have produced predictions for W and Z production at the Tevatron and the LHC in [21, 22], and also for the longitudinal structure function $F_L(x, Q^2)$ (the NNLO coefficient function for this having been estimated [21] in the same way as the NNLO splitting functions). The former are quite stable, showing changes of $\sim 4\%$ when going from NLO to NNLO. However, this is larger than the $\sim 2\%$ 'experimental' error estimated at NLO [6, 10], and is dependent on the quark distributions, which are directly obtained by comparison with the structure functions. In contrast, $F_L(x, Q^2)$ is dependent mainly on the gluon which has no direct constraint at low x, so that when one goes from LO to NLO to NNLO the variation is very large, i.e., $\sim 20\%$ at $x = 0.0001$, even for $Q^2 > 100 \, \text{GeV}^2$. This variation is driven largely by the $\ln(1/x)$ terms, and it is not clear if any sort of convergence has been reached at NNLO. The same is potentially true for other gluon dominated quantities.

---

[3]There is not, at present, a full NNLO calculation of the jet cross section available, but a calculation of leading threshold logarithms [23] suggests that the NNLO contribution is not large or very E_T-dependent.

3.2 Cuts on Data at NNLO

We might hope that at NNLO the cuts on data required to achieve stability are rather less stringent than at NLO. However, in the same way that the quality of the global fit only improves very slightly going from NLO to NNLO, the size of the cuts in W^2, Q^2 and x does not change much. Nevertheless, there are some advantages to be gained at NNLO as we discuss in detail below.

Raising W_{cut}^2 from 12.5 GeV^2 has a similar effect as at NLO. However, this stability at $W_{\mathrm{cut}}^2 = 12.5$ GeV^2 or at most 15 GeV^2 is subject to the same caveat as at NLO, i.e. a certain incompatibility of data for $W^2 \sim 15$ GeV^2. It is also true that the low W^2 data is one of the major constraints on $\alpha_S(M_Z^2)$, which is rather lower at NNLO than NLO, being more consistent between different data sets at NNLO. Also, if one extrapolates the NNLO fit below $W^2 = 12.5$ GeV^2 the departure between data and theory occurs more slowly than at NLO. Hence, although the 'optimal' value of W_{cut}^2 does not change at NNLO there are indications that the theoretical description is improving at low W^2. We will return to this point later.

At NNLO there is a significant improvement each time x_{cut} is raised until $x_{\mathrm{cut}} = 0.005$, as at NLO. However, the improvement in going from $x_{\mathrm{cut}} = 0.001$ to $x_{\mathrm{cut}} = 0.005$ is not as great as at NLO. The dominant improvement in the quality of the fit is due to a much improved fit to the Tevatron jet data due to a readjustment of the high x gluon. This was also the case at NLO, but there was a far more significant improvement in the description of the NMC data and the $x > 0.005$ HERA structure function data in that case. As at NLO, each time the value of x_{cut} is raised there is a reduction in the very low x gluon, and a corresponding general increase in the high and moderate x gluon. However, the effect is far less pronounced at NNLO than it was at NLO, as seen in Fig. 3 which compares the default gluons to the $x_{\mathrm{cut}} = 0.005$ versions at both NLO and NNLO. Similarly, while $\alpha_S(M_Z^2)$ decreased as x_{cut} was lowered at NLO, it is almost completely insensitive to x_{cut} at NNLO. Hence, although the value of x above which we have complete confidence in our partons is the same at NNLO as at NLO, the changes compared to the default set, and hence the uncertainties involved with the default set, are much reduced at NNLO.

It is a similar story for Q^2 cuts. As at NLO there is a continual improvement until $Q_{\mathrm{cut}}^2 = 10$ GeV^2, above which stability sets in, and as at NLO this improvement is relatively gradual between $Q_{\mathrm{cut}}^2 = 2$ GeV^2 and $Q_{\mathrm{cut}}^2 = 10$ GeV^2, suggesting that higher twist is not most obviously the remaining correction to theory required. However, as with x cuts, the total improvement in the quality of the fit and the degree of readjustment of the partons is not quite as large as at NLO.

Finally, we investigate the effects of cuts on both x and Q^2. With our experience at NLO we would expect stability to be achieved when both x_{cut} and Q_{cut}^2 are near their respective individual values for stability. However, there is a slight improvement at NNLO as compared to NLO. In this case Δ_i^{i+1} is negligible when going from $Q_{\mathrm{cut}}^2 = 7$ GeV^2 to 10 GeV^2 at $x_{\mathrm{cut}} = 0.005$. However, $\Delta_i^{i+1} = 0.12$ when going from $x_{\mathrm{cut}} = 0.0025$ to 0.005 at $Q_{\mathrm{cut}}^2 = 7$ GeV^2, and similarly
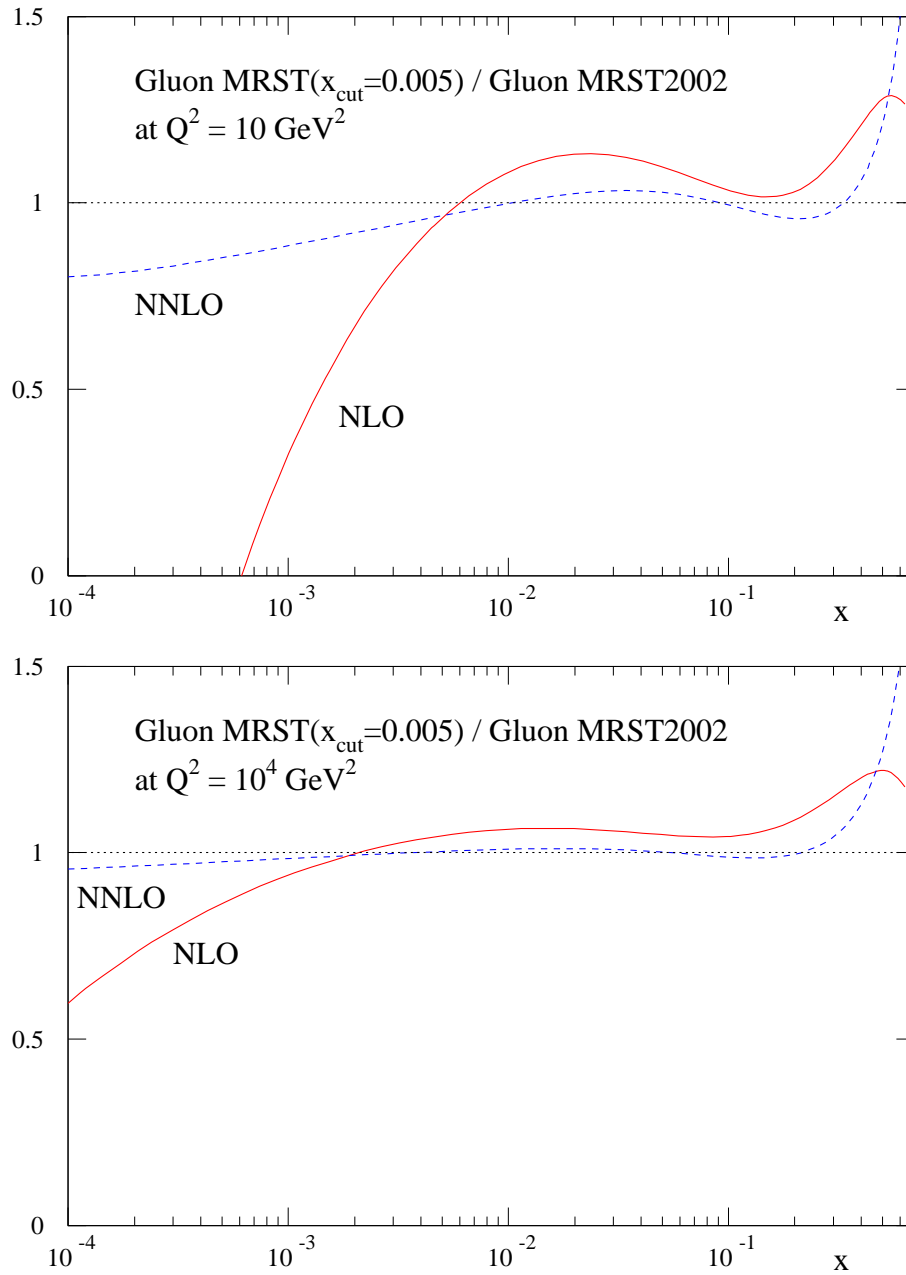
Figure 3: Comparison of MRST gluons obtained from $x_{\mathrm{cut}} = 0.005$ with the default gluons MRST(2002) at both NLO and NNLO.

if $Q^2_{\mathrm{cut}} = 10$ GeV^2. Hence, our conservative set of NNLO partons is obtained with the slightly less severe $Q^2_{\mathrm{cut}} = 7$ GeV^2 and $x_{\mathrm{cut}} = 0.005$. We also note that the modification of the gluon distribution for the MRST(cons) set is far smaller at NNLO than at NLO, being similar to the $x_{\mathrm{cut}} = 0.005$ alone set as shown in Fig. 3, and the value of $\alpha_S(M^2_Z)$ reduces only to 0.1153 ± 0.002, a much smaller reduction than at NLO. As at NLO, the error corresponds to the reduced tolerance $\Delta\chi^2 = 5$. The detailed comparison of the 'conservative' set with the default set of partons is shown in Fig. 4, and one sees that the deviation is rather smaller than at NLO, particularly at small x.

Hence, the investigation of the full range of cuts on the data that are fitted results in a set of 'conservative' NNLO partons which are only completely reliable over a slightly extended range compared to NLO. However, the change in these partons compared to the default set within this range, and more especially outside the range, is much smaller than at NLO (and the same is true for the change in $\alpha_S(M^2_Z)$). This implies that predictions made using NNLO partons are considerably more reliable than those using NLO partons, even before we consider the extra precision obtained simply by using the expressions for cross sections at a higher order. We will discuss this in more detail later.

4 Specific theoretical uncertainties

The results of Sections 2 and 3 imply that there may be significant theoretical corrections beyond NNLO in the standard DGLAP perturbative expansion, which become important at either low x, low Q^2 or low W^2 or some combination of these. Indeed there exist several types of theoretical correction which may be expected to have such effects. These include

(i) Higher powers of $\ln 1/x$ at higher orders in α_S, which become important at low x.

(ii) Increasing powers of $\ln(1-x)$ at higher orders in α_S, which are generally well understood, but which are intrinsically linked to higher-twist corrections. These are important at high x and hence low W, see (1).

(iii) Absorptive corrections which arise from parton recombination and which are higher twist in nature. These should become important at low x and Q^2.

(iv) Residual higher-twist contributions, which will be important at low Q^2. In practice these are often combined with the specific higher-twist contributions, already mentioned, in a purely phenomenological parameterization.

We have performed separate global analyses to investigate the effect of each of these in turn. The results are described below.
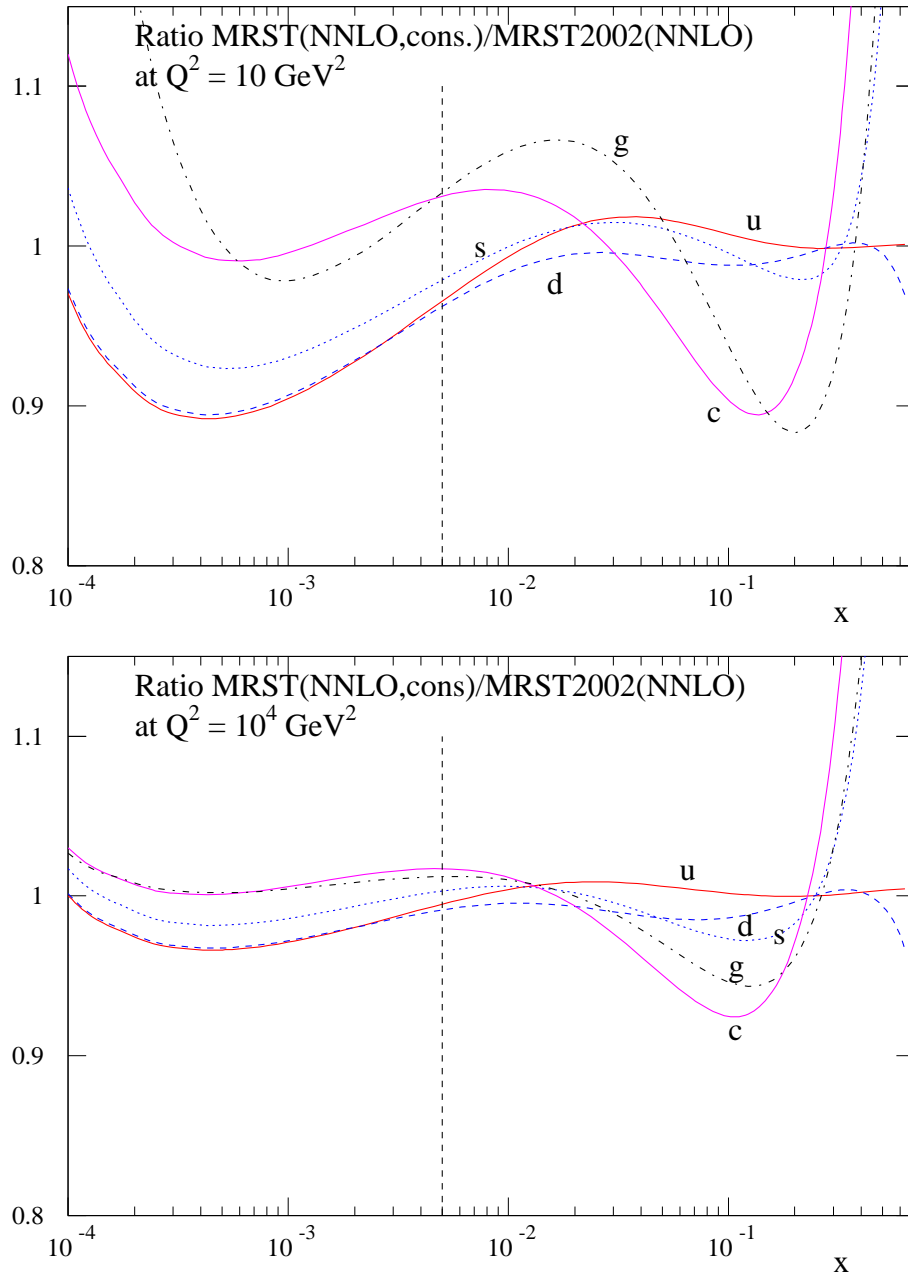
Figure 4: Comparison of the NNLO MRST(cons) partons with the default set MRST(2002) at NNLO. The former partons are only reliable for $x > 0.005$ and $Q^2 > 7$ GeV^2.

4.1 Contribution of higher order $\ln 1/x$ terms

It has long been known that the splitting and coefficient functions typically contain one additional power of $\ln 1/x$ for each additional power of α_S. Many of these terms are known explicitly [24, 25, 26]. However, a full $\ln 1/x$ resummation seems to involve many complications; for example, treatment of the running coupling, kinematic constraints, collinear resummations, etc. Various procedures exist for incorporating a full $\ln 1/x$ resummation but there is, as yet, no agreed prescription.

Hence, in the present study, we take a general approach. We include higher order corrections to the NLO splitting functions, with the correct maximum power of $\ln 1/x$, but we let the coefficient be determined by the global fit to the data. We begin by adding one additional term to P_{gg} and to P_{qg}. For this investigation we choose to include phenomenological $\alpha_S^4 \ln^3 1/x$-type terms of the form

$$P_{gg} \to P_{gg}^{\text{NLO}} + A\overline{\alpha}_S^4 \left(\frac{\ln^3 1/x}{3!} - \frac{\ln^2 1/x}{2!} \right) \tag{2}$$

$$P_{qg} \to P_{qg}^{\text{NLO}} + B\alpha_S \frac{n_f}{3\pi} \overline{\alpha}_S^4 \left(\frac{\ln^3 1/x}{3!} - \frac{\ln^2 1/x}{2!} \right), \tag{3}$$

where $\overline{\alpha}_S = 3\alpha_S/\pi$ and n_f is the number of active quark flavours. Both of the additional terms have been constructed so that momentum is still conserved in the evolution. We also add the same terms multiplied by $C_F/C_A = 4/9$ to P_{gq} and P_{qq} respectively. This factor of C_F/C_A is typical for the results in $\ln 1/x$ resummation [25].

The best NLO global fit, modified as in (2) and (3), gives an improvement in χ^2 of 21 compared to MRST2002 [10] and corresponds to $A = 3.86$ and $B = 5.12$. These are *effective* coefficients and should not be directly compared with the known coefficients from $\ln 1/x$ resummations, since they represent the effect of the two towers of $\ln^n 1/x$ terms. However, the values of A and B are of the magnitude expected from the partial information which exists. Hence the fit seems to benefit from such terms, which increase the speed of evolution at small x. The best-fit value of $\alpha_S(M_Z^2)$ decreases, but only very slightly. The input gluon turns out to be similar to that obtained in the fits when data below $x = 0.005$ are removed; that is, it is larger than the default gluon for large and moderate x, but even more negative at very small x. We can only assume that large positive $\ln 1/x$ terms in the coefficient functions will maintain the positivity of observables sensitive to the gluon, such as F_L [27]. Nevertheless, the increased gluon evolution at small x does result in a positive gluon more quickly as Q^2 increases.

