

3. Results on New Parton Distributions

With the theoretical and experimental input, methods, and parametrizations described above, we arrive at a *standard set* of parton distributions (the nominal “best fit”), along with a *complete set of eigenvector parton distribution sets* which characterize the entire neighborhood of acceptable global fits in parton parameter space. The study is carried out mainly in the MS-bar (\overline{MS}) scheme.⁵ We now discuss the main features of these results and the related physics issues.

3.1 The New Standard PDF Sets

The standard set of parton distributions in the \overline{MS} scheme, referred to as CTEQ6M for convenience, provides an excellent global fit to the data sets described in Sec. 2.1. The overall χ^2 is 1996 for 1811 data points. Comparison with individual experiments are summarized in Table 1. Because there is little change in the agreement with older experimental data, we focus on comparisons with the new data sets.

3.1.1 Comparison with Data

The fact that correlated systematic errors are now fully included in our fitting procedure allows a more detailed study of the quality of fits than was possible in the past. In the following comparison plots of DIS data and the standard fit, we adopt a different approach than the usual one. Instead of combining statistical errors and systematic errors (assumed uncorrelated) in the plotting the error bars, we show only the statistical errors on the data points; however, the points are shifted by the optimal estimates of the systematic errors, which are determined in the fitting procedure. (Appendix A explains how the shifts are calculated.)

Figure 2 shows the comparison of the CTEQ6M fit to the latest data of the H1 experiment [14, 15, 16]. The extensive data set is divided into two plots: (a) for $x < 0.01$, and (b) for $x > 0.01$. The perfect fit seen in the figure is supported by the χ^2 value of 238.4 for 230 data points. Similarly, Fig. 3 shows the comparison to the latest data from ZEUS [17]. The good agreement is also obvious. The χ^2 per data point ($\chi^2 = 277.3$ for 229 data points) is somewhat higher than the ideal expectation of 1.0. However, a close look at Fig. 3 reveals that the large χ^2 is due mainly to fluctuations in the data points, rather than any systematic disagreement. In Appendix A.3, this point is made more precise by studying in detail the systematic shifts obtained in the comparison of ZEUS data with the fit.

The new PDF analysis agrees well with the older fixed-target DIS experiments, as in previous global analyses, as shown in Figs. [4, 5, 102]. Because we are incorporating the fully correlated systematic errors, the data sets used for the BCDMS and NMC experiments are those obtained at each measured incoming energy, rather than the “combined” data sets (for which the correlation information is washed out) that were used in the past in global fits without systematic errors. This fact is relevant for the comparison between

⁵For the convenience of certain applications, we also present one standard set each of parton distributions in the DIS scheme and at leading order. Cf. Sec. 3.1.3.

theory and experiment, because apparent patterns of deviation could be due to systematic errors. (Cf. Sec. 5 and Appendix A.) Among the DIS data sets, the NMC and CCFR F_2 measurements have somewhat higher χ^2 per data point than the others. In Appendix A.3 we also examine the NMC comparison in more detail. We find that the systematic shifts are all reasonable, and no systematic disagreement of any significance is seen.

Of the new experimental input, perhaps the most interesting and significant in its impact on the parton distributions is the one-jet inclusive cross section from D0 [18]. This new data represents a considerable expansion in kinematic range over the previous jet measurements, by providing the measurement in 5 separate η bins. Our fits agree extremely well with these data, in both shape and magnitude over the full kinematic range, as shown in Fig. 6. The agreement is made even clearer in Fig. 7 where the difference between data and theory is plotted. In Figs. 6 and 7, the error bars are combined statistical and systematic errors.⁶ The new PDF's agree reasonably with the CDF data for central jets [19], shown in Fig. 8.

3.1.2 The New Parton Distributions

Fig. 9 shows an overview of the comparison between the new PDF's and the previous generation of CTEQ PDF's, the CTEQ5M1 set, at $Q = 2$ GeV. In order to exhibit the behavior of the PDF's clearly for both large and small x in one single plot, we choose the abscissa to be scaled according to $x^{1/3}$. Correspondingly, we multiply the ordinate by the weighing factor $x^{5/3}$, so that the area under each curve is proportional to the momentum fraction carried by that flavor in the relevant x range. We see that the most noticeable change occurs in the gluon distribution.

The gluon distribution Fig. 10 gives a more detailed picture of the changes in the gluon distribution at three values of Q : 2, 5, and 80 GeV. For low and moderate values of x , say $10^{-5} < x < 0.1$, the most important constraint is due to the rate of Q^2 -evolution of the DIS structure functions. The HERA data in this region are ever improving in accuracy, but they do not result in sizable changes in the gluon distribution, as seen in Figs. 9 and 10a. (Below $Q = 2$ GeV, one may find larger deviations of the new and old distributions, but extrapolation of PDF's into the low Q region is well-known to be unstable. We will return to this point in Sec. X.)

In the moderate to high x range, $x > 0.01$, the inclusive jet data is now playing a very significant role. The combined effects of the precision DIS and jet data result in a significant shift in the shape of the gluon distribution as seen in Figs. 9 and 10b. The new gluon distribution is significantly harder than for CTEQ5M1 and all MRST PDF sets (cf. Sec. 5) at all Q scales. This behavior is mainly dictated by the inclusive jet data which are in the range $50 < Q < 500$ (GeV) and $0.01 < x < 0.5$. (The D0 data in the higher η bins now allow a higher x reach than the central jet data from previous measurements.) The effect becomes amplified at a low Q scale due to the nature of QCD evolution, and the fact that there is no direct experimental handle on the gluon at large x and low Q .

