Published by Institute of Physics Publishing for SISSA/ISAS
Received: *May 21, 2003*
Revised: *September 30, 2003*
Accepted: *October 21, 2003*

Inclusive jet production, parton distributions, and the search for new physics

Daniel Stump, Joey Huston, Jon Pumplin and Wu-Ki Tung
Department of Physics and Astronomy, Michigan State University
East Lansing, MI 48824 U.S.A.
E-mail: stump@pa.msu.edu, huston@msupa.pa.msu.edu, pumplin@msupa.pa.msu.edu, tung@pa.msu.edu

Hung-Liang Lai
Department of Science Education, Taipei Municipal Teachers College
Taiwan
E-mail: hllai@tmtc.edu.tw

Steve Kuhlmann
Argonne National Laboratory
Argonne, IL 60439, U.S.A.
E-mail: stk@hep.anl.gov

Joseph Francis Owens
Department of Physics, Florida State University
Tallahassee, FL 32306 U.S.A.
E-mail: owens@hep.fsu.edu

Abstract: Jet production at the Tevatron probes some of the smallest distance scales currently accessible. A gluon distribution that is enhanced at large x compared to previous determinations provides a better description of the Run 1b jet data from both CDF and DØ. However, considerable uncertainty still remains regarding the gluon distribution at high x. In this paper, we examine the effects of this uncertainty, and of the remaining uncertainties in the NLO QCD theory, on jet cross section comparisons to Run 1b data. We also calculate the range of contributions still possible from any new physics. Predictions are also made for the expanded kinematic range expected for the ongoing Run 2 at the Tevatron and for the LHC.

Keywords: QCD, Jets, Parton Model, Hadronic Colliders.

JHEP10(2003)046

Contents

1. Introduction: jet cross sections at the Tevatron

Jet production at the Tevatron probes the highest momentum transfer region currently accessible. As this region is potentially sensitive to a wide variety of new physics, there was great interest when the inclusive jet cross section measured by the CDF collaboration in Run 1b exhibited an excess in the highest E_T range, when compared to NLO predictions using then-current parton distribution functions [1]. In an attempt to determine if conventional physics could explain the deviation, the CTEQ PDF fitting group carried out a global analysis using information from deep-inelastic scattering and Drell-Yan data, as well as from jet data from the Tevatron, but giving a higher statistical emphasis to the high E_T jet data from CDF [2]. NLO predictions using the resulting fit (CTEQ4HJ) reduced the size of the excess observed by CDF. The jet data from both CDF and DØ were also used in a more conventional fit (CTEQ4M) where no special emphasis was given to the high E_T data [3].

JHEP10(2003)046

The dominant subprocess that contributes to jet production at high E_T is quark-(anti)quark scattering. However, the quark distributions in the corresponding x range are very well constrained by the precise DIS and DY data used in the global fits. Only the gluon distribution has the flexibility to change significantly in the high x region, and indeed the gluon distribution increases by roughly a factor of 2 for x values of about 0.5 in the CTEQ4HJ fit. The gluon-quark scattering subprocesses increase from approximately 20% of the total jet cross section at high E_T using CTEQ4M to 40% using CTEQ4HJ [2].

The next group of PDF's from CTEQ (CTEQ5) [4] also contained two sets: CTEQ5M, the standard lowest χ^2 solution, and CTEQ5HJ, defined by a similar statistical enhancement applied to the high E_T data from CDF [5] and DØ [6].

More recently, DØ has measured the inclusive jet cross section as a function of rapidity y over the range $0 \leq |y| \leq 3$ [7]. This data set, comprised of 90 data points, has a greater statistical power in the global fits than the CDF and DØ jet cross section measurements in the central region alone. The CTEQ6M fit [8] utilizes these DØ jet cross section measurements, along with the CDF measurements in the central rapidity region, as well as the most recent DIS data from HERA and existing fixed target DIS and DY data. An important step was taken to quantify the uncertainties of the parton distributions and their physical predictions with the inclusion of 40 eigenvector basis PDF sets, along with the best fit, to characterize the uncertainties in the neighborhood of the global minimum, using the hessian method.

This paper is devoted to a detailed and more focused study of the uncertainties of the inclusive jet cross section due to global constraints, both at current energies, and more importantly for further QCD studies and the search for new physics at future collider programs. During the course of this work, we have produced an improved version of the CTEQ6 PDF's. The new minimum set, called CTEQ6.1M henceforth, provides a global fit that is almost equivalent in every respect to the published CTEQ6M [8], although some parton distributions (e.g., the gluon) may deviate from CTEQ6M in some kinematic ranges by amounts that are well within the specified uncertainties. The more significant improvements are associated with some of the 40 eigenvector sets, which are made more symmetrical and reliable in CTEQ6.1M[1]

The comparison between the DØ data for the 5 different rapidity intervals, and the predictions using the CTEQ6.1M PDF's, is shown in figure 1. The theory predictions for CTEQ6M, CTEQ5M and CTEQ5HJ parton distributions are also shown. The greater statistical power that results from including the new DØ jet data, and the fact that all of the rapidity intervals prefer a larger gluon at high x, results in the increases in the cross sections found using NLO predictions based on CTEQ6M or CTEQ6.1M rather than CTEQ5M.

This point deserves further emphasis. It is crucial to measure the jet cross section over a wide rapidity range: any new physics should contribute mostly to the central rapidity region, while a PDF explanation will contribute to all rapidity ranges. This will be addressed in more detail below.

[1]The CTEQ6.1M parton distributions will be available in the CTEQ table format at `http://www.phys.psu.edu/~cteq/` and in the LHAPDF format at `http://vircol.fnal.gov`.

JHEP10(2003)046

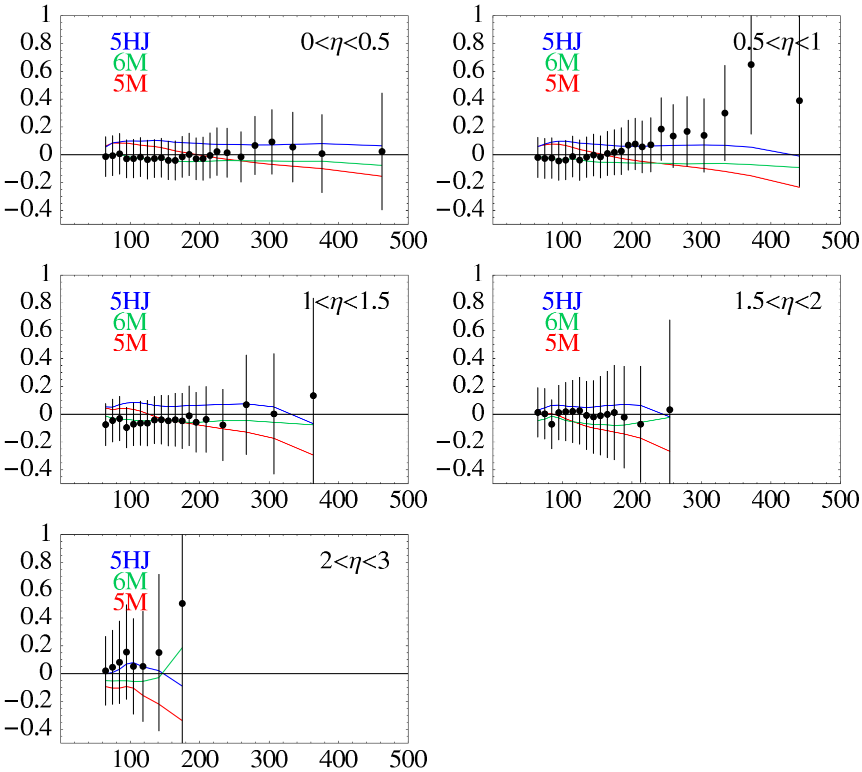

Figure 1: The DØ inclusive jet cross section versus E_T for five rapidity bins compared to NLO predictions using the CTEQ6.1M PDF's. The abscissa is E_T in GeV. The ordinate is (data − theory)/theory. The curves show the CTEQ5M, CTEQ5HJ, and CTEQ6M predictions, as fractional differences compared to CTEQ6.1M. The error bars are the statistical and systematic errors combined in quadrature.

As discussed in ref. [8], it is important to have a parameterization for the PDF's that is flexible enough to allow an increase in the high x gluon, without modifying too much the shape of the gluon distribution at lower values of x. A comparison of the CTEQ5M, CTEQ5HJ and CTEQ6M gluon distributions is shown in figure 2. As can be observed, the CTEQ6M gluon at high x lies between the gluons in CTEQ5M and in CTEQ5HJ, just as the CTEQ6M predictions for the inclusive jet cross sections at high E_T lie between the predictions using the two other PDF's. The CTEQ6.1M gluon distribution is shown by dashed curves in figure 2. At the highest x values the function resembles CTEQ5HJ, and in the intermediate x range it resembles CTEQ6M.

The subprocess contributions to jet production are shown in the next figures, for the central rapidity bin (figure 3) and the forward rapidity bin (figure 4), for the parton distributions CTEQ5M and CTEQ6M.

The dominant difference between the CTEQ5M and CTEQ6M sets is the increased gluon distribution in CTEQ6M at large values of x; the quark distributions are nearly unchanged. This shows up as a decrease in the fractional contributions of the quark-quark subprocesses and as increases in the fractional contributions of the quark-gluon and gluon-gluon subprocesses. It is interesting to note that the gq scattering subprocess is even more important at high E_T for the highest rapidity bin than for the central rapidity bin.