We also investigated the effect of a more flexible parameterization of the $\ln 1/x$ resummation by introducing an additional term, both in (2) and in (3), of the type

$$C\overline{\alpha}_S^5 \left(\frac{\ln^4 1/x}{4!} - \frac{\ln^3 1/x}{3!} \right) \tag{4}$$

$$D\alpha_S \frac{n_f}{3\pi} \overline{\alpha}_S^3 \left(\frac{\ln^2 1/x}{2!} - \ln 1/x \right) \tag{5}$$

respectively. Again the form of the terms has been constructed so that momentum is conserved in the evolution. With the two additional parameters, we have an effective parameterization of both the N^{3}LO and N^{4}LO corrections at small x. The best fit now has a χ^2 that is 36 lower than MRST2002 with

$$A = -0.27, \quad B = 2.79, \quad C = 4.08, \quad D = 1.09. \tag{6}$$

The gluon distribution is very similar to the case with just two extra parameters, as is $\alpha_S(M_Z^2)$. However, neither the small x low Q^2 gluon distribution or the parameters A, B, C and D are determined very precisely, since there is some trade-off between them. For example, since the term in (4) falls off very quickly with Q^2 (since it behaves as $\alpha_S^5(Q^2)$), a more negative input gluon is largely compensated by a larger value of C, resulting in a largely unchanged gluon when one moves far away from the input scale $Q_0^2 = 1$ GeV^2. However, the generally positive input parameters and more negative input gluon is an unambiguous result. The fact that $\alpha_S(M_Z^2)$ is not altered much by the inclusion of terms which are important at small x is not surprising since it is always the evolution of the large x structure functions that has the dominant influence on the extracted value of $\alpha_S(M_Z^2)$, and this is not affected by this type of modification. The studies with up to four additional parameters were deemed sufficient to illustrate the possible effect of $\ln 1/x$ resummation on the NLO analysis.

However, we already know that the approximate NNLO splitting functions have a significant effect on the small x evolution. These contain one extra power of $\ln 1/x$ in comparison to the NLO splitting functions. Hence we repeated the above analysis at NNLO with the same four parameters, $A, \ldots, D$. The results were largely similar. The χ^2 improved by 39 from the standard MRST(NNLO) fit[4] [10], although the parameters take the values

$$A = -0.35, \quad B = -2.00, \quad C = 5.49, \quad D = 2.81, \tag{7}$$

very different from those at NLO, see (6). In addition they have a markedly reduced positive effect for the quarks and a very slightly increased positive effect for the gluon, showing that the NNLO contributions (positive for quark evolution at small x but negative for gluon evolution) have themselves been important at small x.[5] As at NLO, the gluon at high and moderate x is increased as compared to the unmodified NNLO fit, while at small x it is more negative. The gluons in these $\ln(1/x)$-modified fits are shown in Fig. 5. At low Q^2 they are indeed both much reduced compared to the default gluons at small x. However, the additional terms in the gluon splitting functions and the additional gluons at moderate x drive the small-x evolution more quickly than the default and at high Q^2 the $\ln(1/x)$-modified gluons are only a little lower than the default gluons at small x.

Hence, the distinct quality in the improvement in the global fit provides strong evidence that large $\ln 1/x$ contributions beyond even NNLO may be important in the description of the

---

[4]In practice we compared to the NNLO set of partons corresponding to the MRST2002 analysis [10].

[5]The effect is usually presented at quite high Q^2, e.g. 30 GeV^2, and with partons which are steep at small x [20], and is claimed to be quite moderate. At $Q^2 \sim 5$ GeV^2, with partons which are flattish at small x, the contribution from the NNLO splitting functions is proportionally much larger, and certainly very significant.
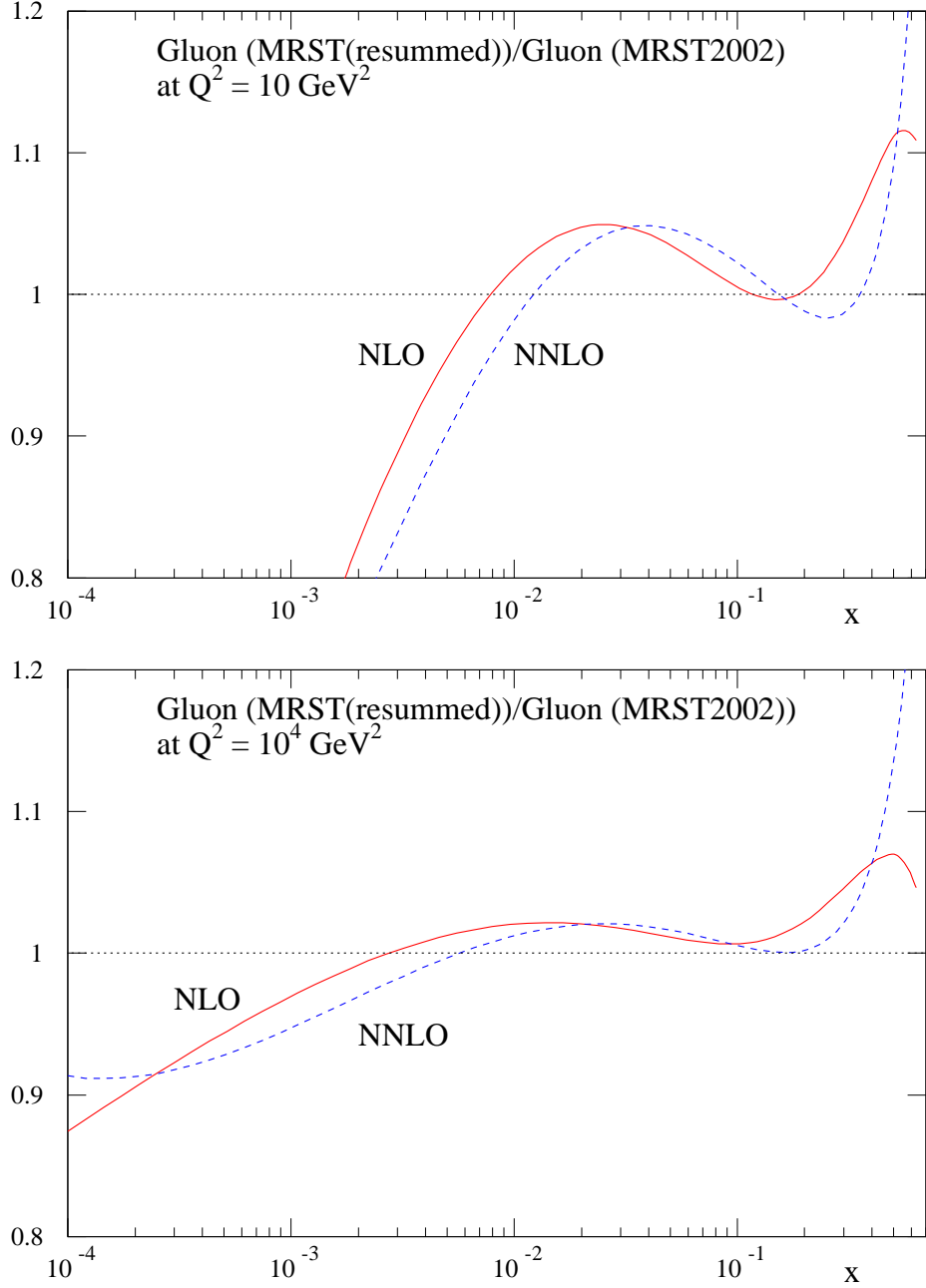
Figure 5: Comparison of MRST gluons obtained from the fits with additional resummation corrections in $\ln(1/x)$ with the default gluons MRST(2002) at both NLO and NNLO.

data. However, these corrections are empirically different in the NLO case than in NNLO, because the latter has its own large terms for $x \to 0$.

4.2 Contribution of higher order $\ln(1-x)$ terms

If we expand the quark coefficient function in powers of α_S, i.e.

$$C_{i,q}(x, Q^2) = \delta(1-x) + \sum_{m=1} \alpha_S^m(Q^2) c_{i,q}^{(m)}(x), \tag{8}$$

then the coefficient functions $c^{(m)}(x)$ and their moments $\tilde{c}^{(m)}(N)$ contain large logarithms as $x \to 1$ ($N \to \infty$) of the form

$$\left[\frac{\ln^{k-1}(1-x)}{1-x}\right]_+, \qquad \frac{(-1)^k}{k}(\ln^k N) \tag{9}$$

where $k = 1, \ldots, 2m$. Combining results on soft gluon resummation [28] and finite order results [17], Vogt has been able to provide explicit expressions for the first four coefficients [29] in an expansion of the form

$$\tilde{C}_{i,q}(N, Q^2) = 1 + \sum_{m=1}^{\infty} \alpha_S^m(Q^2)(c_{m1} \ln^{2m} N + c_{m2} \ln^{2m-1} N + c_{m3} \ln^{2m-2} N + c_{m4} \ln^{2m-3} N + \cdots). \tag{10}$$

Hence, we have detailed knowledge of the leading terms in the coefficient functions for the large x or N limit at all orders, and it is argued in [29] that a more efficient convergence is achieved if the leading terms are arranged as powers of $\ln N$ rather than $\ln(1-x)$. In principle all of the terms in (10) should be included. However, the investigation in [29] suggests that, unless one is at very high x, or equivalently very low W^2, going to a finite order is sufficient. Indeed, the suggestion is that for $m = 3$ the coefficient function is only important above $x \sim 0.7$, and those for $m > 3$ are only important for $x > 0.8$.

In order to investigate this we performed a fit using the NNLO splitting functions and coefficient functions, and including also the $\mathcal{O}(\alpha_S^3)$ contribution to the coefficient function in (10).[6] When using our W^2 cut of 12.5 GeV^2, we find that the effect of the $\mathcal{O}(\alpha_S^3)$ coefficient function at very high x is almost negligible, since the W^2 cut ensures we are at quite high Q^2 for very large x, and $\alpha_S^3(Q^2)$ is fairly small. In fact the most significant effect of the $\mathcal{O}(\alpha_S^3)$ coefficient function is at higher x. Since the coefficient function has a vanishing first moment, its positive effect at very high x must be countered by a negative contribution at lower x. In practice it increases the structure function for $x > 0.55$, but decreases it for $x < 0.55$. This decrease is not large, but it affects much more data. In practice the best fit, when including the approximate NNNLO coefficient function, is very slightly worse than the usual NNLO fit, although the partons and the value of $\alpha_S(M_Z^2)$ are hardly changed at all.

---

[6]In [29] it is demonstrated that the four terms in (10) are a very good approximation to the full NNNLO coefficient function, which can be estimated in detail using the same sort of techniques as in [20].

From [29] it is clear that the contributions in (10) beyond $\mathcal{O}(\alpha_S^3)$ are only important at even higher x. Hence, we conclude that if we use $W_{\text{cut}}^2 = 12.5$ GeV^2, there is no advantage to be gained in including terms beyond the NNLO coefficient function. As we go lower in W^2, however, the NNNLO coefficient function does start to have a non-negligible effect. We will discuss this in more detail in our analysis of higher-twist corrections.

Indeed, with reference to higher twist, we should note that there are a number of ambiguities when performing large x resummations, not just whether to resum large logarithms in N or $(1-x)$. We note that the series expansion in (10) is convergent. However, one could alternatively calculate the first two towers of terms in the expansion

$$\frac{d \ln \tilde{C}_{i,q}(N, Q^2)}{d \ln Q^2} = \sum_{m=1}^{\infty} \alpha_S^m(Q^2)(c_{m1} \ln^m N + c_{m2} \ln^{m-1} N + \cdots), \tag{11}$$

which would give the dominant large-N contribution to an effective anomalous dimension for structure function evolution. In this case the series would have finite radius of convergence. This badly-defined series shows that there are higher twist corrections present, and the ambiguity in the series is taken as an estimate of the size of the higher-twist corrections in renormalon models. The divergence in (11) is reflected in the terms in (10) beyond c_{m4} becoming extremely large. Strictly speaking one cannot simply perform a large $\ln(1-x)$ expansion without encountering this interplay with higher-twist corrections, and beyond about NNNLO it is difficult to disentangle the two.

4.3 Absorptive effects

To investigate the effects of absorption[7] we include bilinear terms in evolution equations as follows

$$\frac{\partial(xg(x, Q^2))}{\partial \ln Q^2} = \ldots - 3\frac{\alpha_S^2(Q^2)}{R^2 Q^2} \int_x^1 \frac{dx'}{x'} [x'g(x', Q^2)]^2 \tag{12}$$

$$\frac{\partial(xq(x, Q^2))}{\partial \ln Q^2} = \ldots - \frac{1}{10} \frac{\alpha_S^2(Q^2)}{R^2 Q^2} [xg(x, Q^2)]^2 \tag{13}$$

These terms take into account the dominant small x contributions at lowest order in α_S, as calculated by Mueller and Qiu [30]. We have approximated the two-gluon correlation function as

$$g^{(2)}(x, Q^2)) = \frac{2}{3\pi^2 R^2} [g(x, Q^2)]^2 \tag{14}$$

as estimated in [30]. We consider two choices of R^2, namely $R^2 = 15$ and 5 GeV^{-2}. The former represents the naïve assumption that R is of the order of the proton radius, whereas the latter represents much stronger absorption motivated by 'hotspot' studies which allow a large contribution from the diagrams responsible for saturation where both gluon ladders couple to the same parton [31]. Of course, inclusion of the absorptive terms leads to a violation of

---

[7]We thank M.G. Ryskin for discussions.

18

momentum conservation. In principle a more complete theoretical treatment would correct for this, but in this study we account for the effect by starting the evolution at low scales with the partons carrying slightly more than 100% of the proton's momentum. The precise amount is chosen so that at high Q^2, when the absorptive corrections have completely died away, the correct value of 100% is obtained. In practice for $R^2 = 15$ and 5 GeV^{-2} we input a total momentum of 100.7% and 103% respectively at $Q_0^2 = 1$ GeV^2.

The effect of absorptive corrections in global parton analyses was investigated many years ago [32]. The corrections were found to have a sizable, observable effect in the small x region accessible to HERA for a parton set B$_-$ in which the gluon was assumed to have a small x behaviour of the form $xg \sim x^{-\frac{1}{2}}$ at the then input scale of $Q_0^2 = 4$ GeV^2. However, with the advent of the HERA data, the size of the small x gluon is now known to be considerably smaller than that of the B$_-$ set, and hence we may anticipate the shadowing effects will be much smaller.

Here we make two alternative studies. First, we repeat the NLO global analysis, with the modifications shown in (12) and (13), starting from MRST2002 partons [10]. For the $R^2 = 15$ GeV^{-2} choice there is almost no change in the fit, the best fit having an essentially identical χ^2. There is no significant change in the partons and $\alpha_S(M_Z^2)$ increases by less than 0.001. For $R^2 = 5$ GeV^{-2} there is a slight deterioration in the best fit of $\Delta\chi^2 \simeq 20$. This is accompanied by an increase in $\alpha_S(M_Z^2)$ to 0.1225 and a slight increase in the small x gluon. This shows that the slowing of the evolution by the absorptive corrections is not consistent with the data, and the increase in α_S is necessary to compensate partially for this.

However, starting from the MRST2002 partons, which have a negative gluon for low x and Q^2, may be regarded to be inconsistent with an absorptive approach, which really assumes a positive input. Hence, as an alternative investigation, we force the input gluon to be positive-definite and, indeed, slowly increasing with decreasing x. Since we know that this causes conflict with the low Q^2 data we also raised the Q^2 cut to $Q_{\text{cut}}^2 = 5$ GeV^2. However, after performing the fits we investigated the extrapolation to lower Q^2.

In detail, we removed the 'negative' input term $[-A_-(1-x)^{\eta_-}x^{-\delta_-}]$ from the gluon, and set the (conventional) small x power $\delta_g = -0.1$, hence ensuring a positive definite starting gluon[8]. For $R^2 = 15$ GeV^{-2} the best fit, for data above $Q^2 = 5$ GeV^2, has $xg(x, Q_0^2) \sim 0.87x^{-0.1}$ at small x and is over 200 worse in χ^2 than MRST2002, most of this deterioration coming from the HERA data. The evolution at low x is far too rapid, even though $\alpha_S(M_Z^2)$ reduces to 0.117. However, there is also some worsening to the fit to high x data due to the gluon at high x being smaller than before. If we extrapolate the fit below $Q^2 = 5$ GeV^2 then the description of the HERA, and even NMC, data becomes very poor indeed, see Fig. 6.

For $R^2 = 5$ GeV^{-2} the increased absorptive corrections moderate the problems at lowest x and $xg(x, Q_0^2) \sim 0.93x^{-0.1}$, i.e., a little larger, and the global χ^2 is 80 worse than for MRST2002

---

[8]It was also necessary to fix ε_g to some value, which in practice was chosen to be 0.74, in order to prevent the interplay between parameters resulting in an effectively valence-like input gluon distribution.
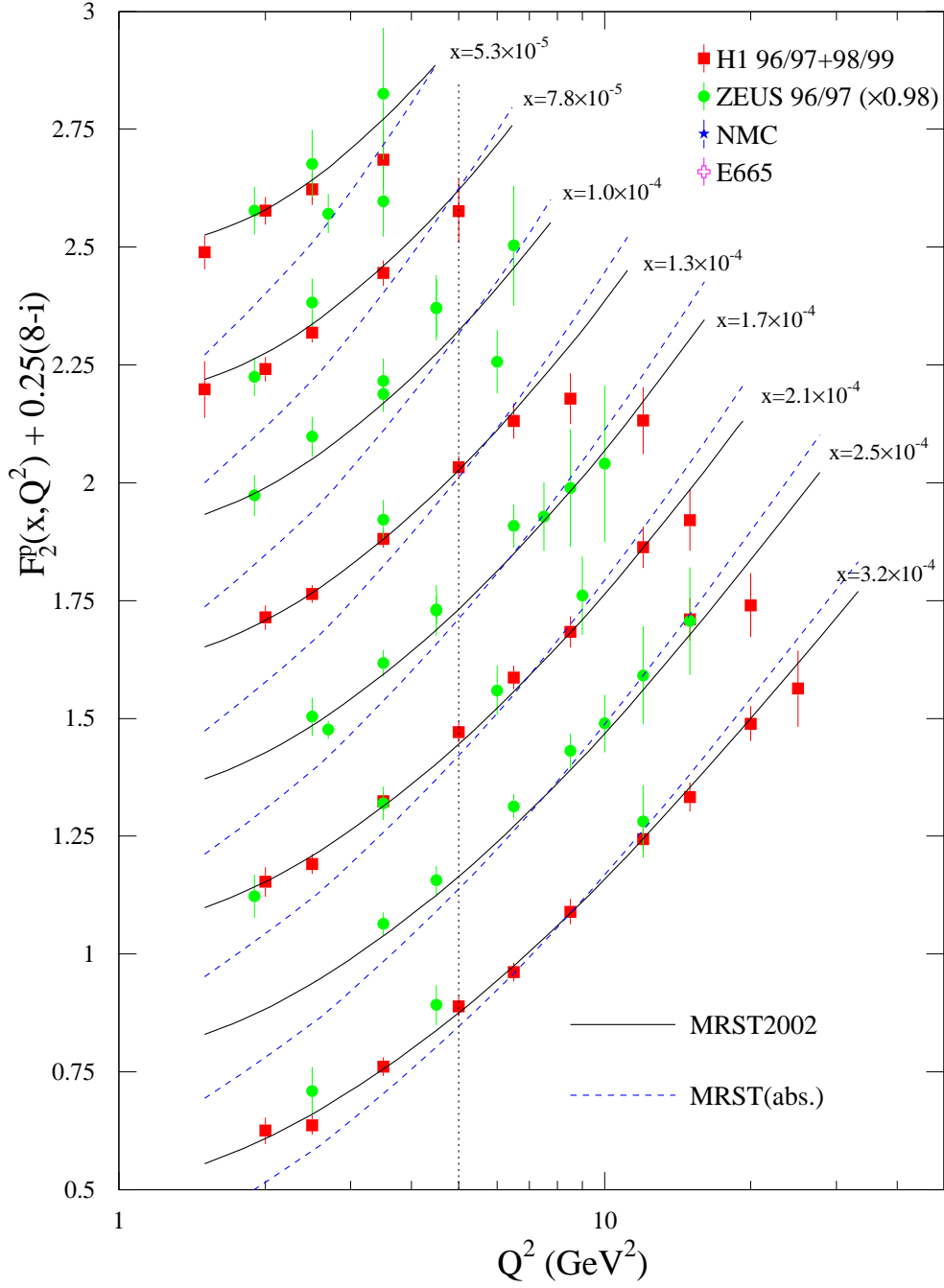
Figure 6: The comparison with data of the fit with a flattish input gluon and saturation corrections with the curves extrapolated below the region of validity of the fit.

with $\alpha_S(M_Z^2) = 0.118$. However, the extrapolation to low Q^2 is still nearly as bad as in the previous case.