⁶The correlation matrices are more subtle for the D0 jet data than for the DIS data, because of partially correlated errors, so we use the χ^2 function published by the D0 collaboration.[18]

The enhanced gluon at large x is similar to the CTEQ4HJ and CTEQ5HJ distributions, although somewhat smaller. However, there is an important difference in the significance of the result: whereas the latter PDF sets were obtained specifically for fitting the high p_T jet data points by artificially inflating the weights of these points in the global fit, the CTEQ6M gluon distribution results naturally from the new global fit without any such special emphasis. As we saw in Figs. 6 and 7, the fit is almost perfect. Since CTEQ6M represents the “best fit” in the global analysis, this gluon behavior is quite consistent with all DIS and Drell-Yan data sets used in the fit.

3.1.3 DIS fit

For most applications, the \overline{MS} parton distributions are the most appropriate. But for certain applications PDF’s in the NLO-DIS scheme are preferred. For these purposes we have obtained CTEQ6D by performing independent global fits in the NLO-DIS scheme. (Although, in principle, one could obtain NLO-DIS parton distributions by a simple transformation from a NLO- \overline{MS} set, such a procedure is known to be unacceptable in x regions where the numerical values of the PDF’s for different flavors, which transform into each other, are orders of magnitude apart.) The quality of fit for the NLO-DIS fit is comparable to that of the \overline{MS} case.

3.2 Eigenvector PDF Sets For Uncertainty Analyses

As mentioned in the Introduction, a significant advance in global PDF analysis is the emergence of practical methods to estimate uncertainties of PDF’s and their physical predictions. Using the method of [10], we achieve this purpose by characterizing the behavior of the global χ^2 function in the neighborhood of the minimum by a set of eigenvector PDF sets.

The eigenvector sets are obtained in two steps. First, the full set of parameters described in Sec. 2.5 is probed with the iterative procedure of [9, 10], in order to identify those parameters that are actually sensitive to the input data set.⁷ With the current data, the most sensitive 20 parameters are selected. We then generate the eigenvector PDF sets in the 20 dimensional parameter space as described in Ref. [9, 10], with the remaining parameters held fixed. This results in 40 PDF sets, a + (up) and a – (down) set for each eigenvector direction in the parameter space, in addition to the central CTEQ6M set. Ideally, in the quadratic approximation of the Hessian approach, the χ^2 curves would be symmetric around the minimum, so that only one displacement would be needed for each eigenvector. However, we do observe some asymmetry in certain directions in practice, so we have decided to generate both up and down sets in each eigenvector direction; this provides more information on the behavior of the constant- χ^2 surfaces in the neighborhood of the minimum. The up/down sets, called $S_1^+, S_1^-, \dots, S_{20}^+, S_{20}^-$, correspond to a *tolerance* of $T = 10$ (cf. Ref. [10]); i.e., their χ^2 value is greater than the minimum by 100 ($= T^2$). The range of uncertainty defined by the choice $T = 10$ represents our estimate for the range of PDF behavior that is allowed by the current data, based on a detailed comparison between data and fits described in Appendix A.

⁷In other words, the sensitive parameters are not close to “flat” directions in the overall parameter space.

3.2.1 Uncertainties of PDF's

We use the eigenvector PDF sets $\{S_i^\pm\}$ to estimate the uncertainty range of physical quantities (and of the PDF's themselves) according to the master equation (2.3) which is derived in [10].⁸ In Figs. 13 and 14 we show fractional uncertainty bands for the gluon and u -quark distributions, respectively, at $Q = 3.16$ GeV. The u distribution is the most accurately known of the parton distributions, because deep-inelastic scattering by photon exchange, being proportional to the square of the quark charge, is most sensitive to the u quark. The gluon distribution is the most uncertain of the PDF's; but its uncertainty is smaller than in CTEQ5 as a result of the new DIS and jet data. The reduced uncertainty of $g(x, Q)$ can be seen by comparing Figs. 13 and 14 with their counterparts for CTEQ5 (Fig. 4 of Ref. [10]; the uncertainties shown in [10] are for $Q = 10$ GeV).

It is important to understand that the uncertainty range shown in Fig. 13 or 14 corresponds to the envelope of possible parton distributions that are consistent with the data. A distribution function that produces the extreme at any particular value of x is generally not extreme at other values of x . Thus, a PDF that follows the upper or lower boundary of the uncertainty region at all x would definitely not be consistent with the data.

The CTEQ5M1 and MRST2001 distributions are also shown on Figs. 13 and 14, for comparison, plotted as ratios to CTEQ6. The change from CTEQ5 to CTEQ6 is generally comparable to the current uncertainty bounds, although the full set of CTEQ5 distributions does not give an acceptable fit to the CTEQ6 data set. The global χ^2 for the CTEQ5 PDF's is greater than the minimum by 500. The MRST2001 distributions are rather similar to CTEQ5, except at small x . Differences between CTEQ6 and MRST2001 distributions are discussed further in Sec. 5.

In the Hessian formalism used in this analysis, the uncertainty of any prediction is linearly proportional to the tolerance T . There remain unanswered questions concerning (i) higher than normal fluctuations in some experimental data sets, and (ii) the compatibility between data sets from different experiments. Therefore, it is perhaps realistic to raise the uncertainty estimates obtained from the eigenvector sets by using a somewhat higher tolerance T than the nominal value of 10. One should also be aware that some theoretical uncertainties—corrections for NNLO, resummation, heavy-quark masses, etc.—have not been incorporated in the current analysis.