JHEP10(2003)046

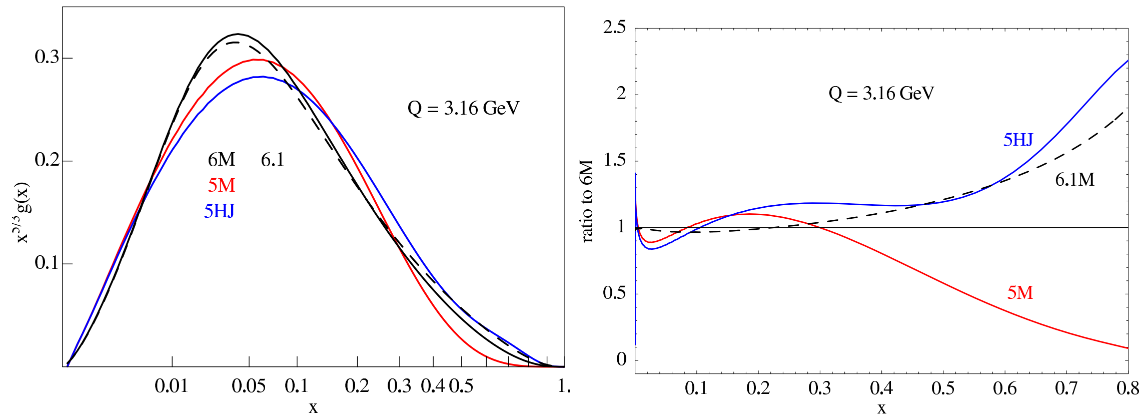

Figure 2: Left: The CTEQ5M, CTEQ5HJ, and CTEQ6M gluon distributions at $Q^2 = 10\,\mathrm{GeV}^2$. Right: The ratios of CTEQ5M and CTEQ5HJ gluon distribution to that of CTEQ6M. The dashed curves show the CTEQ6.1M gluon distribution.

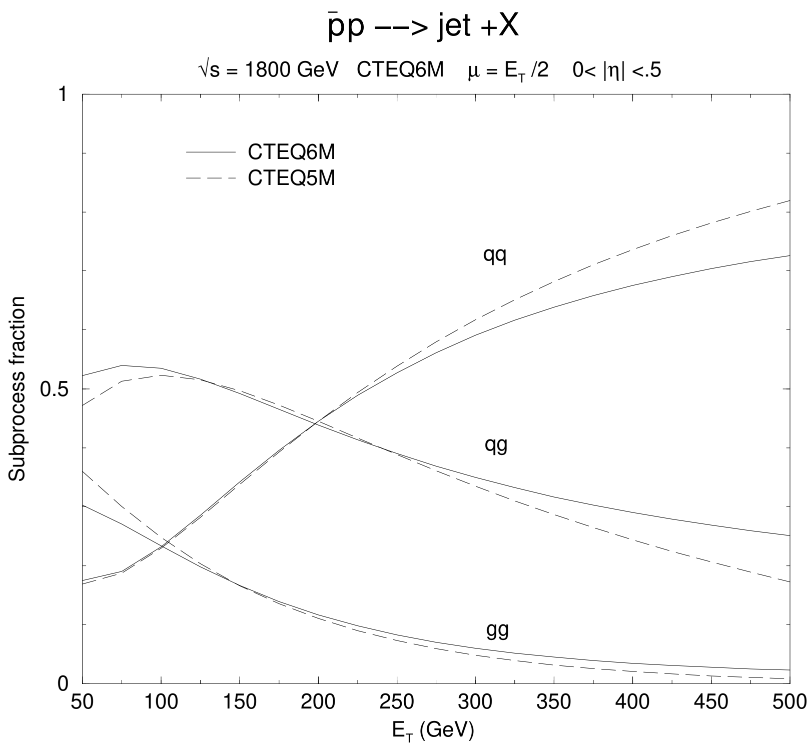

Figure 3: The subprocess contributions to central jet production.

The cross section for production of high-E_T jets at the Tevatron depends on the parton distributions at large x. The relevant range of x may be estimated by considering the leading-order $2 \to 2$ kinematics. Let x_1 and x_2 be the momentum fractions of the two incoming partons. The outgoing parton 4-momenta are

$$p_1'^{\mu} = (p_T \cosh y_1, \vec{p}_T, p_T \sinh y_1)\,, \tag{1.1}$$

$$p_2'^{\mu} = (p_T \cosh y_2, -\vec{p}_T, p_T \sinh y_2)\,; \tag{1.2}$$

the outgoing partons have rapidities y_1 and y_2, and transverse momenta $\vec{p}_T$ and $-\vec{p}_T$,

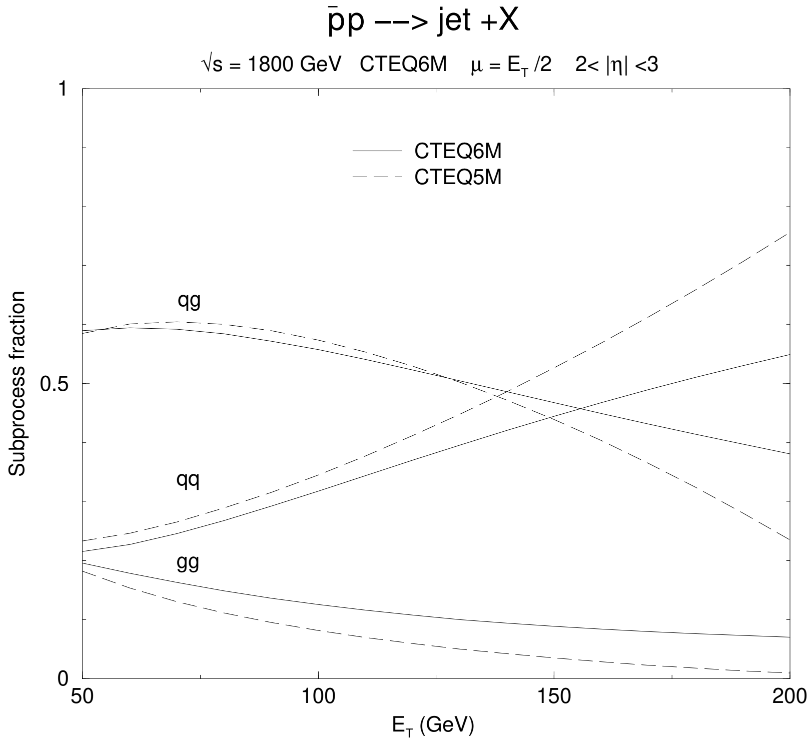

Figure 4: The subprocess contributions to jet production in the forward rapidity region.

respectively. If the first parton is identified with the observed jet then the jet rapidity is $y_j = y_1$ and its transverse momentum is $\vec{p}_T$. The $2 \to 2$ kinematics determines the relation between the momentum fractions and outgoing parton variables,

$$x_1 = \frac{p_T}{\sqrt{s}} \left(e^{y_j} + e^{y_2}\right), \tag{1.3}$$

$$x_2 = \frac{p_T}{\sqrt{s}} \left(e^{-y_j} + e^{-y_2}\right). \tag{1.4}$$

These parametric equations yield a curve in the (x_1, x_2) plane, parametrized by the rapidity y_2 of the second parton, for the specified jet variables.

Figure 5 shows the (x_1, x_2) values for central rapidity ($y_j = 0$) on the left and forward rapidity ($y_j = 2$) on the right. In the central case, the solid curve is the locus of (x_1, x_2) points for $p_T = 400\,\text{GeV}$ and the dashed curve is for $p_T = 200\,\text{GeV}$. Both x_1 and x_2 must be $\gtrsim 0.25$ to produce the high-p_T jet. In the forward case the solid curve corresponds to $p_T = 200\,\text{GeV}$ and the dashed curve to $p_T = 100\,\text{GeV}$. Here one parton must have large momentum fraction to produce the jet at large p_T and forward rapidity. The second parton can have a low value of x, leading to the second jet being close in rapidity to the first jet, or a larger value of x, leading to configurations where the two jets are on opposite sides of the detector. Consider the contribution of the quark-gluon scattering subprocesses. For the configuration where one x is small and the other is large, the smaller x value most likely corresponds to a gluon and the larger to a quark. In this case the difference between the CTEQ5M and CTEQ6M distributions is small. On the other hand, for the configuration where both x values are large the increased gluon in CTEQ6M will enhance the cross