Hence we conclude that if we demand a positive-definite input gluon, which is even slowly increasing as x decreases, then absorptive corrections are not sufficient to compensate the faster evolution when compared to a negative input (MRST2002) gluon[9], and also the necessary reduction of the larger x gluon has a detrimental effect on the fit.

It was hoped that the introduction of absorptive corrections would lessen, or perhaps remove, the need for a negative gluon at very low x and low Q^2. However we see from the above studies that the global fits do not appear to favour the introduction of absorption corrections. This result is surprising [34]. We know that at low x about 10% of F_2 arises from diffractive events. However we cannot simply subtract F_2^D from the inclusive F_2, since the diffractive events which originate from low scales are already accounted for in the parameterization of the starting distributions at Q_0. The diffractive contribution, ΔF_2^D, arising from higher scales is however related to the absorptive correction ΔF_2 to the inclusive F_2. The relation is given by the AGK cutting rules [35]. When $\Delta F_2 \ll F_2$, the relation takes the form

$$\Delta F_2 \simeq -\Delta F_2^D. \tag{15}$$

So some absorptive correction is expected to be present. To quantify the amount will require an enlarged global analysis incorporating the diffractive structure function data.

4.4 Higher-twist effects

In previous papers ([16, 21]) we have examined the effect of including in global fits a simple phenomenological parameterization of the higher-twist contribution in the form

$$F_i^{\text{HT}}(x, Q^2) = F_i^{\text{LT}}(x, Q^2) \left(1 + \frac{D_i(x)}{Q^2} \right), \tag{16}$$

where in practice $D_i(x)$ is taken to be a constant, independent of Q^2, in each of a number of different bins in x. In the fits which include such a parameterization, we have lowered the Q^2 cut to 1.5 GeV^2 and the W^2 cut to 4 GeV^2. Here we repeat the procedure for our NLO and NNLO fits with the most up-to-date data, and also include a fit which has the approximate NNNLO coefficient function, as discussed in Section 4.2. This only alters the NNLO results at high x. We give results for a LO fit, although we have already seen in [21] that such a fit simply fails in many regions of parameter space, and the invoked higher-twist corrections are simply mimicking, as best as they can, corrections which really reproduce the NLO (and possibly higher) leading-twist contributions. In particular, we see that the large negative higher-twist corrections at small x decrease significantly at NLO and even tend to disappear altogether at higher orders. The results are summarised in Table 3.

---

[9]This result is in conflict with observations of Ref. [33] in which only LO partons are considered.

x	LO	NLO	NNLO	NNNLO
0–0.0005	-0.07	-0.02	-0.02	-0.03
0.0005–0.005	-0.03	-0.01	0.03	0.03
0.005–0.01	-0.13	-0.09	-0.04	-0.03
0.01–0.06	-0.09	-0.08	-0.04	-0.03
0.06–0.1	-0.02	0.02	0.03	0.04
0.1–0.2	-0.07	-0.03	-0.00	0.01
0.2–0.3	-0.11	-0.09	-0.04	0.00
0.3–0.4	-0.06	-0.13	-0.06	-0.01
0.4–0.5	0.22	0.01	0.07	0.11
0.5–0.6	0.85	0.40	0.41	0.39
0.6–0.7	2.6	1.7	1.6	1.4
0.7–0.8	7.3	5.5	5.1	4.4
0.8–0.9	20.2	16.7	16.1	13.4

Table 3: The values of the higher-twist coefficients D_i of (16), in the chosen bins of x, extracted from the LO, NLO, NNLO and NNNLO (NNLO with the approximate NNNLO non-singlet quark coefficient function) global fits.

There are a number of conclusions which can be drawn from the Table. First, we can see that there is no clear evidence for any significant higher-twist contributions for $x < 0.005$. Even though it may be argued that the form of the higher-twist corrections at small x is rather more complicated than the simple parameterization of (16) (e.g. [36]), they would have to be of roughly the same qualitative form, and the lack of any indication of them appears compelling.

At NLO there is a strong indication of a negative higher twist contribution for $x \sim 0.005-$ 0.06. This is required in order to make $dF_2(x, Q^2)/d\ln Q^2$ large enough for the NMC data in this region. However, the sizes of the D_i's in this region decrease significantly when going to NNLO, because both the coefficient functions and splitting functions at NNLO lead to increased evolution in this range, and the evidence for higher twist at NLO seems to be really an indication of a lack of important leading-twist, higher-order corrections.

The main higher-twist corrections appear, as expected, at high x. For $x \sim 0.1-0.4$ there is a slight indication of a negative higher-twist correction at NLO, but this diminishes at NNLO and effectively disappears at NNNLO. Hence this is presumably just an indication that leading-twist perturbative corrections are important. A transition is apparent for $x \sim 0.4-0.5$, and for $x > 0.5$ there is a definite positive higher-twist contribution. However, this contribution has a tendency to decrease from one order to the next. Indeed at NNNLO the required higher-twist contribution at very high x is similar to that achieved simply from target-mass corrections [37]. Moreover, in this very high x domain, the correction to the higher-twist coefficient when going

from NNLO to NNNLO is as large, if not larger, as when going from NLO to NNLO (the NLO to NNLO correction is smaller than the LO to NLO). It is easily verified that this is indeed because the NNNLO correction to the structure function is as large at NNNLO as at NNLO for this range of x and Q^2. This shows that the divergent high-x perturbative series discussed in Section 4.2 reaches its minimum at about NNNLO, and this represents the essential ambiguity in this series. Hence the renormalon contribution to the higher twist [38], which comes from this ambiguity of the perturbative series, is roughly of the same size as the NNNLO contribution (at least in the region of parameter space we are probing). Indeed, the prediction for the higher-twist contribution as a function of x in [38] is very similar to that obtained from the approximate NNNLO contribution. This implies that at very high x it is pointless to go beyond NNLO (or certainly NNNLO) in the perturbative series, since at this order the perturbative series and higher-twist corrections become indistinguishable. It is important to realize, however, that this is a special feature of high x, i.e., low W^2, and does not imply the same is true for other regimes of x and Q^2.

To conclude, by going to higher and higher orders in the leading-twist perturbative expansion we see that the only strong evidence for higher-twist corrections is in the region of high x and low W^2. All shortcomings of low-order perturbative calculations seem to be reduced, if not removed, simply by working to higher orders. Moreover, our particular knowledge of high x coefficient functions leads us to believe that we have reached the stage where the perturbative expansion and higher twist corrections are not really separable. This is further illustrated by a method which goes beyond the standard logarithmic resummation using the 'dressed gluon exponentiation' approach of [39], which takes into account some all-order information on the kernel of Sudakov resummation itself. In Ref. [40], this combined resummation of Sudakov logarithms, renormalon contributions and higher twists was confronted by data on the Nachtmann moments (i.e. corrected for target mass) of the structure function extracted from low Q^2 data. While it is impossible to estimate precisely the higher twist contributions, one allowed description is that where the Sudakov resummation alone is necessary to explain the high-moment data.

5 Uncertainties due to input assumptions

5.1 Choice of input parameterization

Over the years we have adopted parameterizations with more and more free parameters, as required by the increase in precision and kinematic range of the data, and by the new types of data that have become available. Perhaps the biggest single extension of the parameterization was the contribution added to the gluon that allowed it to become negative at small x [21]. Recently other groups have investigated the influence of parton parameterizations on parton uncertainties [6, 8], with various conclusions. In [8], which describes a fit to a fairly limited set of data, it is found that there is no real improvement to the quality of the fit once the number

of parameters specifying the input parton distributions has increased past a certain number, 10 in this case. In [6], however, a new type of parton parameterization is used, and it is claimed that this is necessary in order to obtain the best fit to the Tevatron jet data. Our studies do not support this claim.

Although the fit to the Tevatron jet data in [6] is indeed better than in [15], and even in [10] (where the high x gluon has been improved), we believe very little of this is due to the parton parameterization. There are various differences between the CTEQ6 and MRST2002 approaches to global fits including the choice of Q^2 cuts, the data sets included in the fit, the treatment of errors (particularly those of the E605 Drell-Yan data) and even the definition of $\alpha_S(Q^2)$ at NLO. In [41] we discussed in detail the effect of making our starting point for the fit more and more like that for CTEQ6, and discovered that when we did so our fit to the jet data became of almost similar quality to that of CTEQ6 when using our own parameterization (without any of the "kinks" found in some previous best fits to jet data [15]). Some of the further cuts we have introduced in the course of the investigations in this paper have led to even better fits to the Tevatron jet data. Hence, we do not feel that these particular data require us to alter our parton parameters to obtain the best possible fits. However, we do note that the fact that our parameterization does allow the input gluon to be negative does have important consequences compared to the CTEQ analysis. It means that the CTEQ6 gluon is always larger at very small x than ours (even though their starting scale is a little larger, $Q_0^2 = 1.69$ GeV^2, our gluon is still negative at this Q^2), and from the momentum sum rule must be smaller elsewhere, in practice at intermediate x. This does lead to predictions which are significantly different to ours, examples being the Higgs cross section at the Tevatron and LHC as seen in Fig. 15 of [10].

Indeed, there is considerable evidence that we have, if anything, too much flexibility in the parton parameterizations. When attempting to diagonalize the error matrix for partons when all the parameters were left free we discovered that many were extremely correlated (or anti-correlated) leading to some very flat directions in the eigenvalue space [10]. In fact, in order to obtain a stable error matrix we found that it was necessary to observe the departures away from the best fit while allowing only 15 of our nominal 24 parton parameters to vary. A similar effect has, in fact, also been seen in [5], where when producing the error matrix only 16 out of a possible 22 parameters are allowed to vary, and in [6], where only 20 out of 26 parameters are allowed to vary. Similar problems are not seen by other groups [7, 8, 9, 3], all of whom use smaller data sets and a smaller number of parton parameters, implying that they have either the correct number of parameters for the flexibility required by their data, or potentially slightly too few.[10]

The redundancy observed when calculating the error matrix does not mean that all the other parameters can simply be dispensed with, since it is necessary to allow most to vary in

---

[10]As mentioned above, in [8] the number of required parameters is carefully investigated. In [3], however, one can check that more than 2σ variations in the parton parameters would lead to some pathological behaviour (particularly for the gluon), implying that the fit is on the verge of the same type of redundancy problems as in [10] and [5, 6].

order to obtain the best fit itself, and to allow the partons to have a sufficiently flexible shape. It just means that not all the parameters need to vary in order to investigate small deviations from the best fit partons. For example, starting with our previous parameterization for the gluon

$$xg(x, Q_0^2) = A_g(1 - x)^{\eta_g}(1 + \epsilon_g x^{0.5} + \gamma_g x)x^{\delta_g}, \tag{17}$$

it was necessary to add a term of the form

$$-A_-(1 - x)^{\eta_-}x^{\delta_-}, \tag{18}$$

in order to let the input gluon be negative at small x. Only 3 parameters were needed to investigate small variations, but more parameters are certainly needed in order to obtain the gluon for the best fit. However, the total of 7 free parameters we have for the gluon (A_g is fixed by the momentum sum rule) do have some genuine degree of redundancy. We have noticed in the course of performing many fits that, for any fixed value of ϵ_g from -1 to 3, we can obtain optimum descriptions of the data which are practically identical in quality. In these fits all gluon parameters are significantly different, but the gluon distributions produced are practically identical (at least between $x = 0.9$ and $x = 0.00001$). This is perhaps the clearest example of a single parameter which is essentially redundant, but we have other similar examples.

Hence, we conclude that our input parameterizations are sufficiently flexible for the present data. We do not seem to have the optimum parameterization for both finding the best fit and also investigating fluctuations about this best fit. For a fully *global* fit, however, no-one else seems to have it either. To achieve this would require fewer parameters than at present. This might then influence our error analysis using the Hessian approach somewhat, but we feel it is unlikely to affect our best fit partons very much at all.

5.2 Choice of heavy target corrections

When fitting the CCFR $F_2^\nu(x, Q^2)$ and $F_3^\nu(x, Q^2)$ data [42] we have to use some model for nuclear shadowing corrections. The form that we use for the heavy target correction factor is deduced from a Q^2-independent fit to the EMC effect for the scattering of muons on a heavy nuclear target (A=56). To be explicit, we parameterize the correction as

$$R_{\mathrm{HT}} = \begin{cases} 1.238 + 0.203 \log_{10} x & \text{for } x < 0.0903 \\ 1.026 & \text{for } 0.0903 < x < 0.234 \\ 0.783 - 0.385 \log_{10} x & \text{for } 0.234 < x. \end{cases} \tag{19}$$

In order to investigate the uncertainty arising from this correction we perform a series of fits, maintaining the central plateau in R_{HT}, but changing the slopes in $\ln x$ in the high and low x regions. In addition, we allow the normalisation to vary within the experimental error, as usual. We study the quality of the fit as a function of the value of the heavy target correction at the lowest x value ($x = 0.0075$), for which CCFR data exist. This data point receives the
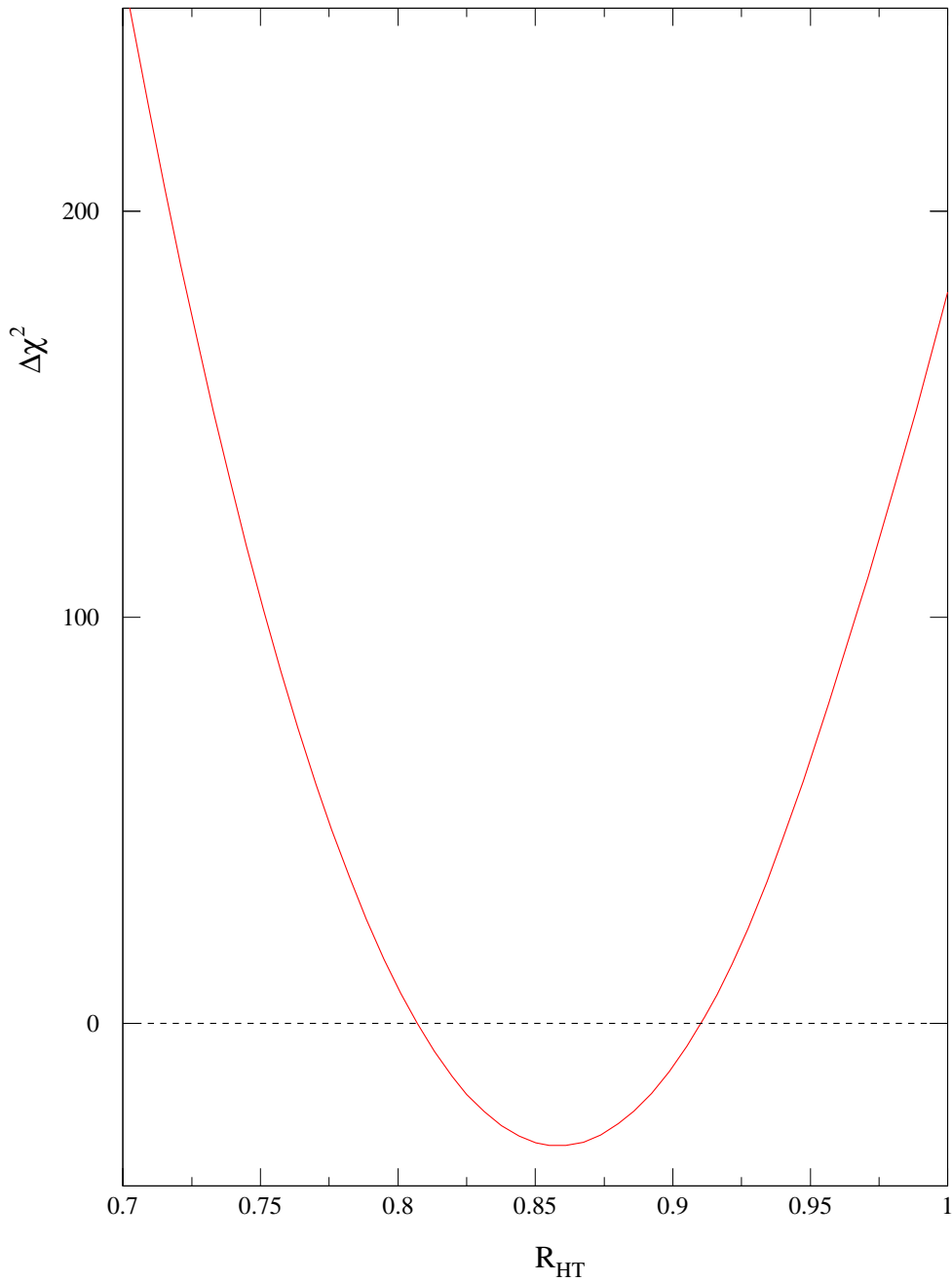
25