3.3 Issues And Comments

3.3.1 How much progress has been made?

The new PDF's, especially the quark distributions, do not look very different from the CTEQ5 functions, so it is natural to ask how much progress is made by the latest round of global analysis. To address this question, we first compare the predictions of CTEQ5M1 with the new data. Figures 11 and 12 show two such comparisons, for the H1 and D0 data sets respectively. Although the fits might be regarded as “acceptable”, judging by eye, we can be much more quantitative for the precision data that is now available, by calculating

⁸A simpler, but adequate, alternative method is to calculate the quantity X of interest for each of the 40 sets and take the uncertainty to be the rms deviation from the standard prediction.

the relevant χ^2 's, using the full information on correlated systematic errors.

The results are _____ NOT READY _____.

We see therefore, _____ NOT READY _____.

In Fig. 13 we also see that the CTEQ5 gluon distribution lies at the extreme edge of the estimated uncertainty band. These quantitative comparisons, more detailed than Fig. 11 or 12, indicate the significant progress that has been made.

In [10] we estimated the uncertainties of the CTEQ5 PDF's for the data sets available at that time. (See Figs. 3 and 4 of [10].) The uncertainty bands of CTEQ6 are smaller.

3.3.2 Higher twist terms?

There have been several studies of higher-twist (HT) effects in DIS experiments. [44, 45, 4] If power-law corrections (to leading-twist PQCD) are needed, they will introduce additional nonperturbative degrees of freedom in the global analysis. This would complicate PDF analysis considerably, because the extracted PDF's would then depend on the HT model. Because there is no accepted theory of HT, and HT terms are probably non-factorizable and process-dependent, PDF's obtained with the inclusion of HT terms would no longer be universal.

In the absence of firm theoretical guidance, we first limit the possible size of higher-twist terms by placing reasonable cuts on the kinematic range of data utilized in the fitting program. We then study phenomenologically the need for HT corrections by comparing the quality of fits with and without HT parameters. Using the same kinematic cuts as in previous CTEQ analyses (in particular, $Q \geq 2$ GeV and $W \geq 3.5$ GeV for DIS data), we find that the inclusion of simple phenomenological HT factors, of the type used in the previous literature, does not improve the quality of the fit significantly. We conclude that HT corrections are not needed, and therefore stay within the twist-2 PQCD formalism for our analysis.⁹ Details on our study of this issue are summarized in Appendix B.

3.3.3 Determination of α_s ?

As mentioned in Sec. 2.5, for our standard analysis we input the strong-coupling strength $\alpha_s(m_Z) = 0.118$, based on dedicated measurements from QCD studies at e^+e^- colliders and sum rules in lepton-hadron processes. It is desirable to check that this value is consistent with our global analysis, and, beyond that, to see if the global analysis can provide a useful independent measurement of α_s .

We can study the sensitivity of the global fits to α_s by repeating the fitting with different choices of $\alpha_s(m_Z)$. The resulting variation of the global χ^2 is shown in Fig. 15. As has been true in previous analyses, the minimum in χ^2 occurs at a value of $\alpha_s(m_Z) \simeq 0.115$, somewhat lower than the world average from precision measurements. However, the reduction in χ^2 is not significant. For $\alpha_s(m_Z) = 0.118$, χ^2 is greater than the minimum by only 20, and is thus well within our estimated tolerance of $\Delta\chi^2 < T^2 \approx 100$.

⁹For the reasons mentioned in the previous paragraph, we believe, should there be evidence for HT effects, that it would be more desirable to raise the kinematic cuts and preserve the universal PDF's, rather than introduce ad hoc HT terms which would reduce the usefulness of the PDF's.

According to Fig. 15, any value for $\alpha_s(m_Z)$ in the range 0.11 – 0.12 would be consistent with our analysis. Thus our analysis does not determine α_s with an accuracy that is competitive with other measurements. Even if we were to assume the tighter condition $\Delta\chi^2 < 50$ for the tolerance, we would obtain only $\alpha_s(m_Z) = 0.115 \pm 0.004$.

The phenomenological analysis of NLO α_s is ambiguous in the determination of $\alpha_s(\mu)$, as a function of μ , by the Renormalization Group Equation (RGE). We normally use the standard formula for $\alpha_s(\mu)$, obtained by solving the NLO RGE as an expansion in inverse powers of $\ln \mu$. [43] An alternative, valid in the NLO approximation, is to determine $\alpha_s(\mu)$ by truncating the RGE equation for $\mu d\alpha/d\mu$ at NLO. The dashed curve in Fig. 15 shows χ^2 versus $\alpha_s(m_Z)$ for this alternative definition of $\alpha_s(\mu)$. The difference between the two treatments is small.

3.3.4 Next-next leading order (NNLO) corrections?

We find that the NLO approximation of QCD fits the data on short-distance processes within the experimental error. In the foreseeable future, cross sections and evolution equations will be calculated to the higher NNLO approximation. What role will the more accurate theoretical results play in global analysis of QCD?

The NNLO corrections will be small, of order 1%, for the kinematic regions of most data. As long as the experimental uncertainties are larger than the NNLO corrections, extending the analysis to NNLO will be unnecessary. Near boundaries of phase space the NNLO corrections may be large. However, that would imply the need for resummation to all orders for those kinematic regions, rather than just the NNLO approximations. Inclusion of NNLO formulas is not of immediate concern.

3.4 User Interface

All PDF sets described above are available (at <http://cteq.org>) in the usual CTEQ format, using external data tables. These can be used in the same way as previous CTEQ PDF's.