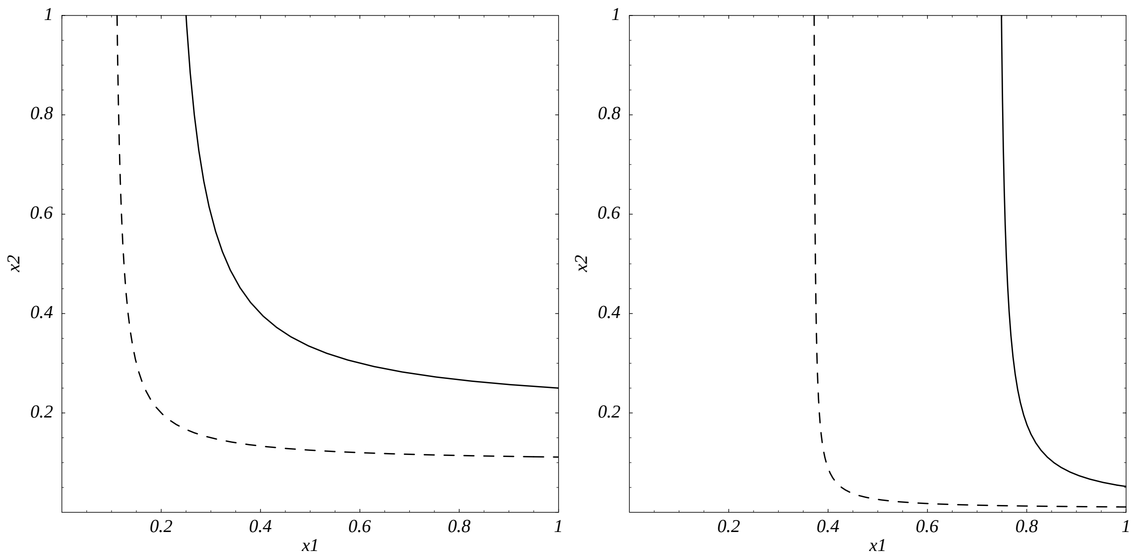

Figure 5: Left: Parton momentum fractions x_1 and x_2 for $y_j = 0$; the solid and dashed curves are for $E_T = 400$ and $200\,\mathrm{GeV}$, respectively. Right: Parton momentum fractions x_1 and x_2 for $y_j = 2$; the solid and dashed curves are for $E_T = 200$ and $100\,\mathrm{GeV}$, respectively.

section. The conclusion is that the increase in the cross section when going from CTEQ5M to CTEQ6M occurs in configurations where the forward jet is balanced in p_T by a jet with roughly the opposite rapidity.

Next, consider the question of where in the jet phase space one is most likely to observe signs of compositeness or other new physics. The angular distributions of the dominant QCD subprocesses are sharply peaked at small scattering angles, whereas the models usually used to estimate compositeness have much flatter distributions. This means that the signal/QCD ratio will be largest for 90° scattering in the parton-parton rest frame. In the overall hadron-hadron frame, this corresponds both jets having the same rapidity. Using the lowest order kinematics described previously, the squared dijet mass is given by

$$M_{jj}^2 = 2p_T^2[1 + \cosh(y_1 - y_2)]\,.$$

Setting $y_1 = y_2$ yields $M_{jj}^2 = 4p_T^2$. This shows that maximizing the dijet mass means maximizing p_T. Since the p_T reach is maximized in the central region, this is the region where the highest dijet masses can be reached with the lowest background for new physics signals (at least those with a flat angular distribution). The model discussed for compositeness later in this paper gives an explicit example of this behavior. On the other hand, as shown by the kinematic examples in figures 5, modifications of the gluon distribution at high values of x will affect the high-p_T jet cross section at all rapidities.

Uncertainties of the gluon and other PDF's arise not only from propagation of experimental errors but also from theoretical uncertainties such as scale dependence, which will be discussed in section 3, and scheme dependence [9]. The overall uncertainty may then be used to set limits on new physics, as described in section 4, and to make meaningful predictions for new experiments in sections 5 and 6.

2. Uncertainty analysis of the cross section for inclusive jet production at the Tevatron

As discussed in the previous section, the cross section for inclusive jet production has been measured by the CDF and DØ collaborations at the Tevatron. These results have been used in the latest CTEQ global analysis of PDF's, leading to the CTEQ6 (and now 6.1M) parton distributions. In this section we analyze the uncertainty of the theoretical QCD cross section due to PDF uncertainty.

2.1 The hessian method of uncertainty analysis

The hessian method of uncertainty analysis has been described in detail previously [10]. For completeness, we summarize the method briefly. The parton distributions are constructed by a method of chi-square minimization with fitting of systematic errors. A chi-square function ${\chi'}^2$ is defined by

$$ {\chi'}^2 = \sum_e \chi_e^2(a,r) \tag{2.1} $$

$$ \chi_e^2(a,r) = \sum_i \frac{\left[D_i - \sum_k r_k \beta_{ki} - T_i(a)\right]^2}{\alpha_i^2} + \sum_k r_k^2 \tag{2.2} $$

where e labels an experimental data set and i labels a data point in the data set. D_i is the data value, α_i is the uncorrelated error, and β_{ki} is the kth correlated systematic error; these numbers are published by the experimental collaboration. Then $T_i(a)$ is the theoretical value, a function of a set of n PDF parameters, $\{a_1, \ldots, a_n\}$. Also, $\{r_k\}$ is a set of gaussian random variables and $r_k\beta_{ki}$ is a (correlated) shift applied to D_i to represent the kth systematic error. We minimize the function ${\chi'}^2(a,r)$ with respect to both the PDF parameters $\{a\}$ and the systematic shift variables $\{r_k\}$. The result yields both the standard PDF model with parameters $\{a_0\}$, and the optimal shifts $\{\widehat{r}_k\}$ to bring theory and data into agreement. This minimum of ${\chi'}^2$ represents the best fit to the data [8].

The hessian matrix is

$$ H_{ij} = \frac{1}{2}\frac{\partial^2 \widehat{\chi}^2}{\partial a_i \partial a_j} \qquad \text{where} \qquad \widehat{\chi}^2(a) \equiv {\chi'}^2(a, \widehat{r}(a))\,. \tag{2.3} $$

This matrix determines the behavior of $\widehat{\chi}^2(a)$ in the neighborhood of the minimum. We have developed an iterative method for calculating the hessian accurately, leading to an improvement in the χ^2 minimization [11]. The point $\{a_0\}$ in the n-dimensional parameter space — where $\widehat{\chi}^2(a)$ is minimum — is the best fit to the global data set. However, points in some small neighborhood of $\{a_0\}$ are also acceptable fits. The variation within the acceptable neighborhood, of a PDF or a physical prediction, is our measure of the uncertainty of the quantity.

In the hessian method of uncertainty analysis, we first define n special directions in the parameter space, namely the n eigenvectors of the hessian. Then for each eigenvector we have two displacements from $\{a_0\}$ (in the $+$ and $-$ directions along the vector) denoted $\{a_i^+\}$ and $\{a_i^-\}$ for the ith eigenvector. At these points, $\widehat{\chi}^2 = \widehat{\chi}_0^2 + T^2$ where

$\widehat{\chi}_0^2 = \widehat{\chi}^2(a_0) =$ the minimum, and T is a parameter called *the tolerance.* We consider any PDF set with $\widehat{\chi}^2 - \widehat{\chi}_0^2 < T^2$ to be an acceptable fit to the global data set. In particular, the $2n$ PDF sets $\{a_i^\pm\}$, which are called the eigenvector basis sets, span the parameter space in the neighborhood of the minimum.

The appropriate choice of tolerance T cannot be decided without a further, more detailed, analysis of the quality of the global fits. After studying a number of examples [10, 12] we conclude that a rather large tolerance, $T \sim 10$, represents a realistic estimate of the PDF uncertainty.[2]

Any quantity X that depends on PDF's has a predicted value $X_0 = X(a_0)$ and a range of uncertainty δX. A simple measure of δX is the spread of the values of $X(a_i^\pm)$ for the $2n$ eigenvector basis sets. However, a more complete uncertainty range is between the minimum and maximum values of X for all points with $\widehat{\chi}^2 - \widehat{\chi}_0^2 < T^2$. It can be shown that in a linear approximation these bounds are $X_0 \pm \delta X$ where

$$(\delta X)^2 = T^2 \sum_{i,j} \left(H^{-1}\right)_{ij} \frac{\partial X}{\partial a_i} \frac{\partial X}{\partial a_j}\,; \tag{2.4}$$

or, in terms of the eigenvector basis sets,

$$(\delta X)^2 = \frac{1}{4} \sum_{k=1}^{n} \left[X(a_i^+) - X(a_i^-)\right]^2 . \tag{2.5}$$

We refer to (2.5) as the Master Equation for the uncertainty of X in the hessian method [10]. Equation (2.5) is based on a linear approximation: $\widehat{\chi}^2(a)$ is assumed to be a quadratic function of the parameters $\{a\}$, and $X(a)$ is assumed to be linear. We find that these approximations are not strictly valid, so instead we calculate asymmetric bounds by

$$(\delta X_\pm)^2 = \sum_{k=1}^{n} \left[X(a_i^\pm) - X(a_0)\right]^2 ; \tag{2.6}$$

the range of uncertainty of X is $(X_0 - \delta X_-, X_0 + \delta X_+)$. (In (2.6) $\{a_i^+\}$ and $\{a_i^-\}$ are the displaced points where $X > X_0$ and $X < X_0$, respectively.)

The use of asymmetric bounds in (2.6) provides a reasonable estimate of the uncertainty when $X(a)$ is not strictly linear, especially for cases where $X(a_+)$ and $X(a_-)$ deviate from $X(a_0)$ by the same sign, so that $X(a_+) - X(a_-)$ in (2.5) could be small. Another, similar procedure to account for departure from linearity was proposed previously [15]. The Sullivan-Nadolsky procedure would also yield asymmetric bounds, somewhat larger but comparable to the bounds $\delta X_\pm$ from (2.6). As will be discussed later, however, the use of asymmetric bounds is not an important issue for our study of jet production.