Figure 7: The variation of χ^2 for global fits with different heavy target corrections to the CCFR neutrino data. χ^2 is plotted against the value of the heavy target correction R_{HT} at $x = 0.0075$, as explained in the text. The default fit has $R_{HT} = 0.807$, whereas for the optimum fit the correction is slightly less, $R_{HT} = 0.86$.

largest heavy target correction, that is $R_{HT} = 0.807$ for our standard fit. This is a very similar, but larger, correction to that for the highest x value ($x = 0.75$) for which data exist. Fig. 7 shows the global χ^2 as a function of the value of R_{HT} at $x = 0.0075$. Clearly our standard shadowing assumption, although near to the minimum, is not the absolutely optimum choice of shadowing correction. The data prefers a value of $R_{HT}(x = 0.0075)$ of 0.86, i.e. a slightly smaller correction than our usual choice, and the improvement in the fit is about 30 units of χ^2. Most of this improvement comes from the fit to the CCFR data on $F_2^{\nu(\bar{\nu})}(x, Q^2)$, which has always had a tendency to lie underneath the theory when shadowing corrections are applied (see for example Fig. 9 of [15]). The main effect of the slightly reduced shadowing is just to bring these data in line with the theory (although there are some other minor improvements) and the parton distributions themselves are not changed much at all, as shown in Fig. 8.[11]

Similarly the deuterium data is corrected for shadowing effects. These prescriptions are not unique. In our fits we normally use the parameterization in [43] which uses a theoretical model to estimate the nuclear shadowing corrections at relatively small x. There are other alternatives for models of this type of shadowing. There has also been a prescription for deuterium shadowing corrections extracted by the SLAC E139/140 experiments [44]. This is an empirical extraction which uses a model [45] in which binding effects are assumed to scale with nuclear density. This gives a relatively large shadowing correction for deuterium, especially at large x, and was used as a basis for an analysis of the d/u ratio in [46]. The validity of this extraction is rather controversial (see e.g. [47]), but we use the results simply as an estimate on the uncertainty of the deuterium shadowing corrections. As such, a fit with this shadowing correction applied to the deuterium data gives an estimate of the model uncertainty from this source on the high x valence partons, particularly $d(x, Q^2)$.[12]

This particular correction leads to a larger neutron structure function at high x. Hence, the largest change in the parton distributions is in the high x down quark distribution, which increases significantly. As seen in Fig. 9 this increase is a little smaller at high x than the uncertainty due to errors on experimental data that was estimated using the Hessian matrix method in [10]. From the constraint on the number of valence quarks the increase in $d_v(x, Q^2)$ at high x must be compensated for elsewhere, and indeed we see that it is smaller than the default for $x = 0.01 - 0.1$, but is well within the bounds of the experimental uncertainty. The change in the $u_v(x, Q^2)$ distribution must lead to a modification of the $u_v(x, Q^2)$ distribution in order to obtain the best overall global fit. This modification is shown in Fig. 10. Although it is proportionally much smaller than the change in $d_v(x, Q^2)$, $u_v(x, Q^2)$ is much better constrained by data, and the change due to the different shadowing correction can be as big as, or even slightly larger than the uncertainty due to experimental errors on data. However, this is mainly

---

[11]The preliminary measurements of $F_2^{\nu(\bar{\nu})}(x, Q^2)$ by the NuTeV collaboration actually tend to lie a little above those of CCFR in the lowest x bins , and hence lie rather closer to the default theory curves with our standard shadowing correction [49].

[12]We note that an examination of theoretical uncertainties due to nuclear effects in the deuteron has recently been studied in [48] with the aim of examining isospin depedence of higher twist corrections.
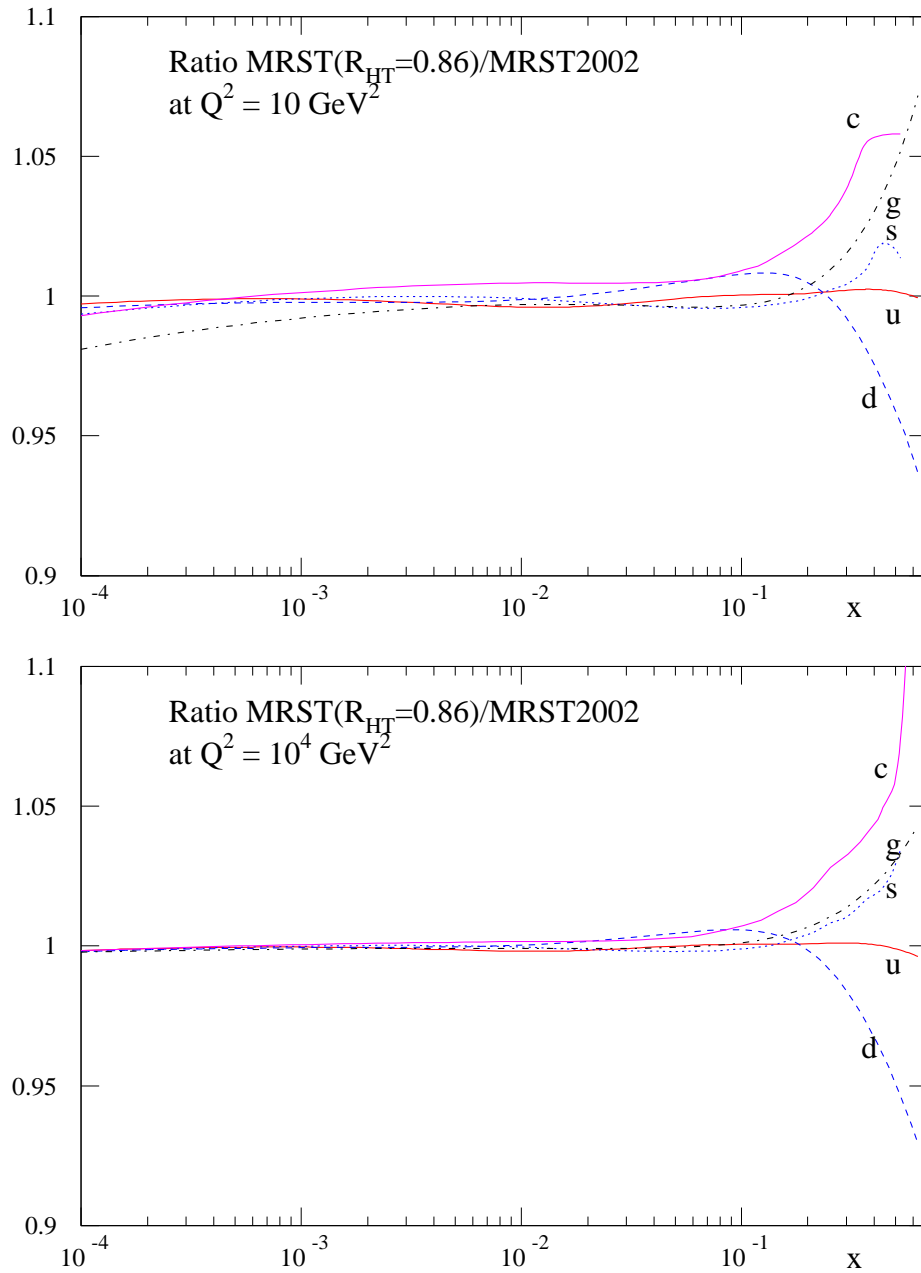
Figure 8: The parton distributions obtained from the fit with the optimum shadowing correction of $R_{\text{HT}}(x = 0.0075)$ of 0.86 compared with the default MRST2002 partons.
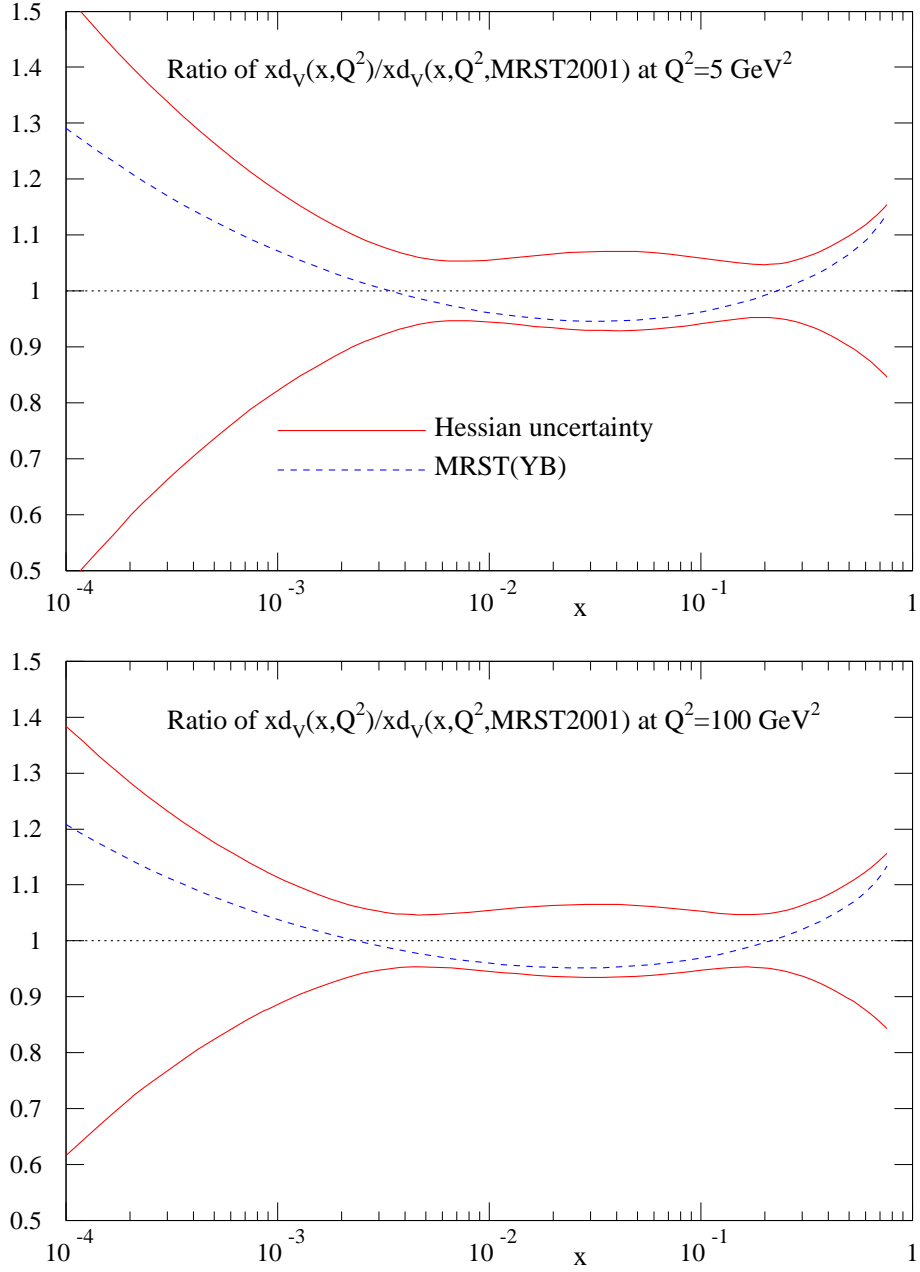
Figure 9: The ratio of $d_v(x, Q^2)$ obtained from the fit with the deuterium shadowing correction of [44] compared with the default MRST2002 partons, and also with the uncertainty in the default $d_v(x, Q^2)$ distribution from errors on experimental data found in [10].
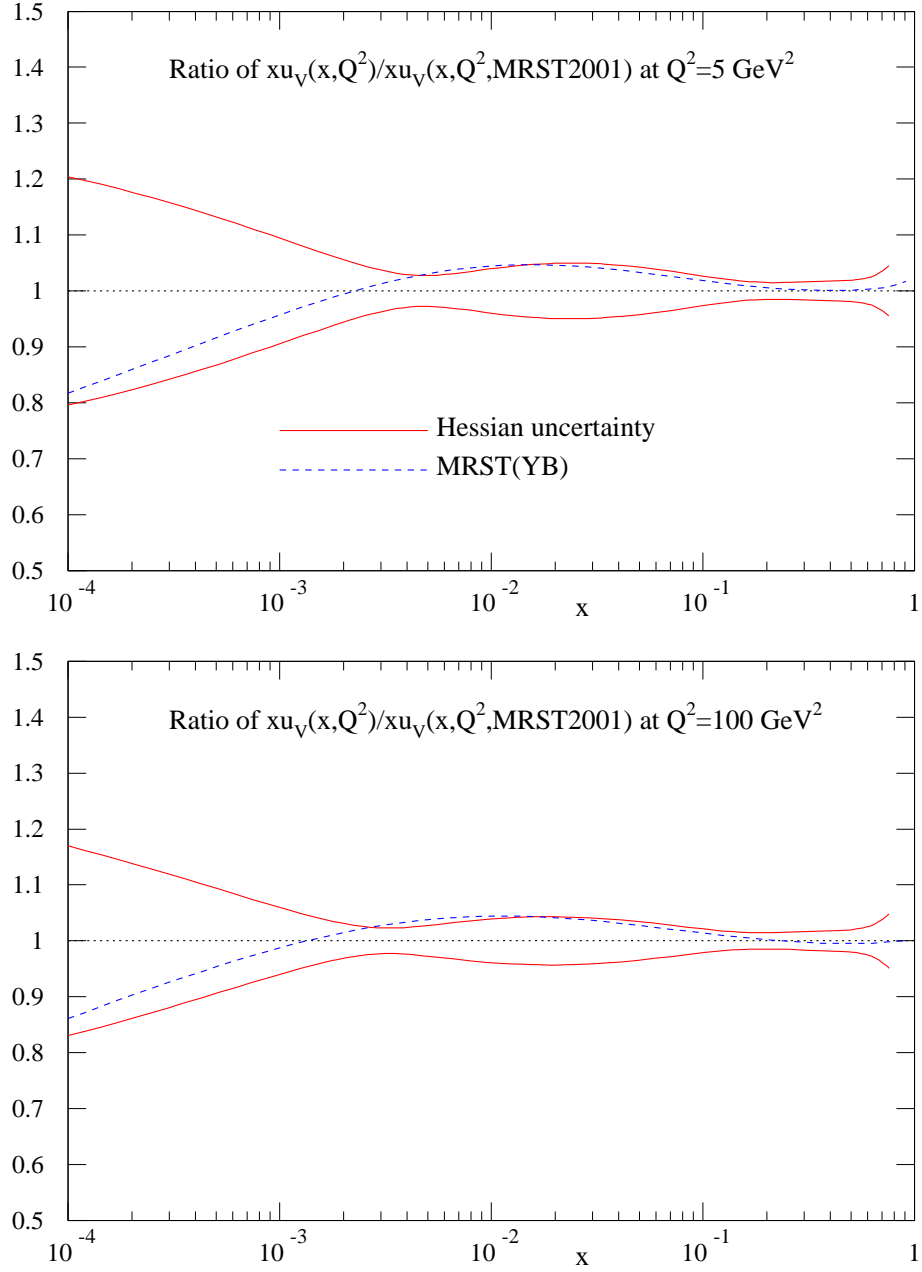
Figure 10: The ratio of $u_v(x, Q^2)$ obtained from the fit with the deuterium shadowing correction of [44] compared with the default MRST2002 partons, and also with the uncertainty in the default $u_v(x, Q^2)$ distribution from errors on experimental data found in [10].

so at $x < 0.1$ where the direct constraint on valence quarks is quite small since sea quarks dominate, and the estimated uncertainty from the Hessian method for valence quarks alone is likely to have limitations related to parameterizations. Hence, for the down and up quark distributions we conclude that model errors on shadowing corrections in deuterium are typically of the same size as the errors due to the experimental errors on deuterium structure function data.

Finally we note that the quality of the global fit when using the deuterium corrections in [44] is 12 units in χ^2, better than the default fit. There is a slight improvement in the total fit to deuterium data, mainly to the BCDMS data, but a slight deterioration in the fit to E605 Drell-Yan data. There is also a slight improvement in the fit to high-E_T Tevatron jet data. This is to be expected since the increased down quark distribution at high x leads to an increase in the highest E_T jet cross-section, which is what is required. However, the improvement is limited since an increase in $d(x, Q^2)$ at high x increases the momentum carried by the down quark. From the momentum sum rule this makes it harder to have a large gluon distribution at high x, which the fit would like. In practice a compromise is reached, i.e. an even greater enhancement in $d_{\rm v}(x, Q^2)$ at high x would slightly improve the fit to deuterium, and so other data, but actually makes the fit to Tevatron jet data worse. The relatively small improvement in the global χ^2 from this model of deuterium shadowing, partially limited by tension between different data sets and parton distributions, leads us to conclude that there is no strong supporting evidence for the model. Nevertheless, the small pull of the evidence is in favour of some deuterium shadowing at high x.

5.3   Size of input strange sea

Global fits have traditionally assumed that the shape of the input strange quark sea distribution is the same as the average of the input $\bar{u} + \bar{d}$ distribution. The primary experimental constraint on the strange distribution is provided by data on dimuon production in neutrino–nuclei deep inelastic scattering [50, 51]. These data are consistent with the shape assumption, and in addition constrain the magnitude of the strange distribution to be approximately half of $(\bar{u} + \bar{d})/2$ at low Q^2. For this reason, we conventionally choose

$$s(x) = \frac{1}{4}(\bar{u}(x) + \bar{d}(x)) \tag{20}$$

at $Q_0^2 = 1$ GeV^2. We now investigate the sensitivity to this input assumption.