In addition, to facilitate progress toward a universal user interface, we are preparing a new format, which uses an evolution program and small external files containing only the coefficients for the initial parton distributions, following the “Les Houches Accord on PDF's.” [13]

The ctq6 parametrization of Dec 10					
data set	N	χ^2_{corr}	χ^2_{corr}/N	χ^2_{uncorr}	χ^2_{uncorr}/N
BCDMS p	339	370.0	1.092	336.3	0.992
BCDMS d	251	275.1	1.096	213.2	0.849
H1a	104	109.7	1.055	58.2	0.559
H1b	126	128.7	1.021	109.8	0.871
Zeus	229	277.3	1.211	202.7	0.885
NMCF2p	201	323.5	1.609	233.1	1.160
NMC ratio	123	111.7	0.908	109.6	0.891
NMC rX	13	–	–	–	1.12
CCFR F2	69	109.6	1.589	69.0	1.000
CCFR F3	87	–	–	–	0.505
E605	119	–	–	–	0.866
NA51	1	–	–	–	1.02
cdfLasy	11	–	–	–	0.783
e866f	15	–	–	–	0.547
D0 jet	90	–	0.684	–	–
CDF jet	33	–	1.50	–	–

Table 1: χ^2 for the comparison of theory and data of the 16 experimental data sets. The second column N is the number of data points. The third column is the χ^2 incorporating correlations of systematic errors, Eq. (A.4). The fourth column shows the χ^2 with statistical and total systematic errors combined in quadrature, which has been used in global analyses in the past to compare theory and data.

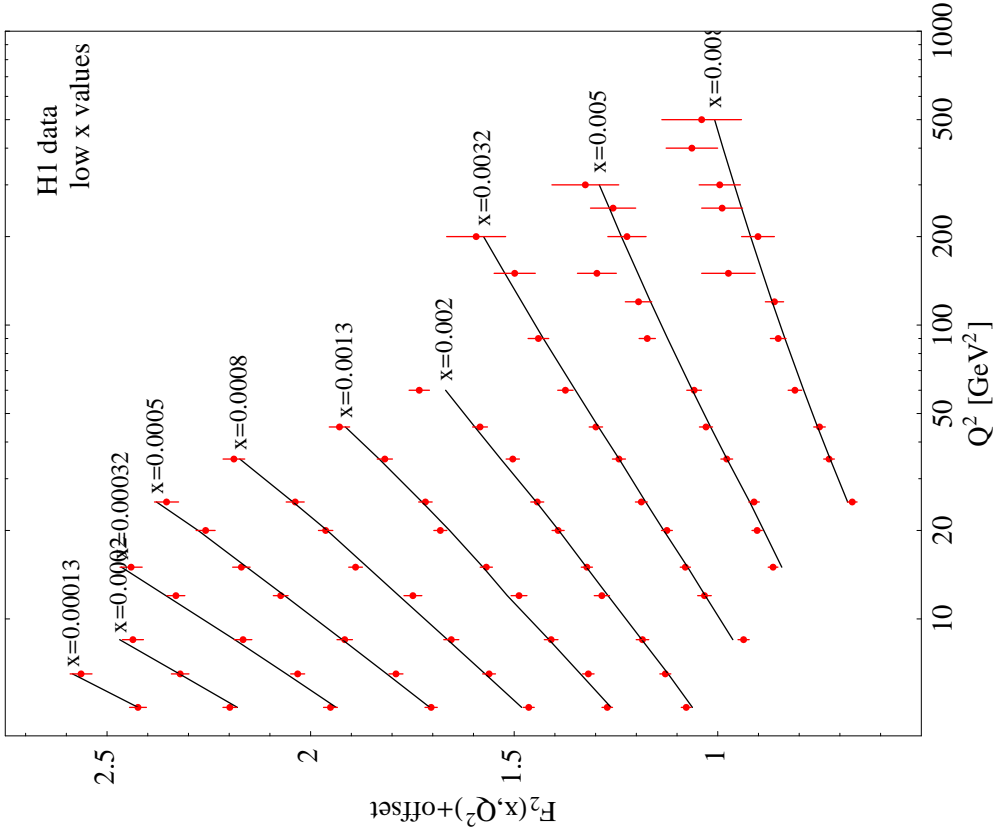
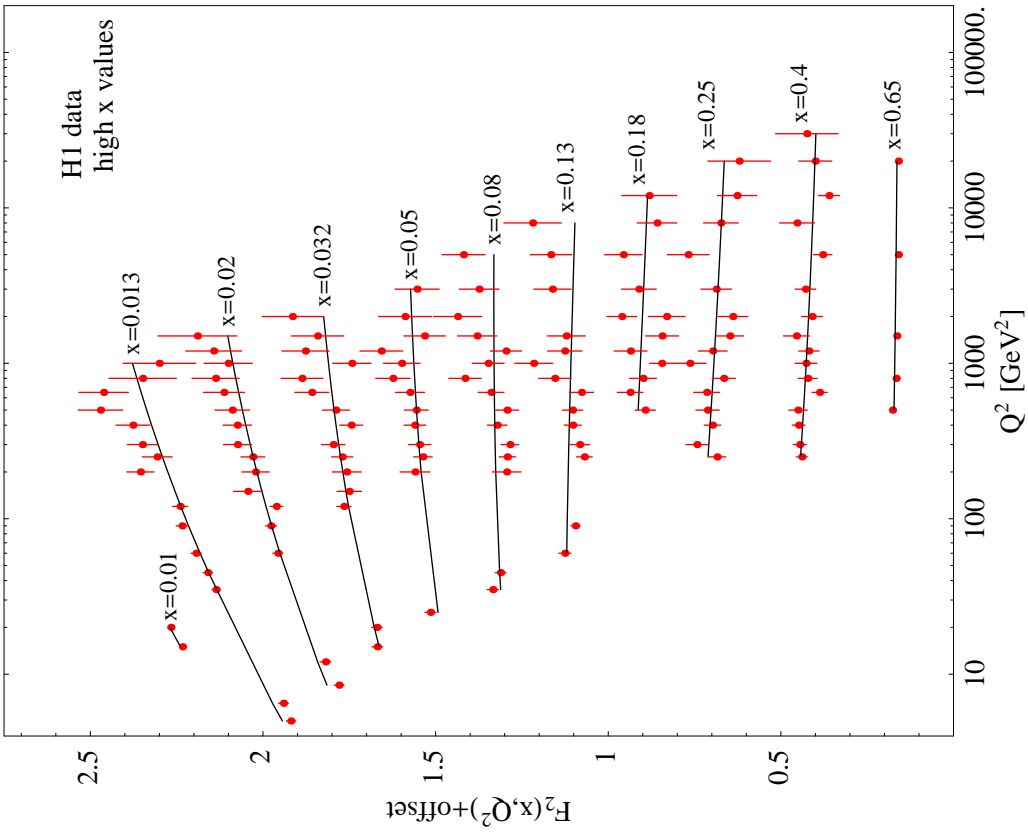