The uncertainties on the gluon and u-quark distributions are shown in figure 6, for $Q^2 = 10\,\mathrm{GeV}^2$. The u-quark distribution is tightly constrained for $x \lesssim 0.8$, whereas the uncertainty on the gluon distribution grows to a factor of 2 or larger for x values greater than ~ 0.4. For comparison, the CTEQ5M, CTEQ5HJ, MRST2001 [14] and MRST2002 [13] PDF's are shown, compared to CTEQ6.1M, in figure 7, for $Q^2 = 10\,\mathrm{GeV}^2$.

[2]Different choices of the tolerance have been made by other groups. For example, the MRST group has used $T \sim \sqrt{20}$ for one standard deviation and $T \sim \sqrt{50}$ for a 90% confidence level [13].

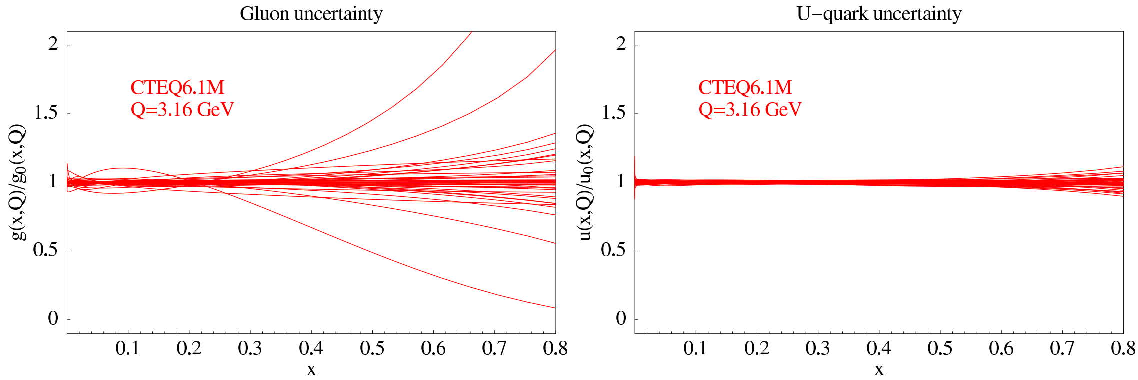

Figure 6: PDF uncertainties. The curves are the ratios of the 40 eigenvector basis sets to the standard set (CTEQ6.1M). The largest variations of the gluon function are the $+$ and $-$ displacements along eigenvector 15.

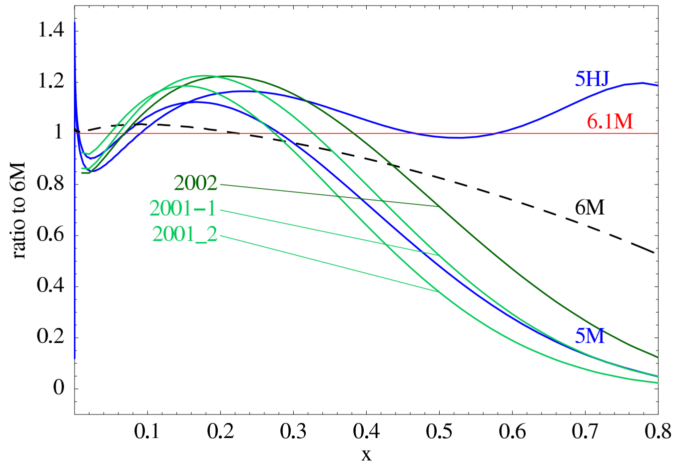

Figure 7: Comparison of several standard PDF sets. For each curve the ratio of the gluon distribution to that of CTEQ6.1M is plotted, at $Q^2 = 10\,\mathrm{GeV}^2$.

2.2 The Tevatron jet cross section

To illustrate our method, and the importance of the systematic errors, figure 8 shows the fractional difference between the data obtained by CDF in Run 1b, and the theory with CTEQ6.1M partons. The error bars are the statistical errors only. The left panel is $(D_i - T_i)/T_i$. Note that there are some systematic differences between the CDF jet data and the theory, both in terms of normalization and in shape.[3] The right panel shows $(\overline{D}_i - T_i)/T_i$ where $\overline{D}_i$ is the shifted data,

$$\overline{D}_i = D_i - \sum_{k=1}^{n} \widehat{r}_k \beta_{ki}\,; \tag{2.7}$$

i.e., the optimal systematic shifts determined by the global fit have been subtracted from

[3] It is important to note all Run 1 CDF cross sections are intrinsically larger than comparable DØ cross sections by approximately 6% due to different assumptions regarding the total inelastic cross section at 1.8 TeV [16].

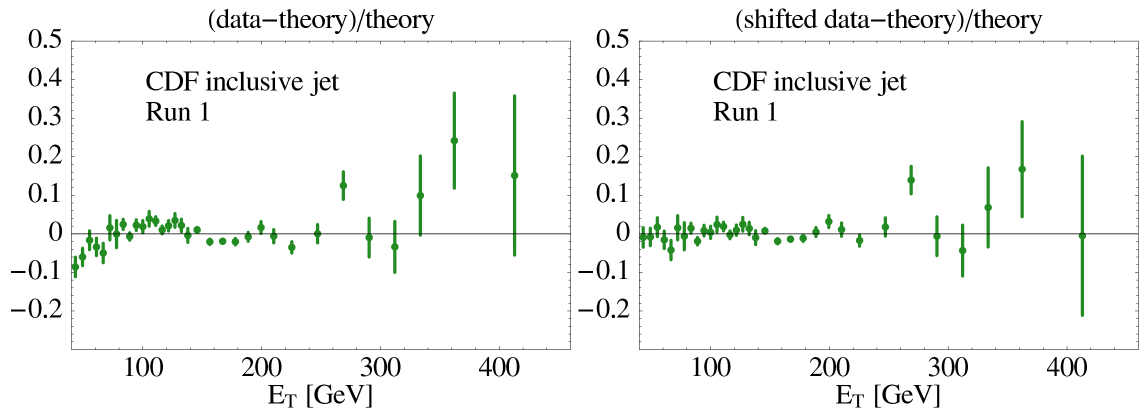

Figure 8: The left panel shows $(D_i - T_i)/T_i$, the fractional difference between data and theory. The right panel shows $(\overline{D}_i - T_i)/T_i$ where $\overline{D}_i$ is the data with systematic shifts subtracted. The error bars are the statistical errors only.

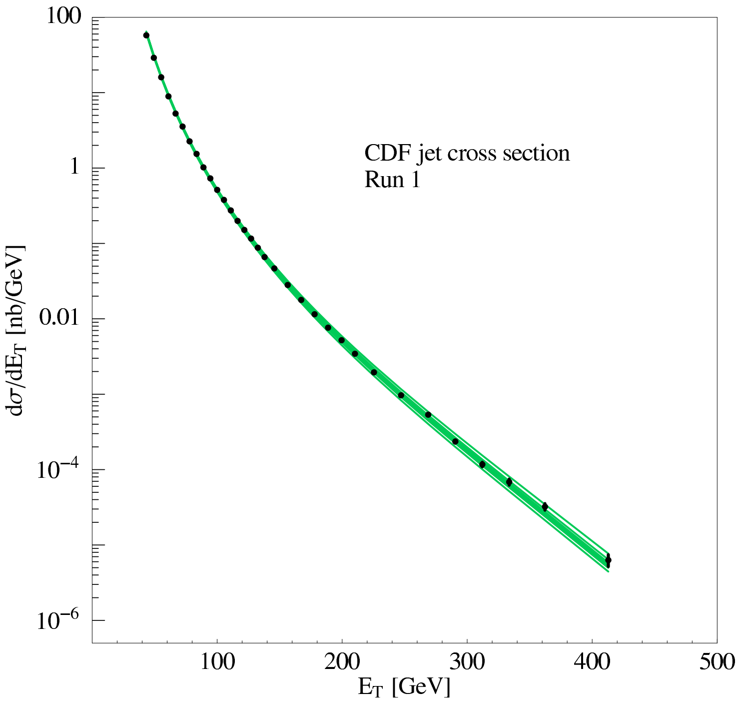

Figure 9: Calculations of $d\sigma/dE_T$ for the 40 eigenvector sets.

the data values. There is no systematic difference between $\overline{D}_i$ and T_i, as expected. The sizes of the systematic shifts are, as expected, comparable to the standard deviations published by the CDF collaboration; that is, $\widehat{r}_k$ is of order 1 for each of the systematic errors.

Figure 9 shows the cross section for inclusive jet production at the Tevatron, integrated over the rapidity range $0.1 < |\eta| < 0.7$, as a function of jet transverse energy E_T, for CTEQ6.1M partons and for the 40 eigenvector sets. (The rapidity interval is that for which the cross section has been published by the CDF collaboration.) The spread of the 41 curves in figure 9 is a simple estimate of the PDF uncertainty. The CDF data points are superimposed on the calculation.