First, we check the validity of our choice by expressing

$$s(x) = \kappa(\bar{u}(x) + \bar{d}(x))/2 \tag{21}$$

and performing global fits, including the 'data' on the strange sea provided by CCFR [50], for various different fixed values of the parameter κ. The variation in χ^2 is shown as a function of κ in Fig. 11. The continuous and dashed curves correspond to including all CCFR data or

Variation of χ^2 with strange/non-strange ratio

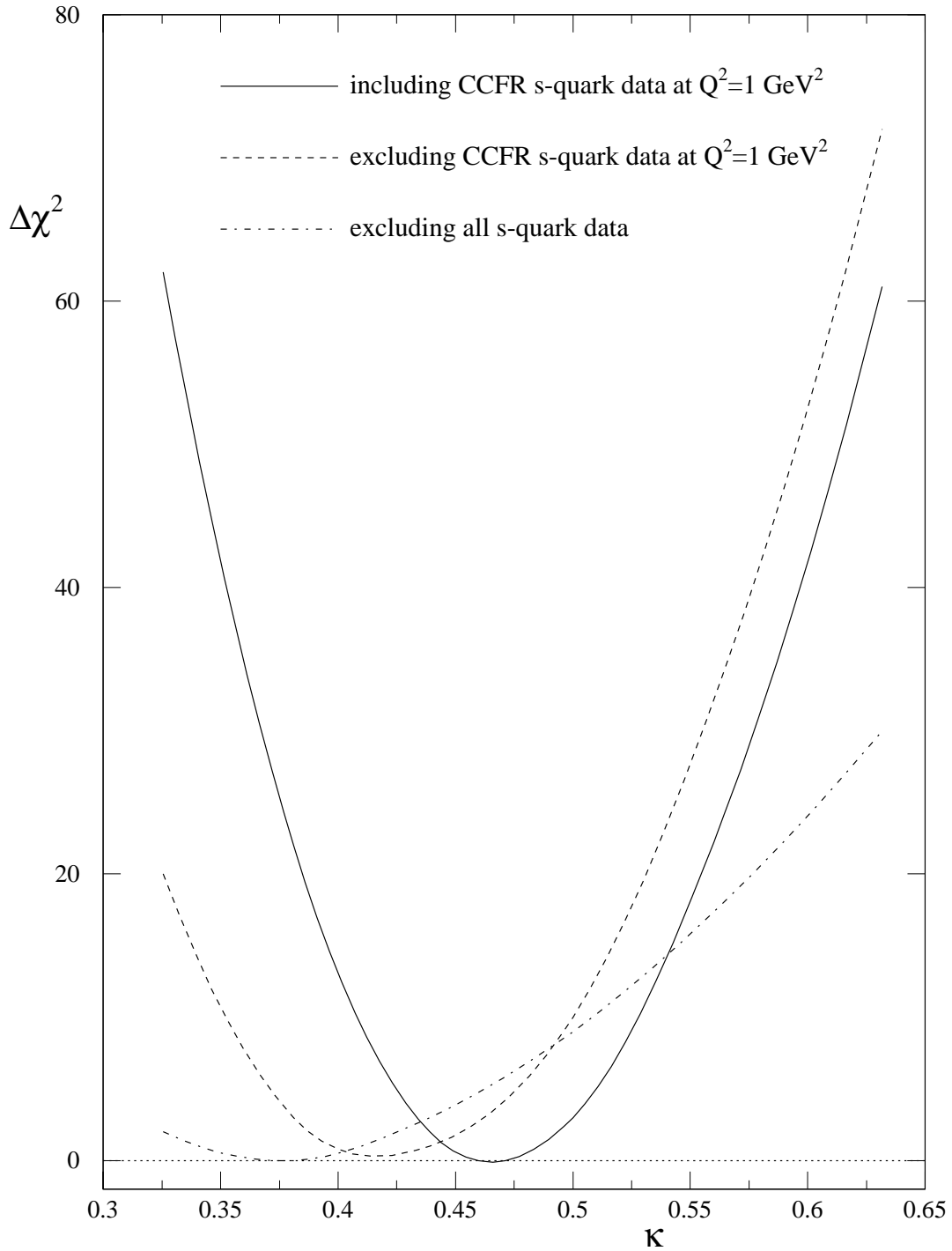Figure 11: The variation of χ^2 for global fits with different values of κ, defined in (21), which gives the strength of the strange quark sea relative to the non-strange sea. The two curves with the deep minima include the CCFR dimuon data [50] in the global χ^2, whereas when these data are omitted the shallower χ^2 profile is obtained.

omitting those CCFR data at the lowest Q^2 value, $Q^2 = 1$ GeV^2. The latter may be more appropriate since NLO leading-twist distributions might become unreliable at such a low scale. We see that our default choice of $\kappa = 0.5$ is near the minimum when the $Q^2 = 1$ GeV^2 CCFR data are included, but is perhaps a little large when these data are omitted. The choice $\kappa = 0.44$ appears optimum for both data selections. We therefore make a set of partons available with this value. In the future, NuTeV dimuon data will better determine the strange sea.

We also investigate the effect of the variation of κ on the quality of our standard global fit (in which the contribution of the dimuon data is not included). This is shown by the (shallower) dash-dotted curve in Fig. 11. We see that the minimum is obtained for $\kappa = 0.38$ and that the χ^2 is about 8 lower than for the MRST(2002) set. Hence this is a rather small improvement in the fit and leads to a very small change in the parton distributions. This value of κ is near the limit of acceptability as determined by the dimuon data. We conclude that the level of uncertainty of parton distributions, and related quantities, is rather small, particularly as it weakened by evolution to higher Q^2.

5.4 Possible isospin violations and $s \neq \bar{s}$

Isospin symmetry implies that the parton distributions of a neutron are obtained from those of the proton simply by swapping the up and down quark distributions, i.e. $d^n(x, Q^2) = u^p(x, Q^2)$ and $u^n(x, Q^2) = d^p(x, Q^2)$. In the absence of any obvious evidence to the contrary this is always assumed to be true in global fits. There are many sets of data in the global fit which would in principle be sensitive to any isospin violation. These are the various sets of deuterium structure function data (SLAC, BCDMS, NMC), the CCFR neutrino structure function data from isoscalar targets, the E605 Drell–Yan data on a copper target, and the NA51 and E866 Drell–Yan asymmetry[13] data. However, in the global fit, all of these data sets are well described, implying that isospin symmetry breaking is small. This may be illustrated as follows. Assuming that the structure functions are dominated by the up and down quarks (and antiquarks), which is largely true, then the measurements of the γ-exchange contributions to F_2 determine the quark contributions

$$F_2^p \propto \frac{4}{9}\left(u^p + \bar{u}^p\right) + \frac{1}{9}\left(d^p + \bar{d}^p\right) \tag{22}$$

$$F_2^n \propto \frac{4}{9}\left(u^n + \bar{u}^n\right) + \frac{1}{9}\left(d^n + \bar{d}^n\right) \tag{23}$$

whilst the structure functions for ν and $\bar{\nu}$ interactions on an isoscalar target constrain the combinations

$$F_2^\nu \propto d^p + \bar{u}^p + d^n + \bar{u}^n \tag{24}$$

$$F_2^{\bar{\nu}} \propto u^p + \bar{d}^p + u^n + \bar{d}^n \tag{25}$$

$$xF_3^\nu \propto d^p - \bar{u}^p + d^n - \bar{u}^n \tag{26}$$

---

[13]The pp and pn asymmetry measurements involve Drell–Yan production on deuterium, as well as hydrogen, targets.

$$xF_3^{\bar{\nu}} \propto u^p - \bar{d}^p + u^n - \bar{d}^n. \tag{27}$$

However the CCFR data [42] for the neutrino structure functions F_2 and xF_3 that are used in the global analyses average over the ν and $\bar{\nu}$ interactions[14]. Note that the F_2^n/F_2^p NMC data determine the ratio of the quark combinations of (23) and (22) with about a 2% error. The data thus allow little flexibility in the neutron parton distributions even without the additional constraints of Drell–Yan data and W-asymmetry data (which constrain the proton valence and sea quarks).

We consider the possibility of isospin violation in the valence quarks and the sea quarks separately. For the sea quarks we assume

$$u_{\text{sea}}^n(x) = d_{\text{sea}}^p(x)(1 + \delta), \tag{28}$$

$$d_{\text{sea}}^n(x) = u_{\text{sea}}^p(x)(1 - \delta). \tag{29}$$

This type of violation is expected from theoretical models, and is consistent with momentum conservation, up to very small violations due to a non-zero value of $(d_{\text{sea}}^p - u_{\text{sea}}^p)\delta$. Strictly speaking it is not preserved by evolution of the partons, but in the kinematic regions of interest the violation is very small. Somewhat surprisingly, we find that a certain amount of isospin violation of this type is actually preferred by the data. The best fit is obtained for $\delta = 0.08$, i.e. an 8% violation of isospin in the sea. This fit has a total χ^2 20 better than the default best fit at NLO. The majority of this improvement comes from the fit to NMC data on F_2^n/F_2^p, which is raised a little by the increase in $u_{\text{sea}}^n(x)$. The fit to the E605 Drell–Yan data is also markedly improved. The change in the up and down sea quarks for this fit compared to our default partons is shown in Fig. 12. This clearly shows the preference of the data for the up sea distribution in the neutron to be enhanced, and the down quark suppressed.

An increase in χ^2 of 50, corresponding to a 90% confidence limit from our arguments in [10], arises when $\delta = 0.18$ or $\delta = -0.08$, see Fig. 13. In the former case the NMC data on F_2^n/F_2^p (which is the most sensitive discriminator of isospin violation) has now become too large, the fit to BCDMS F_2^p data has deteriorated[15] and the description of the E866 Drell–Yan asymmetry has become very poor. The fit to E605 data has continued to improve, however. For $\delta = -0.08$ the prediction for F_2^n/F_2^p has become too small, and the fit to E605 Drell–Yan data has deteriorated.

For the valence quarks a similar model of isospin violation is not possible because it would violate the valence quark number counting for the neutron. Hence we consider a violation of the type

$$u_v^n(x) = d_v^p(x) + \kappa f(x), \tag{30}$$

---

[14]When available, the NuTeV measurements of the ν and $\bar{\nu}$ interactions individually will offer even more stringent checks of the isospin relations.

[15]The deterioration is caused by the decrease of $d_{\text{sea}}^p(x)$, which arises in order to prevent $u_{\text{sea}}^n(x)$ becoming too large.
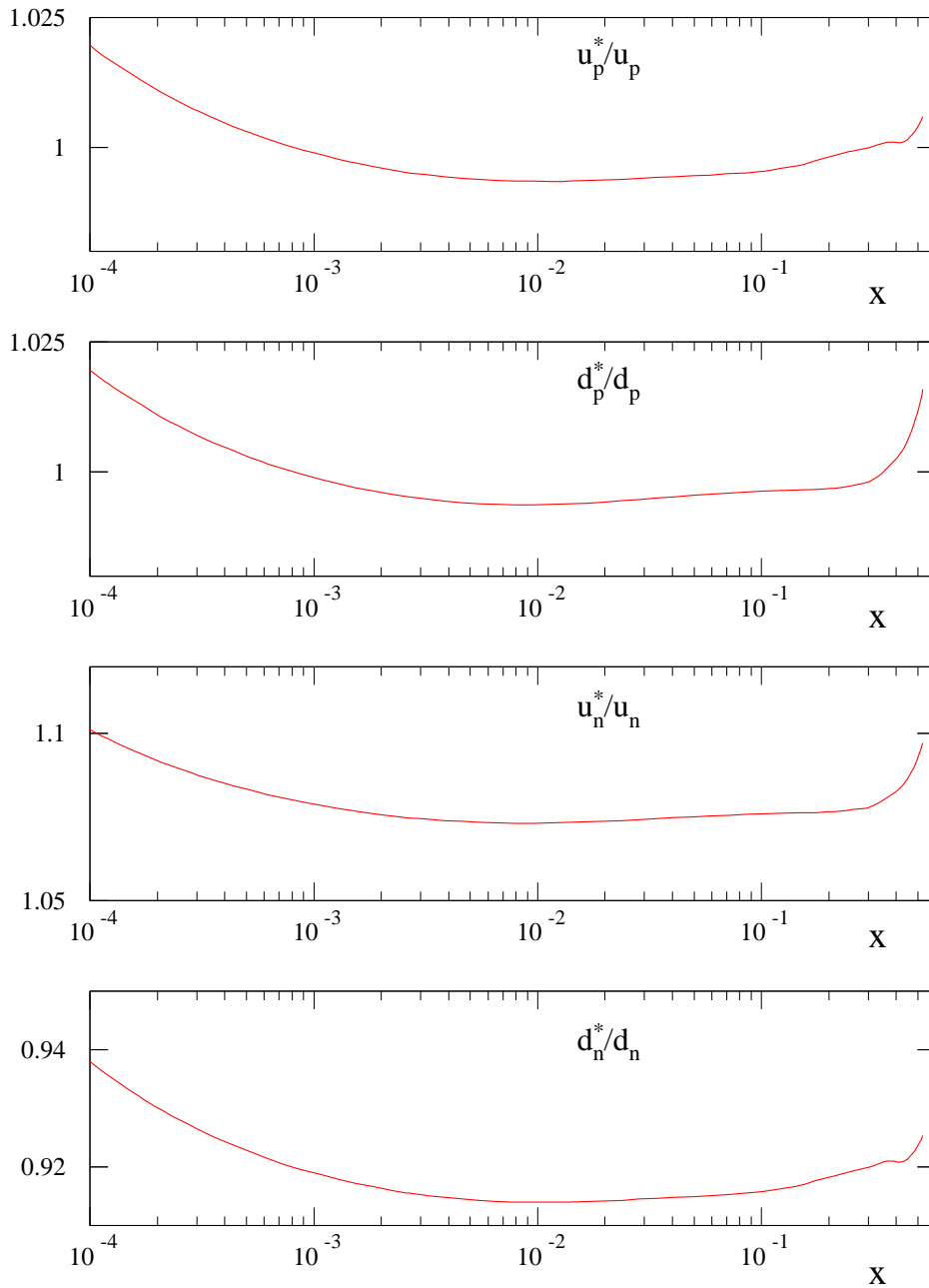
Figure 12: The up and down sea quarks for the best fit with allowed isospin violation in the sea quarks compared to the default partons. The ratio is shown both for the proton and neutron sea quark distributions.

where $f(x)$ is a function which has zero first moment. A suitable function, which has the same type of behaviour as the valence quarks at high and low x is $f(x) = (1-x)^4 x^{-0.5}(x - 0.0909)$. However, there is a further constraint. If we add $\kappa f(x)$ in order to obtain $u_v^n(x)$, then we must subtract $\kappa f(x)$ in order to obtain $d_v^n(x)$, i.e.

$$d_v^n(x) = u_v^p(x) - \kappa f(x), \tag{31}$$

in order to ensure momentum conservation. Hence, we consider isospin violation of this form. Again, the isospin violation is not exactly preserved by evolution, but is correct to a very good approximation.

For valence quarks, there is very little preference for isospin violation. The best fit is obtained for $\kappa = -0.2$, see Fig. 13, but this gives an improvement in χ^2 of only 4. It corresponds to at maximum about a 3% violation of isospin for $u_v^n(x)$. The 90% confidence level is obtained for $\kappa = -0.8$ or $\kappa = 0.65$. In the former case, the main deterioration is in the description of the CCFR $F_2^\nu(x, Q^2)$ data in the region of $x = 0.2$, and in the latter case, it is CCFR $F_3^\nu(x, Q^2)$ data and BCDMS $F_2^d(x, Q^2)$ data which are both badly fit at $x \sim 0.5$. For positive κ, $d_v^p(x)$ decreases at high x to compensate the isospin violating term and in order to fit the $F_2^p(x, Q^2)$ data $u_v^p(x)$ decreases, but less severely (due to the higher charge weighting). Hence, the failure in the fit occurs when $u_v^n(x)$ has increased too much for the BCDMS deuterium data but the larger decrease in $d_v^p(x)$, compared to the increase in $u_v^p(x)$, leaves the prediction for $F_3^\nu(x, Q^2)$ (which is sensitive to $u_v^p(x) + d_v^p(x)$) too small. For negative κ, $d_v^p(x)$ increases at high x, to compensate the isospin violating term, and in order to fit the $F_2^p(x, Q^2)$ data, $u_v^p(x)$ increases less severely. In this case $u_v^p(x) + d_v^p(x)$ decreases, and is too small for $F_2^\nu(x, Q^2)$. These upper and lower limits represent an isospin violation of at most $\mathcal{O}(10\%)$ for $u_v^n(x)$. We offer no theoretical model for why our isospin violation is of the form seen, and note that calculations of the effect, e.g. [52], often indicate smaller results. However, these calculations are very difficult, depending on intrinsically nonperturbative physics, and relying on models and assumptions. We simply examine the empirical evidence given by the data.

It is interesting to compare the level of isospin violation with that needed to explain the NuTeV $\sin^2 \theta_W$ anomaly [53]. The quantity measured[16] by NuTeV is

$$R^- = \frac{\sigma_{NC}^\nu - \sigma_{NC}^{\bar\nu}}{\sigma_{CC}^\nu - \sigma_{CC}^{\bar\nu}}. \tag{32}$$

In the simplest approximation, i.e. assuming an isoscalar target, no isospin violation and equal strange and anti-strange distributions, this ratio is given by

$$R^- \approx \frac{1}{2} - \sin^2 \theta_W, \tag{33}$$

and so the measurement gives a determination of $\sin^2 \theta_W$. NuTeV find $\sin^2 \theta_W = 0.2277 \pm 0.0013(\text{stat.}) \pm 0.0009(\text{syst.})$ [53], compared to the global average of 0.2227 ± 0.0004, i.e. about

---

[16] The NuTeV experiment does not exactly measure R^-, in part because it is not possible experimentally to measure neutral current reactions down to zero recoil energy, see [53, 54].

a 3σ discrepancy. However, if one allows for isospin violation then the simple expression becomes modified to

$$R^- = \frac{1}{2} - \sin^2\theta_W + (1 - \frac{7}{3}\sin^2\theta_W)\frac{[\delta U_v] - [\delta D_v]}{2[V^-]}, \tag{34}$$

where

$$[\delta U_v] = \int_0^1 x(u_v^p(x) - d_v^n(x)), \qquad [\delta D_v] = \int_0^1 x(d_v^p(x) - u_v^n(x)), \tag{35}$$

are measures of the inequality in momentum fraction of the valence quarks induced by isospin violation, and $[V^-] \approx 0.45$ is the overall momentum fraction carried by the valence quarks. One can easily see that given a fixed value of measured R^-, a negative value of κ moves the extracted value of $\sin^2\theta_W$ downwards. The approximate effect can be found simply by using Eq. (34), but a more precise result is obtained by applying the functionals presented in [54], which account for the complications of the measurement. This reduces the naive modification by $\sim 10\%$. Using our best fit value of $\kappa = -0.2$ we obtain $[\delta U_v] = -[\delta D_v] = 0.002$ (corresponding to $< 0.5\%$ of the momentum carried by the valence quarks) and the modification is $\Delta\sin^2\theta_W = -0.0018$. Hence, about $1\sigma - 1.5\sigma$ of the 3σ discrepancy is removed. The determination of $\sin^2\theta_W$ is far less sensitive to the isospin violation in the sea quarks. Our preferred violation with $\delta = 0.08$ is in the wrong direction to account for the discrepancy, but only reduces the effect from the valence quarks slightly to $\Delta\sin^2\theta_W = -0.0015$. Hence, the total result from the best fits with allowed isospin violation is to reduce the 3σ discrepancy to a 2σ discrepancy. However, the allowed range of isospin violation could easily allow the discrepancy to be removed altogether, or even to be made worse. It is nevertheless interesting that the weak indication given by the data in the global fit is such as to reduce the discrepancy by a significant amount.