Figure 2: Comparison between theory and the new H1 data. The optimized corrections for systematic errors (cf. Appendix A) have been subtracted from the data. The error bars are statistical errors. (a) Low- x region; (b) high- x region.

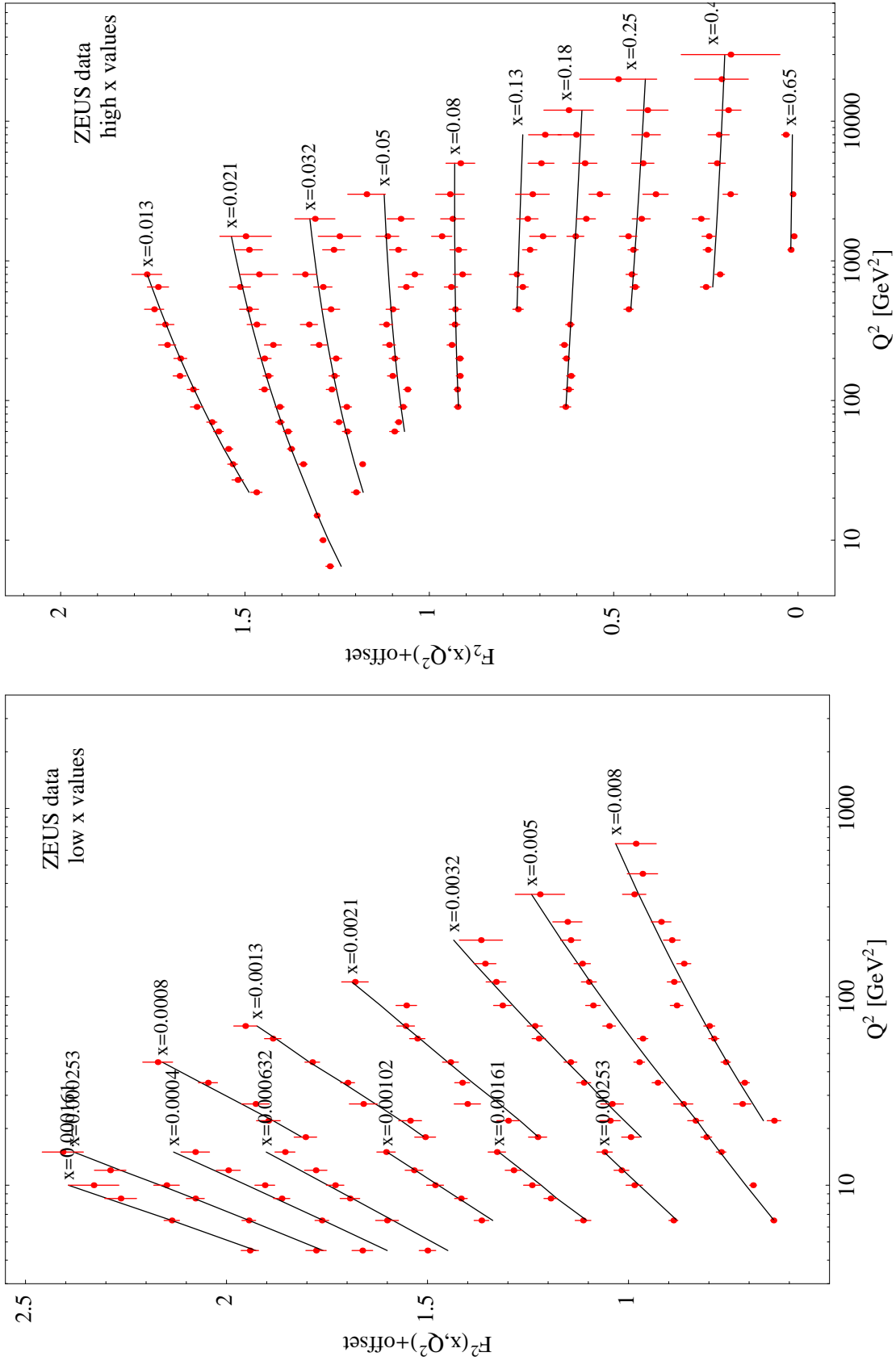


Figure 3: Comparison between theory and the new ZEUS data. The optimized corrections for systematic errors (cf. Appendix A) have been subtracted from the data. The error bars are statistical errors. (a) Low- x region; (b) high- x region.