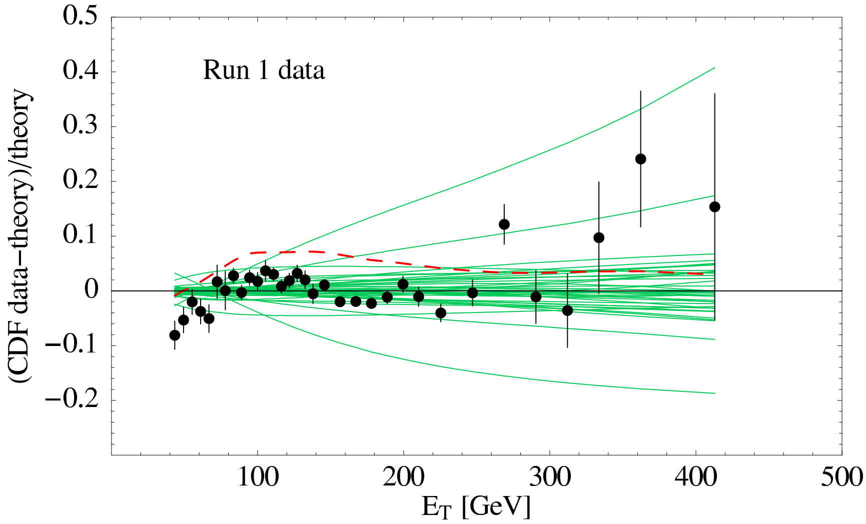

Figure 10: Calculations for the 40 eigenvector basis sets, plotted as fractional differences compared to CTEQ6.1M. The points are the CDF measurements. The error bars are statistical errors only; the systematic shifts are not subtracted from the data. The dashed curve is CTEQ5HJ compared to CTEQ6.1M.

Figure 10 shows the fractional differences $\left[T(a_i^{\pm}) - T(a_0)\right]/T(a_0)$ between the 40 sets and the standard prediction (CTEQ6.1M). The points are the CDF data compared to CTEQ6.1M; the error bars are statistical errors and the data points are plotted *without* the systematic shifts. Also shown is the prediction using CTEQ5HJ partons. The range of uncertainty encompasses both the CDF data and the CTEQ5HJ predictions.

Figure 11 shows the 40 eigenvector basis sets separately. For each eigenvector the two curves are the positive and negative displacements. The hessian method assumes that the variations are approximately linear. If the linear approximation were strictly valid the curves would be mirror images about zero. Figure 11 implies that the linear approximation is reasonably accurate for most eigenvectors, especially those corresponding to the ten largest eigenvalues $(1-10)$. The smaller eigenvalues, which correspond to the "flat directions" in the parameter space, do in some cases show nonlinear behavior; e.g., the $+$ and $-$ deviations are not equal but opposite. The most important eigenvector turns out to be eigenvector 15 (discussed in detail below) which does have a linear behavior.

Figure 12 shows the overall PDF uncertainty calculated from the "Master Equation," but asymmetrically for positive and negative differences. The ordinate is the fractional difference between the extreme cross section and the central (CTEQ6.1M) prediction. The outer curves are the upper and lower bounds of the uncertainty band. The CDF data is superimposed.

The overall uncertainty in the calculation of $(d\sigma/dE_T)_{\rm CDF}$ is dominated by one of the eigenvector directions. The dominant eigenvector is "eigenvector 15," i.e., that for which the eigenvalue of the hessian is 15th in order of magnitude. The $+$ and $-$ displacements in parameter space along this direction have a much larger effect on the value of the cross section than any of the other eigenvector directions. In other words, $(d\sigma/dE_T)_{\rm CDF}$ is most sensitive to the parameter variations along this eigenvector. Figure 13 shows

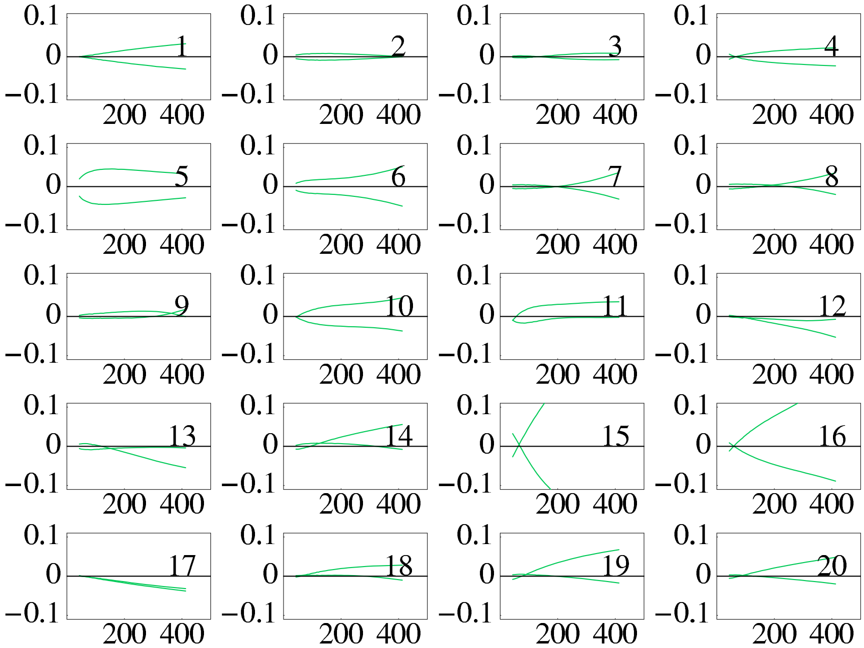

Figure 11: The cross sections for + and − displacements along the 20 eigenvectors, plotted as fractional differences compared to CTEQ6.1M.

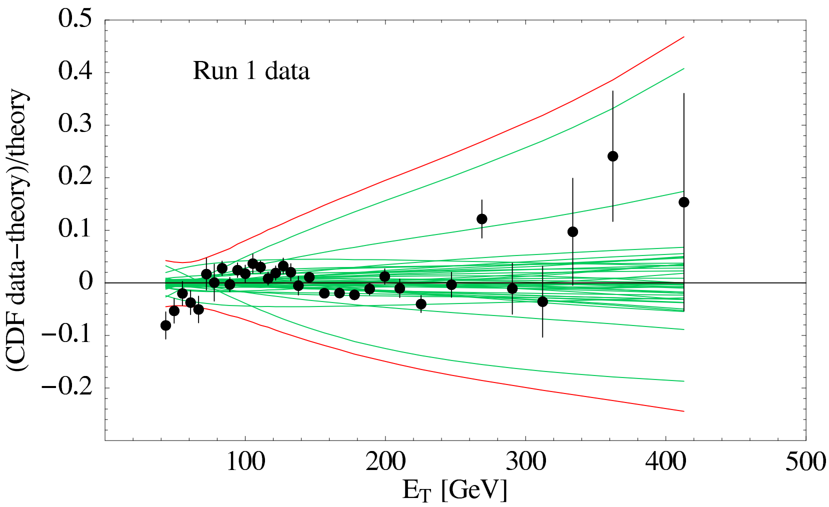

Figure 12: The overall PDF uncertainty of the inclusive jet cross section is plotted as the two red curves. The points are CDF data; the error bars are statistical only and the systematic shifts have not been subtracted.

the gluon distribution for CTEQ6.1M and for the eigenvector basis sets in the + and − directions along eigenvector 15. Eigenvector 15, like the 19 others, corresponds to some linear combination of all of the 20 fitting parameters. The exceptional feature of eigenvector 15 is that it represents the largest change in the high-x gluon behavior.[4]

[4]As noted previously, some of the eigenvectors produce asymmetric errors; however, the uncertainty in the jet cross section is dominated by eigenvector 15 (and to a lesser extent by eigenvector 16) for which the errors are symmetric. Thus the formalism of eq. (2.6) remains valid.

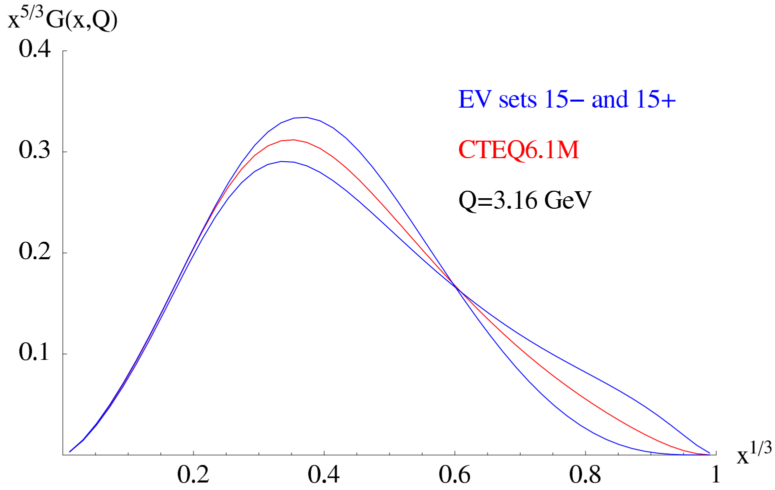

Figure 13: Eigenvector 15. Comparison between the gluon distributions for the + and − displacements along eignevector 15 and the standard set CTEQ6.1M.

PDF set	DØ χ^2	DØ χ^2/N ($N = 90$)	CDF χ^2	CDF χ^2/N ($N = 33$)
CTEQ6.1M	52.26	0.581	57.67	1.748
EV set 15+	111.90	1.243	53.05	1.608
EV set 15−	39.17	0.435	67.35	2.041

Table 1: χ^2/N for the standard PDF's (CTEQ6.1M) and the extreme variations in the + and − directions along eigenvector 15.