One can also investigate the possibility of $s(x) \neq \bar{s}(x)$. The only data in the fit which are sensitive to this difference are the NuTeV dimuon data [51]. These data are more difficult to analyse than the similar CCFR data because they are presented in a much more exclusive form. However, in principle such data will allow for a much more detailed analysis. The NuTeV group themselves have performed an analysis of their data allowing the $s(x)$ and $\bar{s}(x)$ distributions to have different normalisations and different $(1-x)^\eta$ behaviour at high x. Their analysis indicates that the data would prefer a slight (11%) excess of $\bar{s}(x)$ over $s(x)$ [51]. However the analysis is at leading order, and does not impose the quark counting rule, i.e. equal number of strange and anti-strange quarks. It is complicated to improve this analysis to the full NLO level. CTEQ have performed similar preliminary analyses [55], obtaining very similar results. Most predictions of physical quantities are insensitive to the potential imbalance of $s(x)$ and $\bar{s}(x)$, but the NuTeV $\sin^2\theta_W$ anomaly is affected by any difference [53]. However, the excess of $\bar{s}(x)$ over $s(x)$ actually makes the discrepancy slightly worse. More precise data and a full theoretical treatment will hopefully lead to an improved understanding of this question in future.[17]

---

[17]CTEQ have produced fits where the quark counting is imposed and find that the excess in $\bar{s}(x)$ over $s(x)$ in the region of data must then be countered by an excess of $s(x)$ over $\bar{s}(x)$ at higher x, leading to a positive momentum excess of $s(x)$ over $\bar{s}(x)$ [56]. This is then in the correct direction to reduce the NuTeV $\sin^2\theta_W$ anomaly.
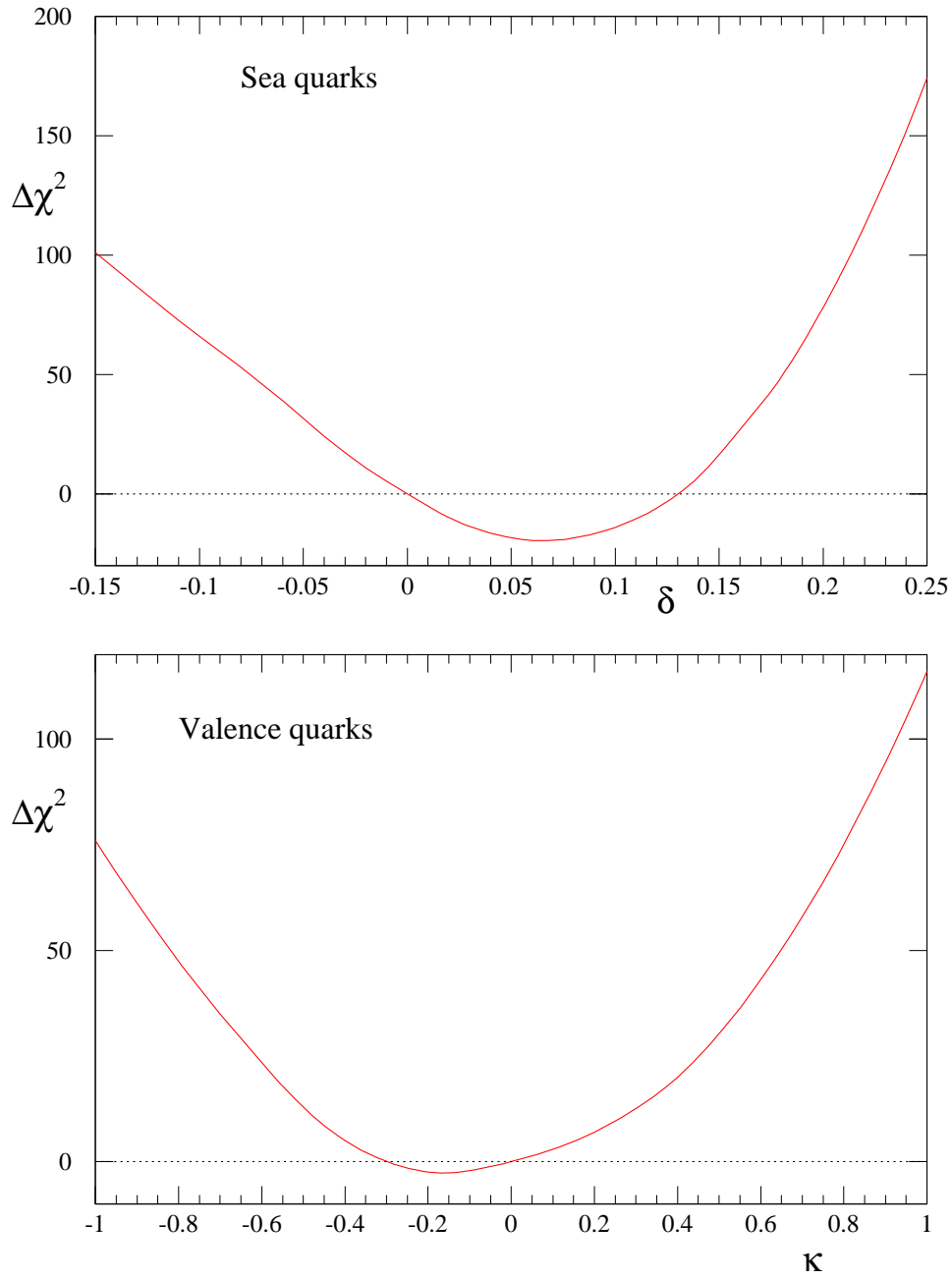
Figure 13: χ^2 profiles showing the effect of isospin violation in the sea and valence quark sectors respectively. δ and κ are defined in eqs. (28) and (30) respectively.

	$\Delta\chi^2$	$\alpha_S(M_Z^2)$ $\pm$ expt $\pm$ theory $\pm$ model
NLO		
CTEQ6	100	0.1165 ± 0.0065
ZEUS	50	0.1166 ± 0.0049 ± 0.0018
MRST03	5	0.1165 ± 0.002 ± 0.003
H1	1	0.115 $\pm 0.0017 \pm 0.005^{+0.0009}_{-0.0005}$
Alekhin	1	$0.1171 \pm 0.0015 \pm 0.0033$
NNLO		
MRST03	5	$0.1153 \pm 0.002 \pm 0.003$
Alekhin	1	$0.1143 \pm 0.0014 \pm 0.0009$

Table 4: The values of $\alpha_S(M_Z^2)$ found in NLO and NNLO fits to DIS data. The experimental errors quoted correspond to an increase $\Delta\chi^2$ from the best fit value of χ^2. CTEQ6 [6] and MRST03 are global fits, where the latter correspond to the 'conservative' sets of partons of Sections 2.5 and 3.2. H1 [7] fit only a subset of F_2^{ep} data, while Alekhin [57] also includes F_2^{ed} and ZEUS [9] in addition include xF_3^ν data.

6   Implications for α_S and predictions for observables

6.1   Determination of α_S

As we have already demonstrated, there is a significant amount of variation in our extracted value of $\alpha_S(M_Z^2)$ when we vary the input assumptions for our fitting procedure, particularly when we vary the x and Q^2 cuts. Previously we have always determined $\alpha_S(M_Z^2)$ from a global fit using our default cuts, that is no cut on x and a Q^2 cut of 2 GeV^2. As demonstrated earlier, the evidence suggests that when fitting this full range of data, standard NLO (or NNLO) perturbation theory is not completely sufficient, and the variation in $\alpha_S(M_Z^2)$ is due to the parameters in the fit compensating for the deficiencies of the theoretical treatment. Hence, the 'true' value of $\alpha_S(M_Z^2)$ should be that corresponding to the 'conservative' partons at both NLO and NNLO. We present these values below in Table 4, labelled MRST03, together with other recent determinations. Note that the conservative NLO partons themselves use $\alpha_S = 0.1162$. However the χ^2 profile versus α_S is very flat between 0.116 and 0.117. We therefore choose the mid-point as the best value of α_S, and $\Delta\chi^2 = 5$ to give the 1σ error of ± 0.002.

As mentioned in Section 2, a Bayesian approach to determining parton uncertainties [2] gives $\alpha_S(M_Z^2) = 0.112 \pm 0.001$. However, in order to satisfy the strict requirements of consistency between the data sets, this Bayesian analysis only uses the BCDMS [13], E665 [14] and H1(94) [12] data for F_2^p. Bearing in mind that the H1(94) data have relatively large errors,

the value of $\alpha_S(M_Z^2)$ that is obtained simply reflects the original BCDMS determination of Ref. [62].

From Table 4 we see that the various determinations of $\alpha_S(M_Z^2)$ have approximately converged to a common value, even though they are based on different selections of the DIS and related data. Averaging the two global NLO analyses we have

$$\alpha_S(M_Z^2) = 0.1165 \pm 0.004. \tag{36}$$

Previously, the MRST value [15] was larger, $\alpha_S(M_Z^2) = 0.119$. This was due to the attempt to fit data in regions where the theoretical corrections to NLO DGLAP appear to be important. It is well illustrated by Fig. 16 of [15], where we see that the optimum value of $\alpha_S(M_Z^2)$ varies considerably from data set to data set. However, when the 'conservative' data cuts are applied, the tension between the data sets is reduced enormously, as can be seen by comparing Fig. 14 with Fig. 16 of [15]. Some data sets, particularly the SLAC data, which prefer a high value of $\alpha_S(M_Z^2)$, and the E605 Drell-Yan data and the BCDMS data, which prefer a low value of $\alpha_S(M_Z^2)$, still pull strongly away from the minimum value. However, the other data sets are now at, or near, their minimum χ^2 for the best fit value of $\alpha_S(M_Z^2)$, which was certainly not the case for the default fit. The D0 and CDF jet data are a particularly good example, where not only are the data better fit by the 'conservative' sets than the default sets, but they are no longer pulled to extreme values of $\alpha_S(M_Z^2)$ in order to obtain their best individual fits. The same marked improvement occurs for the NMC data which survive the x and Q^2 cuts. This increased compatibility between the data sets, and also between the data and the theory, is why the tolerance $\Delta\chi^2$ has been reduced from 20 (in [15]) to 5 in the present study. The reduction in tension between data sets also occurs when $\ln(1/x)$, $\ln(1-x)$ or higher-twist corrections are included, as discussed in Section 4.

There are fewer extractions of the value of $\alpha_S(M_Z^2)$ using NNLO global or semi-global fits to DIS and related data. The results are also shown in Table 4. Again we see good agreement, and a small, but definite, reduction from the NLO value. We found, in this case, far less sensitivity to the data cuts, indicating that some important theoretical corrections are already accounted for at NNLO.

6.2 Predictions for W and Higgs hadroproduction

Predictions for physical quantities are, like the value of $\alpha_S(M_Z^2)$, sensitive to the 'theoretical' uncertainties in the global parton analysis. For illustration, we show in Figs. 15 – 18 the variation of the W and Higgs cross section predictions for the Tevatron and LHC as a function of x_{cut} and Q^2_{cut}.

The cross sections for W and H production at the Tevatron sample partons down to $x \simeq 0.005$, and so are only directly sensitive to partons within the range of our most conservative cuts. However the cross sections can still vary if we change the values of the cuts due to the
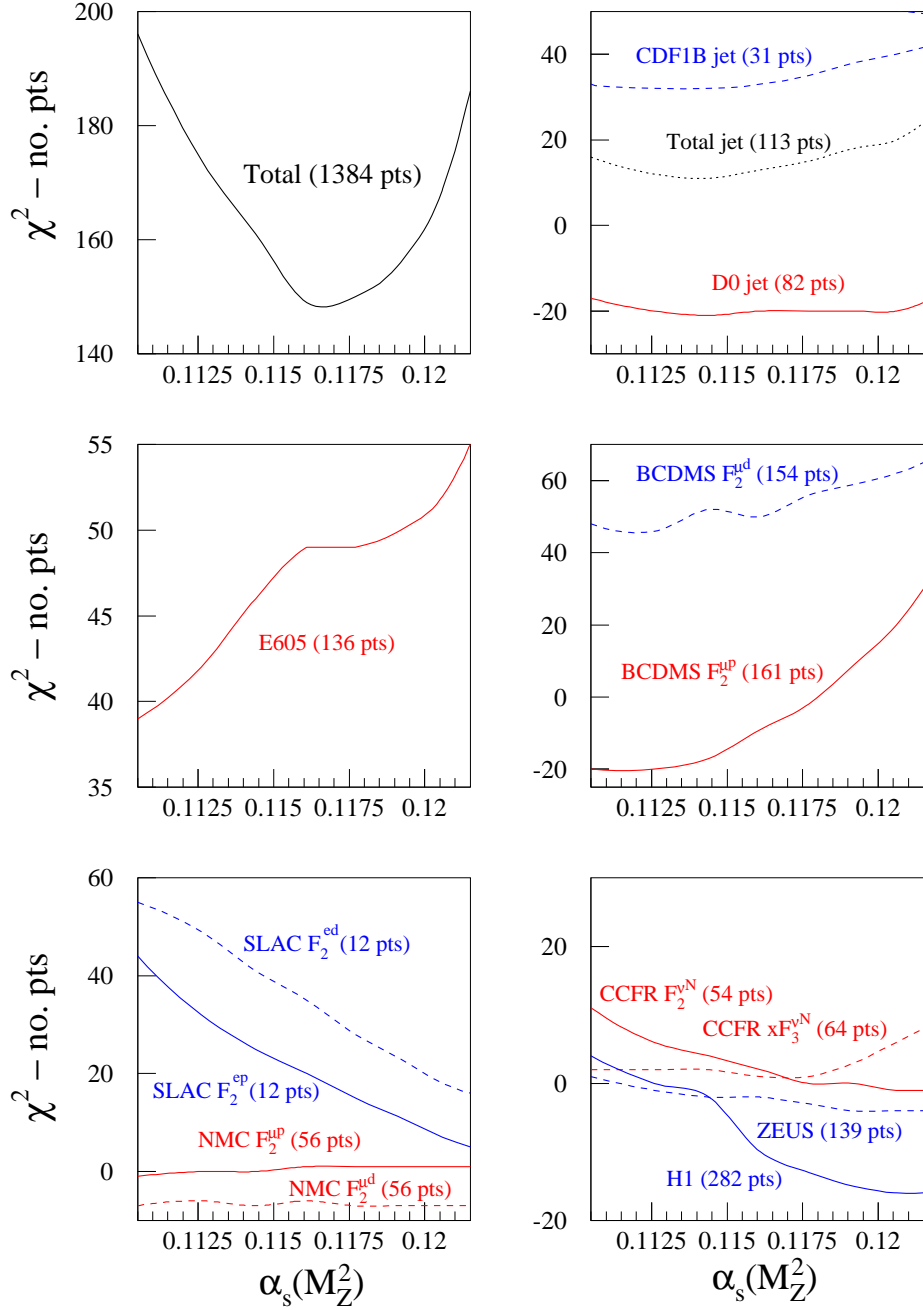
Figure 14: The quality of the fit to the individual data sets included in the NLO global analysis with $x_{\text{cut}} = 0.005$ and $Q^2_{\text{cut}} = 10$ GeV^2, shown together with the overall total χ^2, as a function of $\alpha_S(M_Z^2)$.
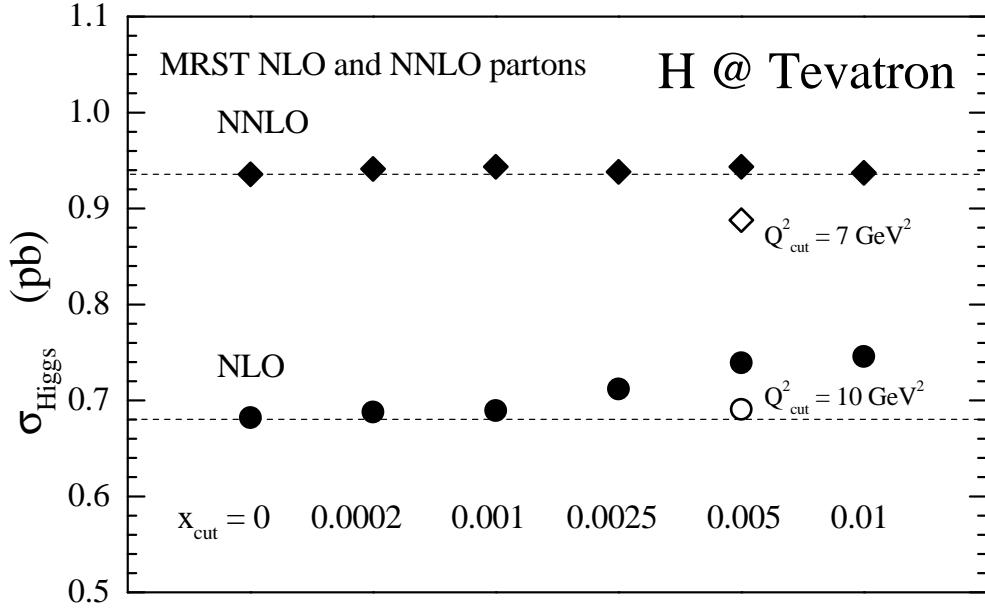
Figure 15: Predictions for the ($M_H = 120$ GeV) Higgs cross section at the Tevatron ($\sqrt{s} = 1.96$ TeV) at NLO and NNLO for various values of x_{cut}, and for the 'conservative' partons with a cut on both x and Q^2 (shown as open symbols).
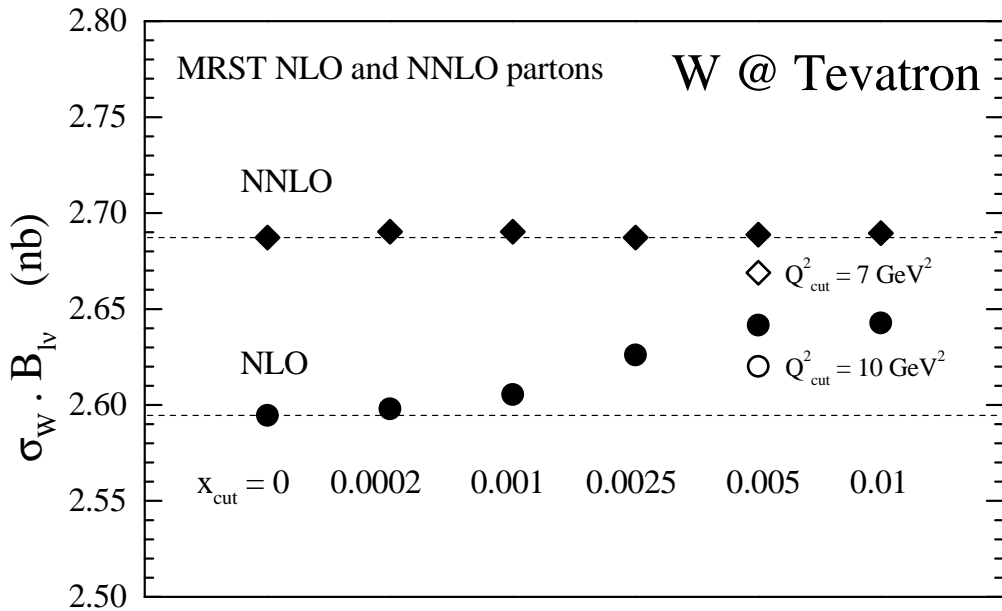