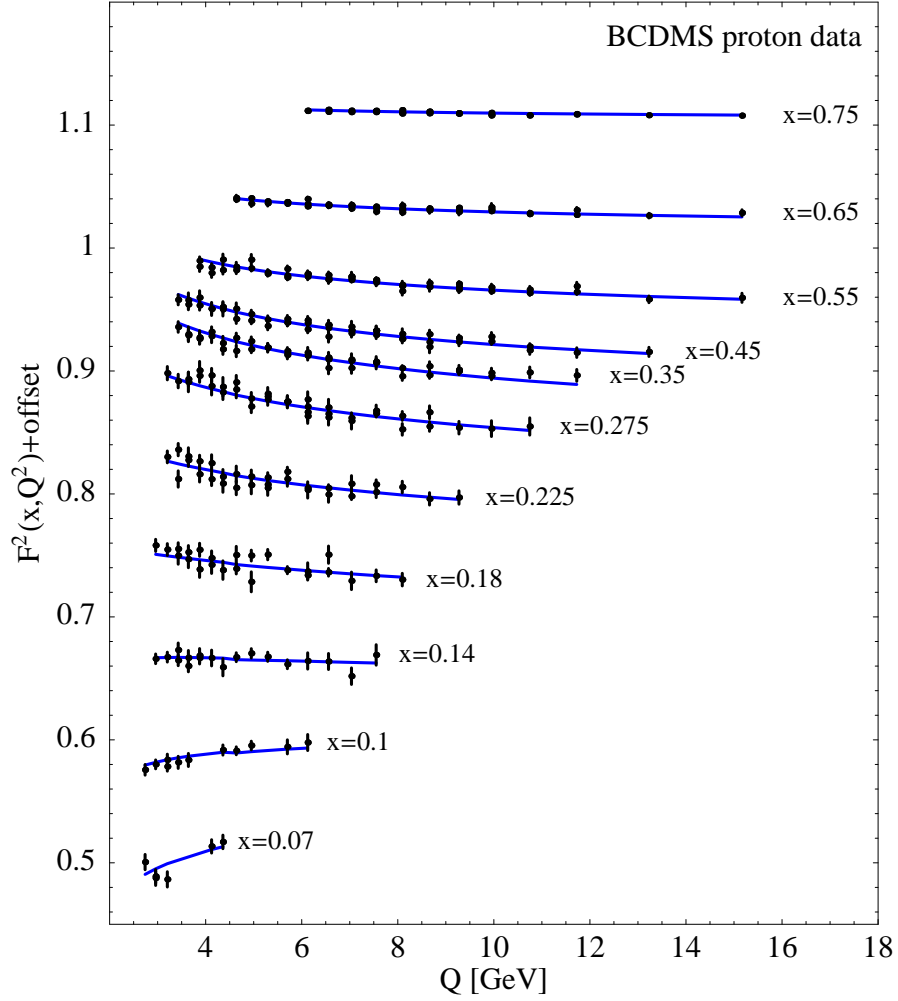


Figure 4: Comparison between theory and the BCDMS data on μp DIS. The optimized corrections for systematic errors (cf. Appendix A) have been subtracted from the data. The error bars are statistical errors.