In the new CTEQ6.1M analysis, the eigenvector basis sets have been carefully constructed to have $\widehat{\chi}^2 = \widehat{\chi}_0^2 + 100$.[5] But $\widehat{\chi}^2$ is the *global* chi-square function. The criterion $\Delta\widehat{\chi}^2 = 100$ is only a reasonable rule of thumb, and a more thorough analysis — inspecting the fit for each data set — is needed to justify the extreme sets a_{15}^+ and a_{15}^-. We must ask whether these variations could be ruled out by any specific experiment.

The theoretical cross section most affected by displacements along eigenvector 15 is the inclusive jet cross section. Table 1 shows the $\widehat{\chi}^2$ values for the CDF and DØ data for three PDF sets (a_0, a_{15}^+ and a_{15}^-.) Because of the significant systematic uncertainties in the experimental data, represented by the error correlation matrices published by the collaborations, we might conclude that the extreme sets a_{15}^+ and a_{15}^- are still acceptable fits to the data within the systematic errors.

As implied by table 1, much of the increase in χ^2 (50 out of 100) for eigenvector 15+ comes from the DØ jet cross section. The Run 1b DØ cross section can tolerate some increase compared to the CTEQ6.1M predictions, but the increase corresponding to eigenvector 15+ is in fact too extreme. It is not surprising that such a large portion of the increase in χ^2 comes from the DØ jet data: as shown in figure 13, this eigenvector probes

[5]The earlier CTEQ6 analysis [8] made an approximation of the NLO cross sections — holding the K-factor tables for DIS and DY processes constant as the PDF parameters varied in the neighborhood of the minimum — such that $\Delta\widehat{\chi}^2$ was not strictly constant at T^2 for the hessian PDF sets. The CTEQ6.1M analysis is preferred although the changes from CTEQ6 are small.

JHEP10(2003)046

primarily the high-x gluon distribution and the other data sets (with the exception of the CDF jet data) are fairly insensitive to this direction. The opposite direction, eigenvector 15−, gives the lowest χ^2 for the DØ jet data. (Of course the χ^2's of other experiments are larger so that the net change is 100.) The cross section in this case is less than that for CTEQ6.1M, by as much as 10 – 20%. The fact that $\widehat{\chi}^2/N$ is much less than 1 for the DØ data suggests that the size and correlation of systematic errors are not accurately understood and thus CTEQ6.1M gives a satisfactory fit although not the best fit. The changes in χ^2 from the two jet experiments approximately balance in the case of eigenvector set 15−; so the overall increase in χ^2 comes from other experiments in that case. However, the increase is shared by many experiments and no single experiment would rule out the PDFs of eigenvector 15−.

Figures for the DØ jet data analogous to figures 9 and 10 are shown in figures 14 and 15. Figure 14 shows the inclusive jet cross section as a function of E_T for the five rapidity bins of the DØ measurements. Calculations for all 40 eigenvector basis sets are superimposed. The points are the DØ data and the error bars equal the single-point statistical and systematic errors added in quadrature.[6] Figure 15 shows the inclusive jet cross section as a function of E_T for the five rapidity bins, plotted as fractional differences compared to the central (CTEQ6.1M) model.

3. Reliability of next-to-leading-order predictions

Thus far, the discussion of uncertainties has dealt primarily with the issue of the propagation of experimental errors on the observables and their effect on the parameters of the parton distributions. There are, of course, issues related to theoretical uncertainties both in the global fit itself and in the theoretical calculations for jet production. However, it is far more difficult to quantify these. We have previously described studies of some theoretical uncertainties [8]. For example, we have chosen cuts on Q and W for the DIS data sets used for CTEQ6 and earlier analyses in order to reduce the theoretical uncertainties due to effects from higher twist and related terms. Also, we have varied the functional form of the parametrization to ensure that sufficient flexibility to cover the wide kinematic range of x and Q is available. In this section we concentrate on the theoretical uncertainties related to NLO calculations for jet production. The response of the predictions to variations in the renormalization and factorizations scales yields some information on the theoretical uncertainties. In this section the scale dependence of the jet cross sections and the effects on the fits of varying the scales will be discussed. Some estimates of the possible effects of two-loop corrections will also be presented.

3.1 Scale dependence of the jet cross section

Leading-logarithm calculations of the jet cross section at high-E_T generally decrease monotonically as the factorization and renormalization scales are increased. The scale varia-

[6] The systematic errors for the DØ data are supplied by the collaboration only in the form of correlation matrices. Unlike the CDF data, for which the standard deviations of individual systematic errors have been published, we are unable to separate the optimal systematic shifts for subtraction from the DØ data.

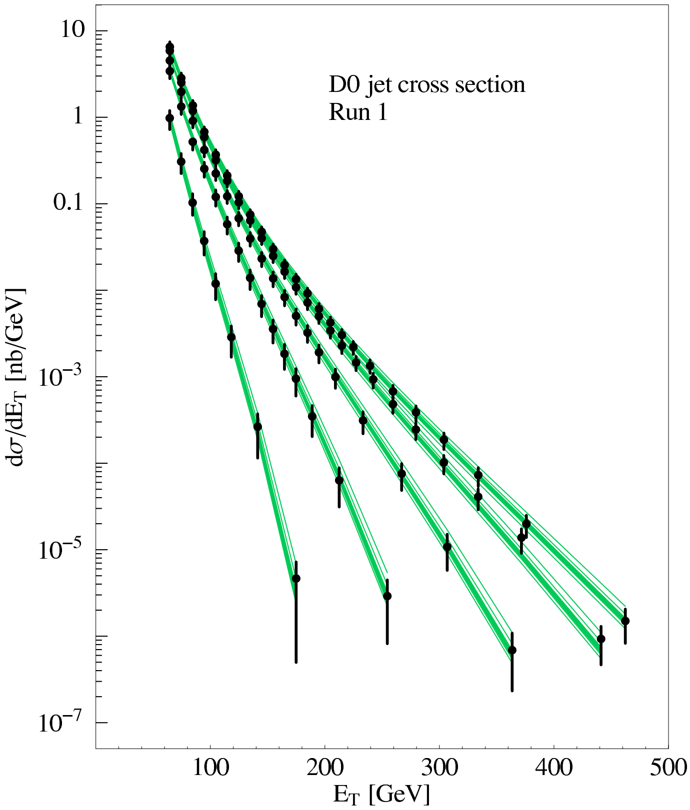

Figure 14: The inclusive jet cross section as a function of E_T for the five rapidity bins of the Run 1b DØ measurements. Predictions of all 40 eigenvector basis sets are superimposed. The points are the DØ data, and the error bars are the statistical and systematic errors combined in quadrature.

tion of next-to-leading-logarithm calculations is generally reduced as compared to that of the leading-log calculations. A brief review of how this comes about is presented in the appendix.

The increased rapidity coverage available with the recent DØ data [7] has provided additional constraints on the gluon distribution. However, if the theoretical uncertainties become large at high rapidity, then these constraints may not be as useful as first believed. As one approaches the edge of phase space, it might be anticipated that there will be large logarithms due to the constraints on multiple gluon emission. The effects of these large logs will be investigated in section (3.3). Their effects may show up already in the next-to-leading-log calculations used in the global fits, so it is important to look for any anomalous behavior in the predictions.

The first step is to examine the scale dependence of the jet cross section predictions in the highest rapidity bin of the DØ data. For ease of presentation, the factorization and renormalization scales have been chosen to be the same. In the jet cross section calculations used for the CTEQ PDF fits, and in this paper, a scale $\mu_R = \mu_F = E_T/2$ has been used. Figure 16 shows the comparison of the scale dependences of the leading-log and next-to-leading-log calculations at $E_T = 70\,\mathrm{GeV}/c$. The monotonically decreasing behavior of the leading-log results is evident, as is the roughly quadratic dependence of the next-to-leading-log curve. Both behaviors are consistent with the expectations discussed

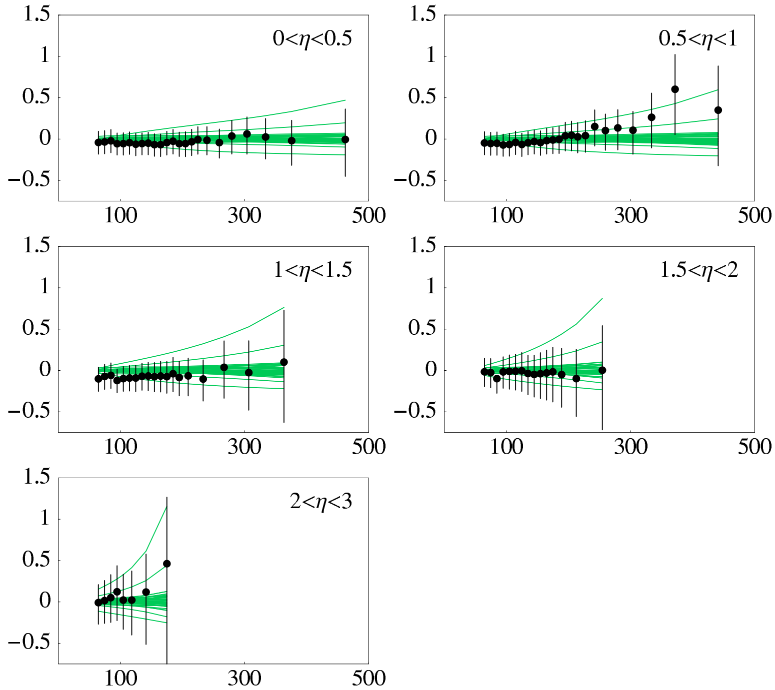

Figure 15: The DØ inclusive jet cross section as a function of E_T for five rapidity bins, plotted as fractional differences compared to the central (CTEQ6.1M) model. The abscissa is E_T in GeV and the ordinate is $[T(a) - T(a_0)] / T(a_0)$ for the curves and $(D - T)/T$ for the points.