Figure 16: Predictions for the W cross section (times the leptonic branching ratio $B_{l\nu} = 0.1068$) at the Tevatron ($\sqrt{s} = 1.96$ TeV) at NLO and NNLO for various values of x_{cut}, and for the 'conservative' partons with a cut on both x and Q^2 (shown as open symbols).
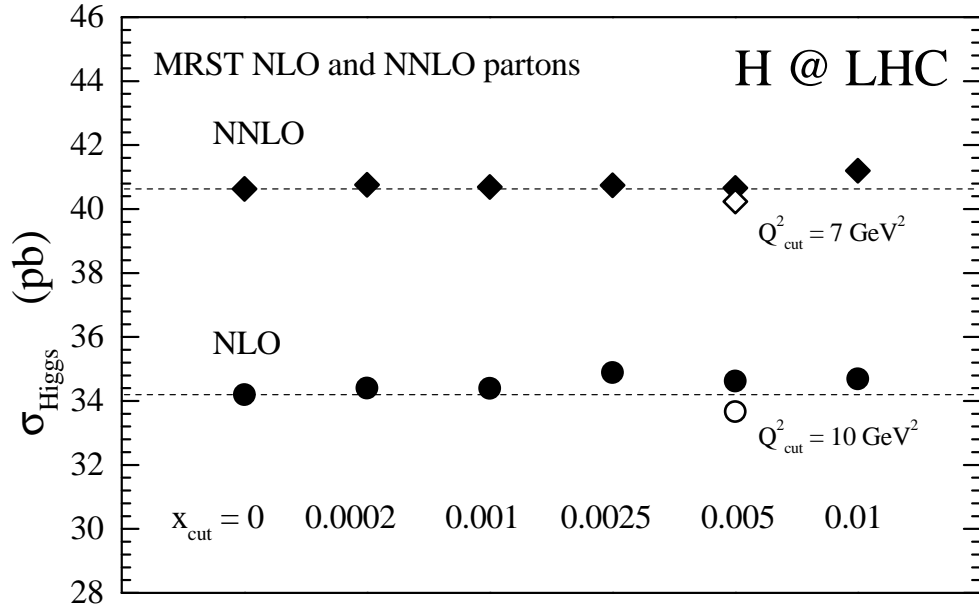
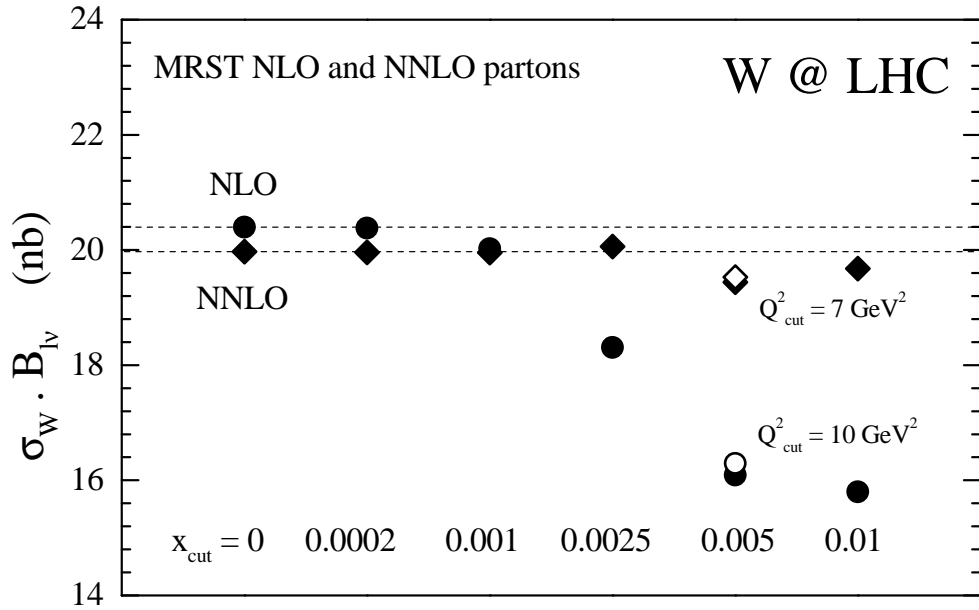Figure 17: The same as Fig. 15, but for the LHC energy of $\sqrt{s} = 14$ TeV.

Figure 18: The same as Fig. 16, but for the LHC energy of $\sqrt{s} = 14$ TeV.

43

readjustment of partons above x_{cut} and Q^2_{cut}, and for Higgs production, due to the different values of $\alpha_S(M^2_Z)$ extracted. This is evident in Figs. 15 and 16.

The NLO prediction[18] for the cross section for the production of a Higgs of mass 120 GeV at the Tevatron rises steadily as x_{cut} is increased, saturating at an increase of about 9% for $x_{\text{cut}} \simeq 0.005$. This rise is due to the increase of the gluon distribution at moderate x values as the value of x_{cut} is raised, see Fig. 1. There is a slight decrease in the value of $\alpha_S(M^2_Z)$ with increasing x_{cut}, but the effect of this on the Higgs cross section is completely outweighed by the large increase in the gluon in the relevant x range. We also see from Fig. 17 that, when Q^2_{cut} is increased to 10 GeV^2, corresponding to our conservative set of NLO partons, the Higgs cross section is only increased by 1%. This is mainly due to the drop in $\alpha_S(M^2_Z)$ to 0.116, but also due to a slight decrease in the gluon for the relevant value of $x \sim 0.06$ (for central rapidity production), compared to the case where only the cut in x is applied. This arises because the Q^2 cut of 10 GeV^2 eliminates the NMC $F_2^{p(d)}(x, Q^2)$ data for $x \sim 0.05$, which prefer a steeper rise of $F_2(x, Q^2)$ with Q^2, and hence a larger gluon in this range (as well as a larger value of $\alpha_S(M^2_Z)$). Nevertheless the value for the 'conservative' set represents a non-negligible increase in the Higgs cross section compared to the prediction of the default set.

From Fig. 16 we also see that the NLO prediction for the cross section for W production at the Tevatron also rises steadily as x_{cut} is increased, saturating at an increase of about 2% for $x_{\text{cut}} \simeq 0.005$. This is due to the increased evolution of the quarks driven by the increase of the gluon in the relevant range. Again, when we also impose $Q^2_{\text{cut}} = 10$ GeV^2 the predicted value of the cross section decreases. The increased value of Q^2_{cut} allows the input quarks to be larger for $x \simeq 0.05$ (too large for the NMC data now cut out), and the improvement in the quality of the fit requires less increase in $\partial F_2/\partial \ln Q^2$. Compared to the quarks at $x \simeq 0.05$ obtained using $Q^2_{\text{cut}} = 2$ GeV^2, the quarks corresponding to $Q^2_{\text{cut}} = 10$ GeV^2 cross in magnitude in the region of $Q^2 = 250$ GeV^2, and become 0.4% smaller at $Q^2 = M^2_W$.

At NNLO we see from Figs. 15 and 16 that the predictions for W and H production at the Tevatron are much more stable to variations of x_{cut}. The reason is that the increased evolution of quarks for $x \simeq 0.05$, due to the NNLO contributions to the splitting and coefficient functions, requires much less change in the gluon, and consequently in $\partial F_2/\partial \ln Q^2$ in this range. The additional imposition of $Q^2_{\text{cut}} = 7$ GeV^2, however, reduces the Higgs cross section significantly, by about 5%, even though at NNLO it leads to only a slight decrease in the value of $\alpha_S(M^2_Z)$. This is because the loss of the NMC $F_2^{p(d)}(x, Q^2)$ data for $x \sim 0.05$ below Q^2 of 10 GeV^2 has resulted in a large reduction in the gluon for $x \sim 0.1$, as seen in Fig. 4. The prediction for the W cross section also falls, by about 1%, for the same reason as at NLO. Hence at NNLO the 'conservative' parton set predicts a small decrease in both σ_H and σ_W compared to the prediction of the default parton set, while at NLO there is a small increase in both.

---

[18]For the Higgs cross section calculations described in this section we use the full NLO QCD correction in the $m_t \gg M_H$ limit [58] and the soft-virtual-collinear x–space (SVC$_x$) approximation to the full NNLO correction, again for $m_t \gg M_H$, taken from Ref. [59]. In both cases the factorisation and renormalisation scales are set equal to M_H.

As a consequence, the 'conservative' parton predictions show slightly greater convergence with increased perturbative order, than the default predictions.

At the LHC, the cross sections for W and H production at central rapidity sample partons at $x = 0.006$ and $x = 0.0085$ respectively, so these are safely predicted using the 'conservative' partons. However, the *total* cross sections sample a wide range of rapidity, and in fact probe down to $x \sim 0.00008$ for W production and $x \sim 0.0006$ for H production, and hence the predictions using the 'conservative' partons are not guaranteed to be reliable. We present such predictions to exhibit where the reliability does break down, and to what extent.

We notice (Fig. 17) that the predictions for the Higgs cross section at the LHC are actually rather stable to changes in the value of x_{cut} at both NLO and NNLO. From Fig. 3 this is not too surprising. At the high scales that are needed for Higgs production, the NNLO gluon is hardly changed compared to the default. The NLO gluon, with $x_{\mathrm{cut}} = 0.005$ applied, is a little larger in the central rapidity region, and only falls to a much smaller value than the default at values of x that are only making very small contributions to the total cross section. Hence, there is only a slight increase in the prediction for the Higgs cross section compared to the default. When the additional cut in Q^2 is applied, in order to obtain the conservative sets, both the NLO and NNLO predictions decrease slightly due to the decreases in $\alpha_S(M_Z^2)$. In fact, at NLO and NNLO the predictions using the 'conservative' sets finish very close to the default predictions (-1.5% and -1% for the changes at NLO and NNLO respectively). This implies that there is little theoretical error associated with Higgs production at the LHC due to the partons themselves. Indeed the variation with x_{cut} and Q^2_{cut} at the LHC, exhibited in the figure, is evidently much smaller than the correction in going from NLO to NNLO, that is, smaller than from the NNLO coefficient function contribution.[19]

For W production at the LHC, Fig. 18, the story is rather different. At NLO there is a steady, and rather dramatic decrease in the predicted cross section as x_{cut} is increased. This culminates in a drop of 20% for $x_{\mathrm{cut}} = 0.005$ (which is insensitive to the additional Q^2 cut). The reason for this is clear from examination of Fig. 2. For the sets with $x_{\mathrm{cut}} = 0.005$ the quark distributions for $x < 0.005$ are much smaller than those of the default set. Also the gluon at small x and low Q^2 is reduced, and therefore the evolution of the quarks is slower than the default set. Hence, although at $Q^2 \sim 10^4$ GeV^2 the quarks at the central rapidity value of $x = 0.006$ are actually much the same as in the default set, they quickly become reduced at smaller x, and very much so for $x < 0.001$. However, much of the total cross section comes from W rapidities corresponding to such low x for one of the two quarks contributing to the reaction, and, as a consequence, the contribution to the cross section at high rapidities is much reduced. This is illustrated in Fig. 19, which shows the differential cross section as a function of the W rapidity. At central rapidity, the cross section is even increased slightly, but it falls away very quickly with increasing $|y_W|$, resulting in the 20% loss for the total. However, at NNLO there is far greater stability. There is a slight decrease in predicted cross section with

---

[19]By way of calibration, the scale dependence of the $M_H = 120$ GeV Higgs cross section at the LHC is quoted as $\pm 10\%$ (at NNLO) and $\pm 8\%$ (using NNLL soft-gluon resummation) in the recent study of Ref. [60].
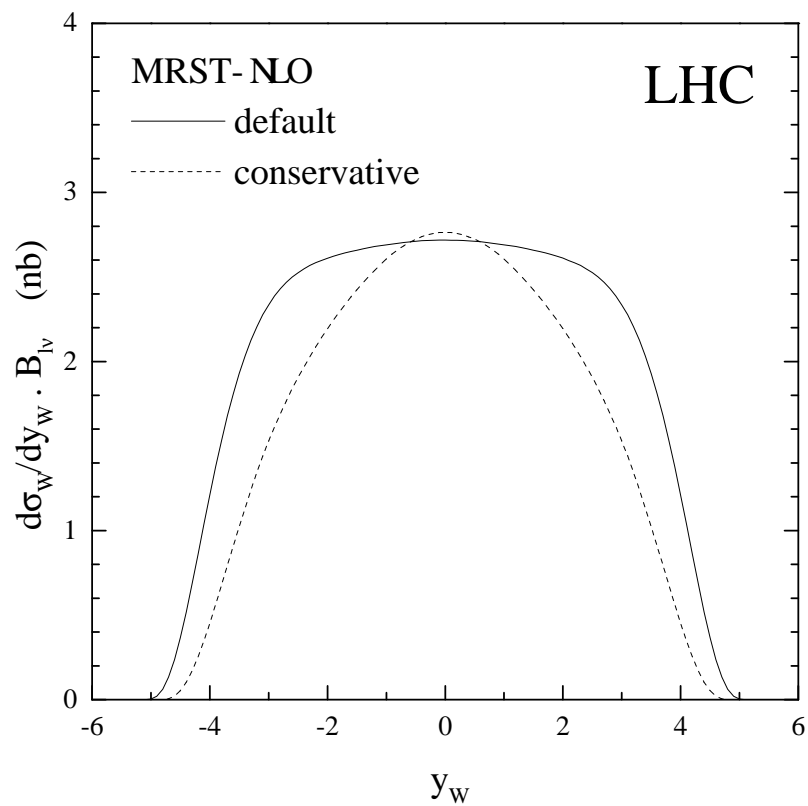
Figure 19: The predictions for the rapidity distribution of the W cross section at the LHC for both the default MRST2002 set and for the 'conservative' set.

increasing $x_{\rm cut}$, which can be understood from the small decrease in the quarks for x in the region 0.0001–0.005, as compared to the default set, as seen in Fig. 4. However, as we have already commented for the gluon, the change in the very low x partons with increasing $x_{\rm cut}$ is very much reduced at NNLO, because the NNLO contributions themselves include important corrections to the small-x behaviour of the partons and structure functions.

We believe that the variation of the predicted cross sections with the value of $x_{\rm cut}$ and $Q^2_{\rm cut}$ gives a rough indication of the theoretical uncertainty due to the parton distributions. From the NNLO prediction for the cross section for W production at the Tevatron, we see from Fig. 16 that, not only is the value stable, but that the *theoretical* uncertainty is of order 1–2%. This is similar in magnitude to the uncertainty due to experimental errors, which was obtained in Ref. [10]. This gives a *total* error of about $\pm 2\%$, which would imply that observing the W production rate at the Tevatron (and comparing with the NNLO prediction) could serve as a valuable luminosity monitor. The theoretical error, from the partons, for NNLO Higgs production at the Tevatron is rather larger (about 5%), due to the adjustment of the gluon over the whole x range and changes in α_S, when the input assumptions to the fit are varied.

The variation at the LHC is relatively small for the Higgs cross section, which does not probe partons at too low x, and is swamped by the higher-order corrections to the partonic cross section. However, it is potentially much larger for the W cross section at the LHC, which relies on lower x partons. The 20% variation in the W cross section at the LHC at NLO implies that important theoretical corrections are required. The reduction of this variation to $\sim 3\%$ at NNLO implies than much of this correction has occurred in going from NLO to NNLO. However, in this case, the uncertainty is still larger than the change due to the NNLO contribution to the partonic cross section and is the dominant theoretical uncertainty.[20] The results in Section 4 imply that even at NNLO additional theoretical precision could be obtained by further theoretical corrections, e.g. a correct inclusion of $\ln(1/x)$ terms at higher orders. However, NNLO seems to be a great improvement on NLO, if one wishes to predict quantities sensitive to partons for x much lower than 0.005.