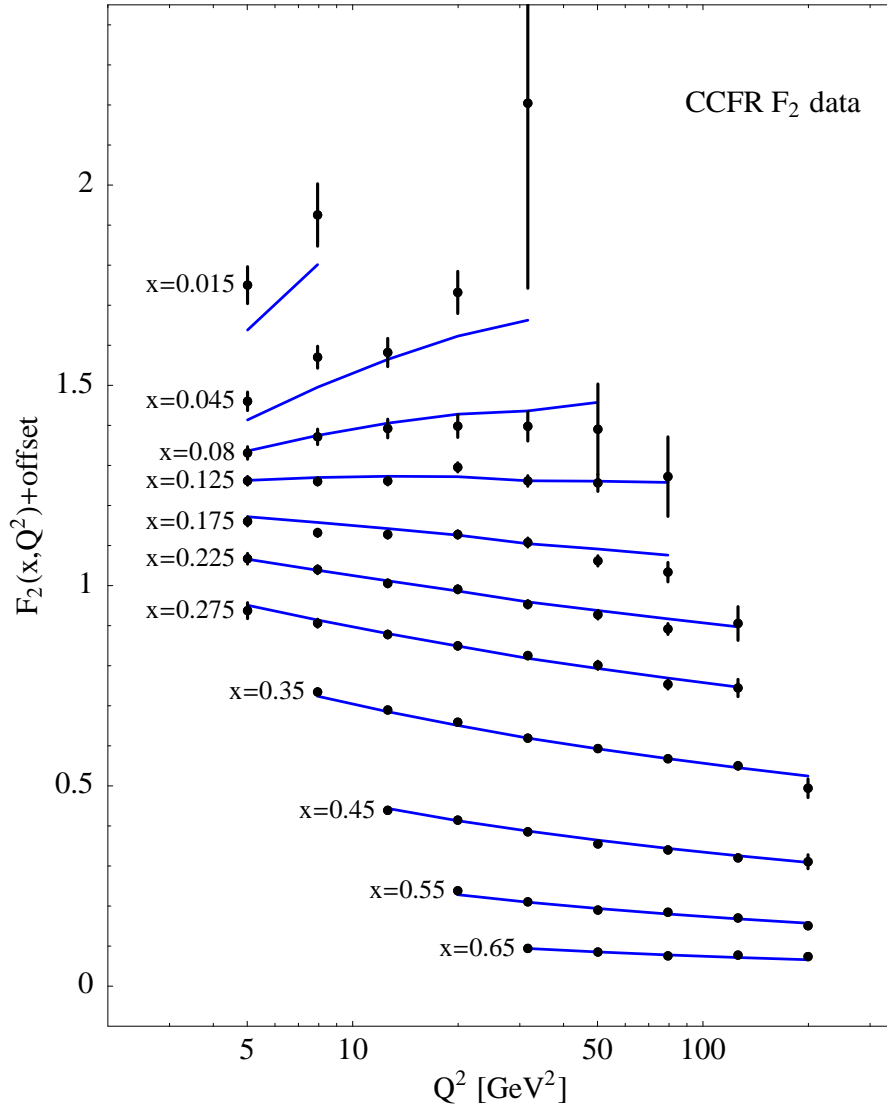


Figure 5: Comparison between theory and the CCFR measurements of F_2 . The optimized corrections for systematic errors (cf. Appendix A) have been subtracted from the data. The error bars are statistical errors.

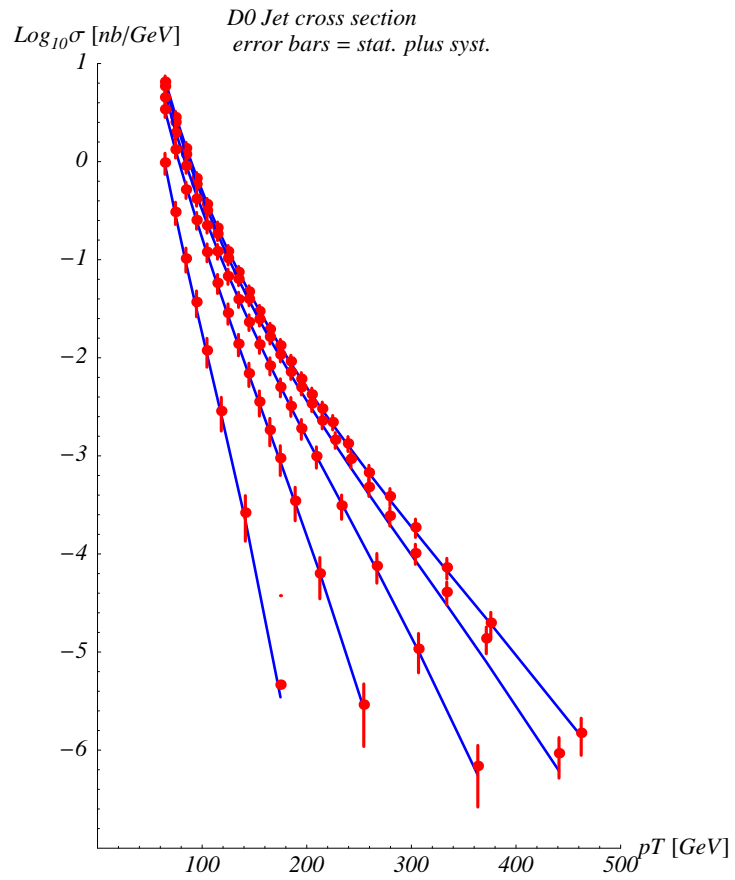


Figure 6: Comparison between theory and the D0 jet data. The error bars are statistical and systematic errors combined.