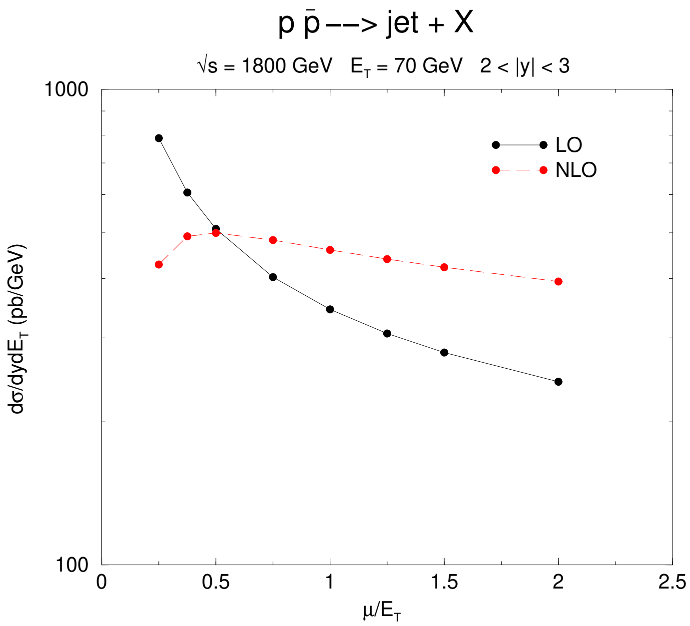

Figure 16: Comparison of the scale dependence of the jet cross section at E_T=70 GeV/c for the leading- and next-to-leading-logarithm calculations.

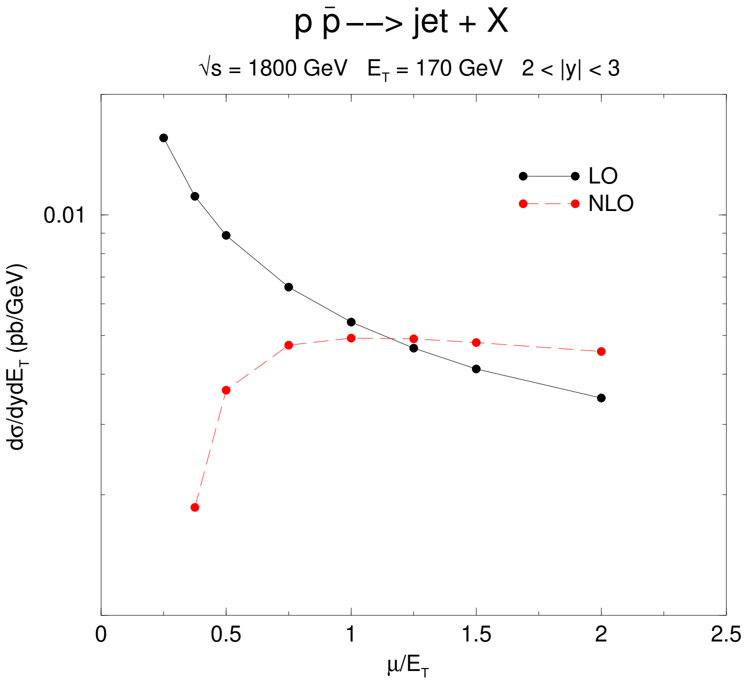

Figure 17: Comparison of the scale dependence of the jet cross section at E_T=170 GeV/c for the leading- and next-to-leading-logarithm calculations.

in the appendix. Similar results are shown in figure 17 for $E_T = 170\,\text{GeV}/c$. Both figures show that the scale dependence of the next-to-leading-log calculation is greatly reduced with respect to that of the leading-log calculation, provided that one avoids scale choices much smaller that $\mu = E_T$.

It is interesting to note that the K-factor (NLO/LO) for $\mu = E_T/2$ is approximately unity for $E_T = 70\,\text{GeV}/c$, while it is significantly less for $E_T = 170\,\text{GeV}/c$. The reasons for this are discussed in more detail in the appendix.

3.2 Fits with different scales

The next point to be addressed is the effect on the fits of varying the choice of scale. To investigate this point, the CTEQ6M fits[7] were repeated with scale choices of $\mu = E_T$ and $2E_T$. As compared to the original fit with $\mu = E_T/2$ the latter two yielded increases in chi-square of 16 and 70, respectively. Thus, the choice of $E_T/2$ used in the CTEQ6M fits yields the best chi-square of the three choices, but this does not constitute a fit for the optimal scale. Note, however, that the rapid increase in the scale dependence for choices below about $E_T/2$ shown in figures 16 and 17 precludes fits with significantly smaller scales while the larger scale choices show an increase in chi-square.

The gluon distributions from these fits are shown in figure 18 relative to the standard CTEQ6M distribution corresponding to $\mu = E_T/2$. The effects at values of x below about

[7]This procedure has not been repeated with the CTEQ6.1M PDF's. However, since the changes in the region of the DØ and CDF data were minimal, the scale dependence in the 6M and 6.1M sets should be essentially the same.

JHEP10(2003)046

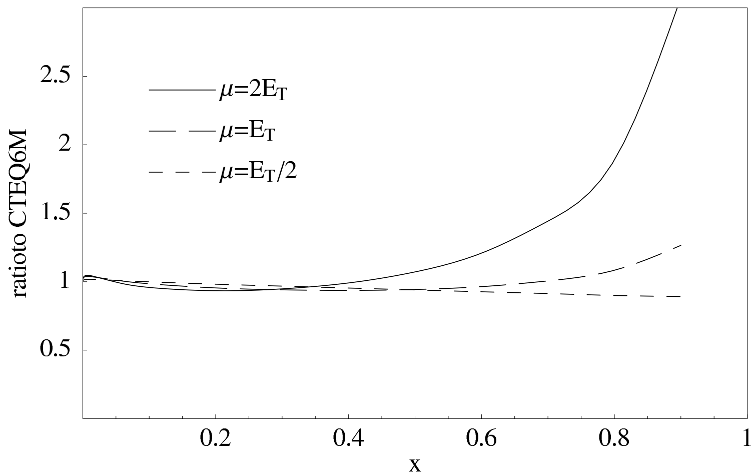

Figure 18: Comparison of fitted gluon distributions for fits made with different scale choices for the jet cross section. Comparison to figure 6 shows that this theoretical error is smaller than the uncertainty due to experimental errors.

0.4 are minimal, but the effects at larger x values are significant. In particular, the reduction in the cross section due to the larger scale choice of $2E_T$ is compensated by an increase of the gluon distribution. However, it should be noted that the chi-square increased by 70 units for this fit. This is still less than the chi-square increase of 100 that determines the 40 error PDF's.

A comparison of figure 18 to figure 6 indicates that the uncertainty on the gluon distribution due to the scale choice for jet production in the global fits is everywhere less than the uncertainty from the treatment of the experimental errors.

It is important to bear in mind that the gluon distribution itself is not an observable quantity. Rather, the gluon and quark distributions contribute via convolutions with appropriate hard scattering cross sections. Accordingly, the impression provided by figure 18 may be somewhat misleading as much of the gluon variation is offset by changes in the hard scattering cross section with which it is convoluted, with the end result being that the variation in the physical observables is much less.

The jet cross section predictions using scales of $E_T/2$, E_T and $2E_T$ are shown in figures 19 and 20 for the DØ rapidity regions of $0.0-0.5$ and $2.0-3.0$. These curves may be compared to those for bins 1 and 5 in figure 15. The scale dependence variations are generally within the bands formed by the 40 eigenvector basis sets, indicating that the uncertainty introduced by the scale dependence of the jet cross sections is less than or on the order of that coming from the experimental errors.

3.3 Threshold resummation

As noted in the introduction to this section, there are potentially significant contributions to the jet cross resulting from phase space limitations on gluon emission as one goes to large E_T and large rapidity. Such corrections can be treated by threshold resummation techniques [17]. In ref. [18] these corrections for jet production were estimated at the two-loop level for central rapidity values and found to be small. The calculation has been

JHEP10(2003)046

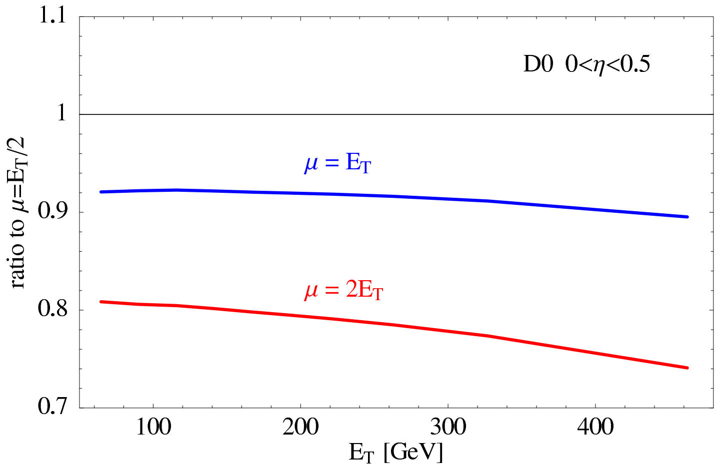

Figure 19: Scale dependence of the jet cross section prediction for $0 < y < 0.5$.