7   Conclusions

In a previous paper [10] we have already studied the uncertainties of parton distributions, and related observables, arising from the errors on the experimental data used in the global parton analysis. However in that paper we had already commented that in many cases the major uncertainty could be due to corrections to the standard DGLAP evolution and due to

---

[20]A very recent prediction of the W and Z cross sections at NLO and NNLO, with errors, has appeared in [61]. This disagrees with our predictions by $\mathcal{O}(5\%)$, which is larger than the total quoted error in [61] (2% for the Tevatron and 3% for the LHC). The main reason for this discrepancy is almost certainly due to the absence of a number of sets of data which determine the precise form of the quark distributions in the partonic fit [57] that is used. Hence, in our spirit of determining parton distributions, we would deem this to be a removable discrepancy.

assumptions used in the fit procedure, in other words due, collectively, to so-called *theoretical* errors. In this paper we have studied a wide range of sources of theoretical error and their potential consequences.

To begin, we investigated the possible corrections to standard fixed-order DGLAP analyses. The investigation was performed in two alternative ways. First we made an empirical study in order to find those kinematic regions where DGLAP evolution was fully consistent. This was done, at both NLO and NNLO, by gradually eliminating data until the analyses were stable to further cuts. We found that this was best achieved by imposing separate cuts on x, Q^2 and W^2 of the data fitted.

- For W^2, stability was achieved by raising the cut from our default value of 12.5 GeV^2 to just 15 GeV^2. However, in the region of low W^2, we noted that there exists some incompatibility between the data sets.

- For x, stability was achieved only for the relatively high value of $x_{\text{cut}} = 0.005$. When this cut was applied, the gluon distribution above this x value increased, relative to the default set, at the expense of the loss of the gluon at smaller x. This improved the quality of the fit to the Tevatron jet data, and also to the structure function data for $x \sim 0.01$ due to an increase in $\partial F_2/\partial \ln Q^2$. The x_{cut} required was the same at NNLO as at NLO, but the modification of the gluon was considerably smaller at NNLO.

- For Q^2, stability is reached for $Q^2_{\text{cut}} \sim 7$–10 GeV^2 at both NLO and NNLO, the convergence to stability being quite gradual as Q^2_{cut} is raised. The slow convergence indicates that higher-order corrections, rather than higher-twist effects, are important at relatively low Q^2.

At both NLO and NNLO we also considered combinations of the above types of cuts. Thus we found the full kinematic region where fixed-order DGLAP analysis is appropriate, together with the corresponding sets of conservative partons. At NLO the domain is given by $W^2 > 15$ GeV^2, $Q^2 > 10$ GeV^2 and $x > 0.005$, whereas at NNLO it is given by $W^2 > 15$ GeV^2, $Q^2 > 7$ GeV^2 and $x > 0.005$.[21] Within these regions these conservative sets of partons are most reliable, but should not be used outside the domain. Indeed, outside the region they may be completely incompatible with the data. However, we note that, although the NLO and NNLO regions of stability are similar, the NNLO partons are far closer to the default partons outside the stable domain, indicating smaller theoretical uncertainty at NNLO.

Complementary to the above empirical study, we also made an explicit investigation of the following variety of possible theoretical corrections.

---

[21]In practice the conservative partons were obtained using $W^2_{\text{cut}} = 12.5$ GeV^2, but raising W^2_{cut} to 15 GeV^2 has a negligible effect on the partons.

- Higher-order parametric $\ln(1/x)$ corrections to the splitting functions were found to improve the quality of the fit. This occurs both at small x, as expected, and also at larger x where the partons are allowed to re-adjust to a form similar to that of the conservative parton set (with $x_{\text{cut}} > 0.005$). The required corrections were not the same at NNLO as at NLO, being overall a little smaller at NNLO, reflecting the important small x contributions introduced at NNLO.

- Shadowing (or absorptive) corrections were introduced, but found to have little effect when the default partons were used as the starting point. If the input gluon was forced to be even very slowly increasing at small x, then evolution is too rapid even with shadowing corrections included. Further study would require the simultaneous description of diffractive structure function data.

- Parametric higher-twist contributions were introduced and quantified. These decreased as we progressed to higher perturbative order, and, at NNLO, were only evident for $x \gtrsim 0.5$. However, at both NLO and NNLO, the change in the input partons is similar to that invoked by increasing Q^2_{cut}. The high x, or equivalently low W^2, higher-twist contribution is intrinsically related to the $\ln(1-x)$ resummation. This is because the latter is inherently divergent and the ambiguity in the perturbation series must cancel that in the higher-twist contribution. We found that the effect of the well-determined NNNLO high x contribution was important, and reduced the high x, higher-twist contribution still further. However this NNNLO contribution demarks the perturbative order beyond which the series fails to converge. Hence we conclude that resummation beyond NNNLO becomes indistinguishable from higher-twist corrections at high x.

As well as the theoretical corrections to the standard DGLAP evolution there are other potential sources of uncertainty in the fitting procedure. Among these we considered the following.

- We found our input parameterization was sufficiently flexible to accommodate the data, and indeed there is a certain redundancy evident. Hence we conclude that the form of our parameterization does not significantly constrain the description, though a more efficient parameterization may be possible. Allowing our input gluon to be negative at small x does have important consequences however.

- Heavy-target corrections are required when fitting to neutrino data. We found that, although our default choice was not quite optimum, changes resulted in different data sets becoming more compatible with each other, and led to only minimal changes in the partons.

- Shadowing corrections are also required when fitting to deuterium structure functions. We found that there was a small amount of evidence for some high x shadowing, and that

the uncertainty in valence partons, particularly $d_v(x, Q^2)$ at high x, due to the model uncertainty in deuterium shadowing is similar to the uncertainty due to the experimental errors on the data.

- The input strange quark sea distribution is parameterized as $\kappa(\bar{u}+\bar{d})/2$, where the default choice was $\kappa = 0.5$. We found that both the global fit and the description of the CCFR dimuon data prefer a smaller value of $\kappa \simeq 0.44$. This choice, which gives a smaller strange sea distribution and slightly larger $\bar{u}, \bar{d}$, will be implemented in future fits. However it leads to only very small changes in most physical quantities.

- The possibility of isospin violations was investigated in both the sea and valence quark sectors. In the former case, we found an that increase of u_{sea}^n (and a corresponding decrease of d_{sea}^n), in comparison to the default set, was preferred by the data at the 8% level. In fact the acceptable range of u_{sea}^n allowed an increase up to 18% and a decrease down to 8%. For the valence quarks there was no significant improvement in the fit due to possible isospin violations, though a slight preference for $u_v^p(x) > d_v^n(x)$ and $d_v^p(x) < u_v^n(x)$ at high x was noted. Changes in u_v^n of up to $\pm 10\%$ were permitted. From conservation of quark number the percentage violation for d_v^n is half that for u_v^n. For both valence and sea quarks the consequent change in the proton distributions in the isospin-violating sets of partons is always 2% or less. We observed that the possible isospin violation was certainly sufficient to account for the NuTeV $\sin^2 \theta_W$ anomaly, and that the slight preference in the type of violation for valence quarks is indeed in the correct direction to reduce this anomaly. We also discussed the possibility that $s \neq \bar{s}$, noting that other studies have indicated a small inequality, the most recent of these again being in the correct direction to help resolve the NuTeV $\sin^2 \theta_W$ anomaly.

We note that the most significant theoretical errors come from possible corrections to the fixed-order DGLAP framework. Hence within the region of applicability, the conservative partons are the most reliable, and the variation of partons, and of physical observables, under changes in x_{cut} and Q_{cut}^2, give some indication of the theoretical uncertainties, both inside and outside the conservative domain. The most reliable extractions of $\alpha_S(M_Z^2)$ therefore follow from the conservative fits and are equal to

$$\alpha_S^{\mathrm{NLO}}(M_Z^2) = 0.1165 \pm 0.002(\mathrm{expt}) \pm 0.003(\mathrm{theory}), \tag{37}$$

$$\alpha_S^{\mathrm{NNLO}}(M_Z^2) = 0.1153 \pm 0.002(\mathrm{expt}) \pm 0.003(\mathrm{theory}). \tag{38}$$

In the NLO case this is considerably below our default determination, showing that the increase in α_S is mimicking theoretical corrections beyond NLO DGLAP.

As additional tests of theoretical stability, we examined W and Higgs cross sections at the Tevatron and the LHC. At the Tevatron we only sample partons in the conservative domain. Nevertheless the NLO predictions vary significantly as x_{cut} and Q_{cut}^2 are changed due to the re-adjustment of the partons. The NNLO predictions are much more stable, implying less

theoretical uncertainty at NNLO. Indeed, the estimated total theoretical and experimental uncertainty of about $\pm 2\%$ on σ_W at the Tevatron offers an attractive and precise luminosity monitor.

For non-central W and Higgs production at the LHC, we probe partons below the conservative $x_{\mathrm{cut}} = 0.005$. For σ_H, which samples the gluon not too much lower than $x = 0.005$, we still have reasonable stability, especially at NNLO. On the other hand σ_W samples quarks further below $x = 0.005$ and the NLO prediction can vary by up to 20% with different choices of cuts. This implies that the theoretical uncertainty at small x at NLO is rather large. However, at NNLO, σ_W is much more stable (varying by about 3%), suggesting that the theoretical uncertainty has been considerably reduced by the inclusion of NNLO splitting and coefficient functions. We hope to confirm that this is still the case when the complete NNLO splitting functions become available. Assuming that this is so, the total theoretical and experimental uncertainty at the LHC is about $\pm 4\%$. Therefore again it can serve as a good luminosity monitor. Work on resummations may be able to reduce the theoretical uncertainty still further.

Therefore we conclude that theoretical uncertainties can be dominant in some kinematic regions, especially when the physical quantity probes partons at small x and/or small Q^2. There seems to be considerable advantage in working at NNLO, as compared to NLO, but for real precision a few observables may have to await theoretical developments in low x and/or low Q^2 physics.

Acknowledgements

References

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

[2] W.T. Giele and S. Keller, Phys. Rev. **D58** (1998) 094023;
    W.T. Giele, S. Keller and D.A. Kosower, `hep-ph/0104052`.

[3] S.I. Alekhin, Phys. Rev. **D63** (2001) 094022.

[4] CTEQ Collaboration: D. Stump *et al.*, Phys. Rev. **D65** (2002) 014012.

[5] CTEQ Collaboration: J. Pumplin *et al.*, Phys. Rev. **D65** (2002) 014013.

[6] CTEQ Collaboration: J. Pumplin *et al.*, JHEP 0207:012 (2002).

[7] H1 Collaboration: C. Adloff *et al.*, Eur. Phys. J. **C21** (2001) 33.

[8] H1 Collaboration: C. Adloff *et al.*, `hep-ex/0304003`.

[9] A.M. Cooper-Sarkar, `hep-ph/0205153`, J. Phys. **G28** (2002) 2669;
ZEUS Collaboration: S. Chekanov *et al.*, Phys. Rev. **D67** (2003) 012007.

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

[11] M.A.G. Aivazis *et al.*, Phys. Rev. **D50** (1994) 3102;
J.C. Collins, Phys. Rev. **D58** (1998) 094002;
R.S. Thorne and R.G. Roberts, Phys. Rev. **D57** (1998) 1998, Phys. Lett. **B421** (1998) 303, Eur. Phys. J. **C19** (2001) 339;
W.-K. Tung, S. Kretzer and C. Schmidt, J. Phys. **G28** (2002) 983.

[12] H1 Collaboration: C. Adloff *et al.*, Z. Phys. **C72** (1996) 593.

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

[14] E665 Collaboration: M.R. Adams *et al.*, Phys. Rev. **D54** (1996) 3006.

[15] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Eur. Phys. J. **C23** (2002) 73.

[16] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Phys. Lett. **B443** (1998) 301.

[17] E.B. Zijlstra and W.L. van Neerven, Phys. Lett. **B272** (1991) 127; ibid. **B273** (1991) 476;
ibid. **B297** (1992) 377; Nucl. Phys. **B383** (1992) 525;
S. Moch and J.A.M. Vermaseren, Nucl. Phys. **B573** (2000) 853.

[18] S.A. Larin, P. Nogueira, T. van Ritbergen and J.A.M. Vermaseren, Nucl. Phys. **B492** (1997) 338;
A. Retey and J.A.M. Vermaseren, Nucl. Phys. **B604** (2001) 281;
S. Moch, J.A.M. Vermaseren and A. Vogt, Nucl. Phys. **B646** (2002) 181.

[19] S. Catani and F. Hautmann, Nucl. Phys. **B427** (1994) 475;
V.S. Fadin and L.N. Lipatov, Phys. Lett. **B429** (1998) 127;
G. Camici and M. Ciafaloni, Phys. Lett. **B430** (1998) 349;
J. Blümlein and A. Vogt, Phys. Lett. **B370** (1996) 149;
J.A. Gracey, Phys. Lett. **B322** (1994) 141;
J.F. Bennett and J.A. Gracey, Nucl. Phys. **B517** (1998) 241.

[20] W.L. van Neerven and A. Vogt, Nucl. Phys. **B568** (2000) 263; Nucl. Phys. **B588** (2000) 345.

[21] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Eur. Phys. J. **C18** (2000) 117.

[22] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Phys. Lett. **B531** (2002) 216.

[23] N. Kidonakis and J.F. Owens, Phys. Rev. **D63** (2001) 054019.

[24] L. N. Lipatov, Sov. J. Nucl. Phys. **23** (1976) 338;
E. A. Kuraev, L. N. Lipatov, V. S. Fadin, Sov. Phys. JETP **45** (1977) 199;
I. I. Balitsky, L. N. Lipatov, Sov. J. Nucl. Phys. **28** (1978) 338.

[25] S. Catani and F. Hautmann, Ref. [19].

[26] V.S. Fadin and L.N. Lipatov, Phys. Lett. **B429** (1998) 127.
G. Camici and M.Ciafaloni, Phys. Lett. **B430** (1998) 349.

[27] R.S. Thorne, Phys. Lett. **B474** (2000) 372; Phys. Rev. **D64** (2001) 074005.

[28] G. Sterman, Nucl. Phys. **B281** (1987) 310;
D. Appel, P. Mackenzie and G. Sterman, Nucl. Phys. **B309** (1988) 259;
S. Catani and L. Trentadue, Nucl. Phys. **B327** (1989) 323;
S. Catani, G. Marchesini and B.R. Webber, **B349** (1991) 635.

[29] A. Vogt, Phys. Lett. **B471** (1999) 97.

[30] A.H. Mueller and J. Qiu, Nucl. Phys. **B268** (1986) 427.

[31] E.M. Levin and M.G. Ryskin, Phys. Rep. **189** (1990) 267.

[32] J. Kwiecinski, A.D. Martin, R.G. Roberts and W.J. Stirling, Phys. Rev. **D42** (1990) 3645.

[33] J.J. Eskola, H. Honkanen, V.J. Kolhinen, J. Qiu and C.A. Salgado, Nucl. Phys. **B660** (2003) 211.

[34] M.G. Ryskin, plenary talk at XI International Workshop on Deep Inelastic Scattering, DIS03, St Petersburg, April 2003.

[35] V. Abramovsky, V.N. Gribov and O.V. Kancheli, Sov. J. Nucl. Phys. **18** (1974) 308.

[36] J. Bartels and C. Bontus, Phys. Rev. **D61** (2000) 034009.

[37] H. Georgi and H.D. Politzer, Phys. Rev. **D15** (1976) 1829.

[38] Yu.L. Dokshitzer and B.R. Webber, Phys. Lett. **B352** (1995) 451.
M. Dasgupta and B.R. Webber, Phys. Lett. **B382** (1996) 96.

[39] E. Gardi, Nucl. Phys. **B622** (2002) 365.

[40] E. Gardi and R.G. Roberts, Nucl. Phys. **B653** (2003) 227.

[41] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Acta. Phys. Polon. **B33** (2002) 2927.

[42] CCFR Collaboration: W.G. Seligman *et al.*, Phys. Rev. Lett. **79** (1997) 1213;
CCFR Collaboration: U.K. Yang *et al.*, Phys. Rev. Lett. **86** (2001) 2742.

[43] B. Badelek and J. Kwiecinski, Phys. Rev. **D50** (1994) R4.

[44] J. Gomez *et al.*, Phys. Rev. **D49** (1994) 4348.

[45] M. Frankfurt and M. Strikman, Phys. Rep. **160** (1988) 235.

[46] U.K. Yang and A. Bodek, Phys. Rev. Lett. **82** (1999) 2467.

[47] W. Melnitchouk *et al.*, Phys. Rev. Lett. **84** (2000) 5455.

[48] S.I. Alekhin, S.A. Kulagin and S. Luiti, `hep-ph/0304210`.

[49] D. Naples *et al.*, `hep-ex/0307005`.

[50] CCFR Collaboration: A.O. Bazarko *et al.*, Z. Phys. **C65** (1995) 189.

[51] M. Goncharov *et al.*, Phys. Rev. **D64** (2001) 112006.

[52] E.N. Rodionov, A.W. Thomas and J.T. Londergan, Mod. Phys. Lett. **A9** (194) 1799;
F.G. Cao and A.I. Signal, Phys. Rev. **C62** (2000) 015203.

[53] G.P. Zeller *et al.*, Phys. Rev. **88** (2002) 091802.

[54] G.P. Zeller *et al.*, Phys. Rev. Lett. **D65** (2002) 111103; Erratum *ibid* **D67** (2003) 119902.

[55] F. Olness, talk at XI International Workshop on Deep Inelastic Scattering, DIS03, St Petersburg, April 2003.

[56] W.K. Tung, private communication, submitted for XXI International Symposium on Lepton and Photon Interactions at High Energies, Fermilab National Accelerator Laboratory, Batavia, Illinois, USA, 11-16 August 2003.

[57] S. Alekhin, Phys. Rev. **D68** (2003) 014002; JHEP **0302**, (2003) 015.

[58] S. Dawson, Nucl. Phys. **B359** (1991) 283.
A. Djouadi, M. Spira and P.M. Zerwas, Phys. Lett. **B264** (1991) 440.

[59] S. Catani, D. de Florian and M. Grazzini, JHEP **0105**, (2001) 025.

[60] S. Catani, D. de Florian, M. Grazzini and P. Nason, `hep-ph/0306211`.

[61] S. Alekhin, `hep-ph/0307219`.

[62] A. Milsztajn and M. Virchaux, Phys. Lett. **B274** (1992) 221.