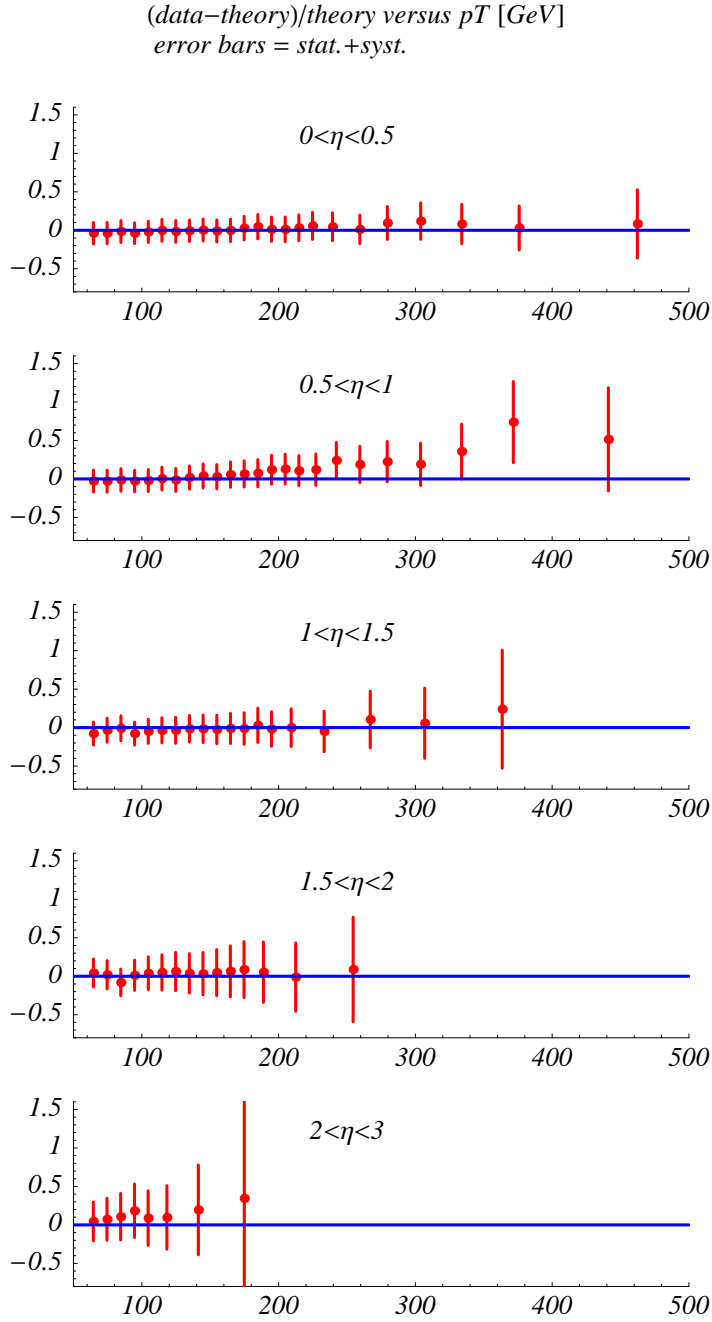


Figure 7: Comparison between theory and the D0 jet data vs. p_T in GeV. The ordinate is $(\text{data}-\text{theory})/\text{theory}$. The error bars are statistical and systematic errors combined.

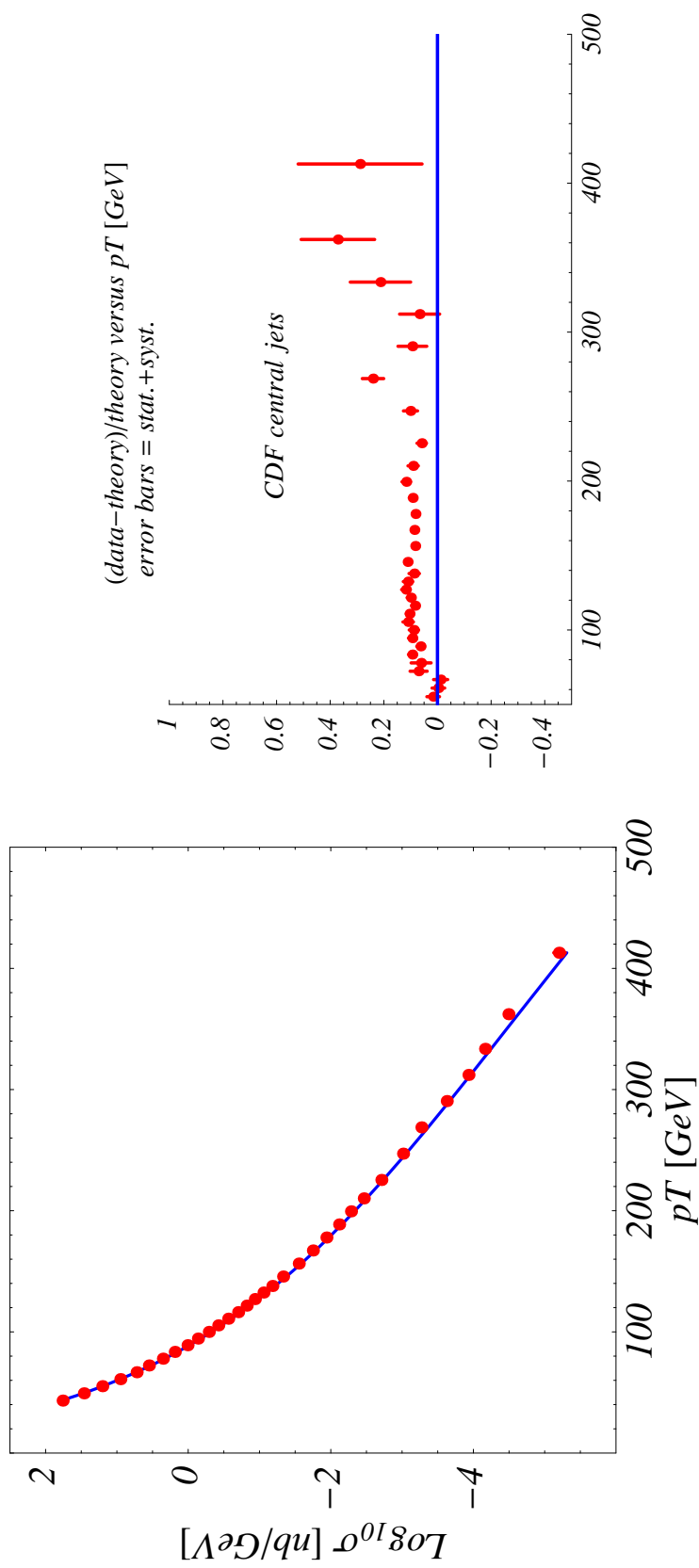


Figure 8: Comparison between theory and the CDF jet data. The error bars are statistical and systematic errors combined.

Figure 9: Comparison of CTEQ6M to CTEQ5M1 PDFs at $Q = 2$ GeV. The x -axis is scaled according to $x^{1/3}$, so that behavior in both the small and large x regions can be seen. The y -axis corresponds to $x^{3/5} f(x, Q)$. In this way of plotting, the area under the various curves represents the momentum carried by the corresponding parton.

Figure 10: Comparison of CTEQ6M to CTEQ5M1 gluon distributions at $Q = 2, 5$ and 80 GeV. The upper graph focus on the small- x region, and the lower one on the large- x region.

Figure 11: Comparison of CTEQ5M1 predictions to the recent H1 data.

Figure 12: Comparison of CTEQ5M1 predictions to the recent D0 data.