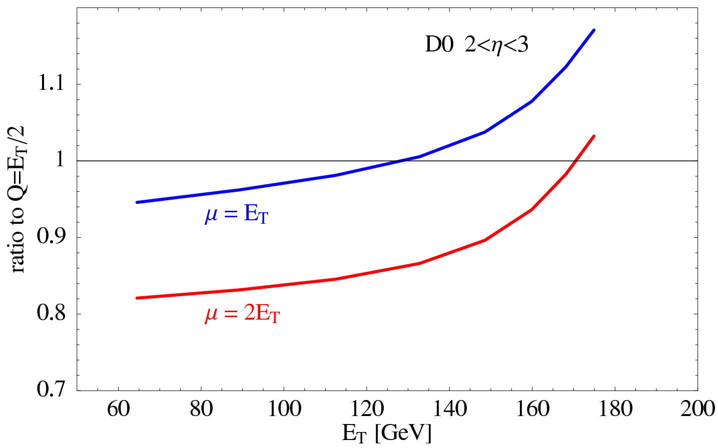

Figure 20: Scale dependence of the jet cross section prediction for $2 < y < 3$.

repeated for the largest rapidity bin covered in the DØ data [7] and the results are shown in figures 21 and 22 by the dashed lines. As was the case in the central region, the two-loop estimates of the next-to-next-to-leading-logarithm contributions are not large for the scale choices considered in this analysis.

To summarize the results of this section, the next-to-leading-logarithm jet calculations show the expected reduction in scale dependence, even for large values of E_T in the highest rapidity bin in the existing data set. Nevertheless, the remaining scale variation does show up in an increased uncertainty in the gluon distribution at large values of x. This uncertainty is, of course, partially offset by the compensating scale dependence of the hard scattering cross section. Even so, the scale dependence of the next-to-leading-log calculations remains as a contributing factor for the overall uncertainty of the predictions for new processes. Finally, the next-to-next-to-leading-log corrections estimated by expanding the threshold resummation predictions for jet production remain small, even at high E_T in the highest rapidity bin considered, lending support to the use of the next-to-leading-log expressions in the global fit.

JHEP10(2003)046

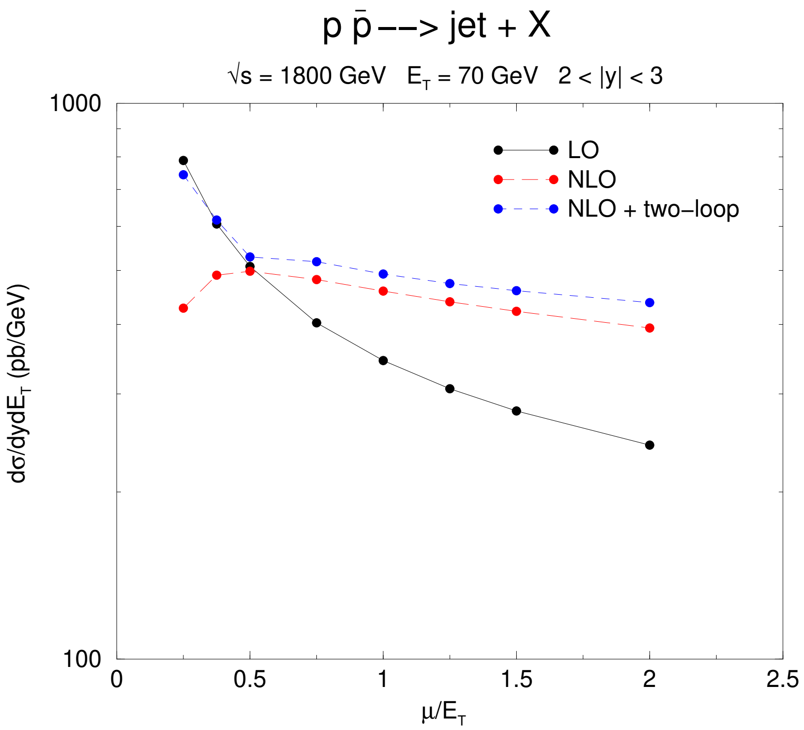

Figure 21: Same as figure 16 with the addition of the estimated two-loop contribution expected from threshold resummation.

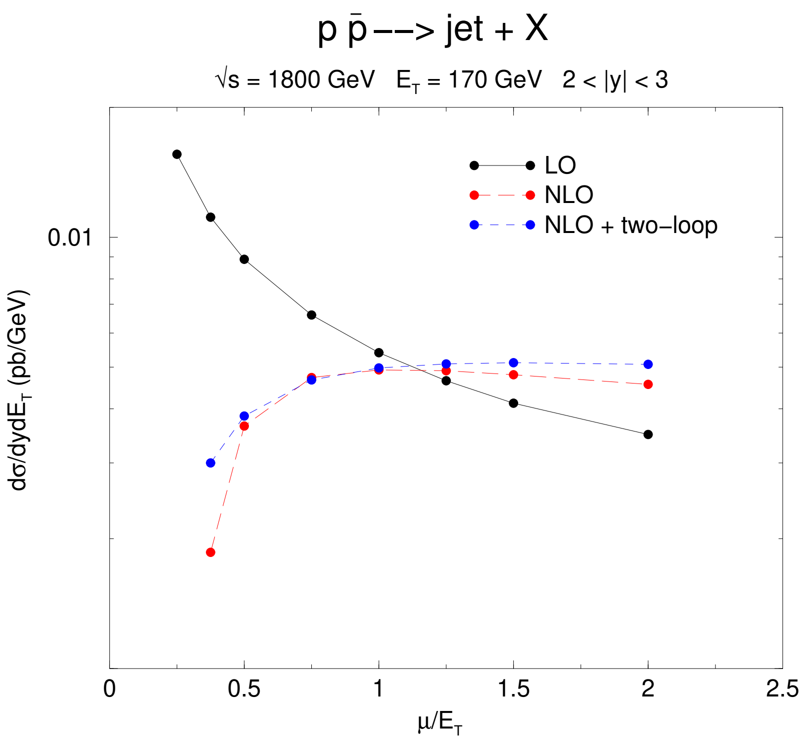

Figure 22: Same as figure 17 with the addition of the estimated two-loop contribution expected from threshold resummation.

JHEP10(2003)046

4. Room for new physics in run 1b

Hadron collider jet data have traditionally been used to constrain models of new physics, usually models of quark substructure. The measurements include dijet angular distributions that are insensitive to theoretical uncertainties due to parton distributions [19]. Inclusive jet cross sections or dijet mass distributions typically provide better probes of quark compositeness, but are sensitive to parton distributions at large x which have large uncertainties [20]. Currently the best limits are provided by a hybrid approach that uses the ratio of dijets in different rapidity bins [21]. This approach has the advantage that many theoretical and experimental systematic errors cancel out in the ratio. It has the disadvantage that it removes some useful information, such as the absolute values of the cross sections in the two rapidity intervals.

There are several reasons to revisit this issue. First, the published limits are based on parton distributions several generations old, and the best-fit gluon distribution at large x is significantly changed. Second, the recent jet measurements at large rapidity provide constraints on the large x partons while being largely insensitive to new physics. Finally, and most important, the global analysis machinery has never been used in these analyses to refit the parton distributions in the presence of a new component to the QCD calculation. If the constraints on the parton distributions are not sufficient, it is possible that new physics could hide inside the PDF uncertainties, or current limits may be weakened.

Compositeness signals are typically parameterized by a mass scale Λ which characterizes the quark substructure coupling. The coupling is approximated by a four-Fermi contact interaction giving rise to an effective lagrangian. Only the term describing left-handed coupling of quarks and anti-quarks has been calculated [22],

$$\mathcal{L}_{qq} = A\left(\frac{2\pi}{\Lambda^2}\right)(\overline{q}_L\gamma^\mu q_L)(\overline{q}_L\gamma_\mu q_L) \qquad \text{where} \qquad A = \pm 1\,. \tag{4.1}$$

The simplest squared amplitude is the quark anti-quark s-channel annihilation process, with the first term the normal QCD interaction and the second due to substructure:

$$\left|A(u\overline{u} \to d\overline{d})\right|^2 = \frac{4}{9}\alpha_S^2(Q^2)\frac{(\widehat{u}^2 + \widehat{t}^2)}{\widehat{s}^2} + \frac{A\widehat{u}}{\Lambda^2}\,. \tag{4.2}$$

Since $\widehat{u}$ is negative, constructive (destructive) interference in this process occurs when $A = -1$ $(+1)$. This is generally true for the other scattering terms as well, which will be included in the calculations. As in the experimental papers, we will use the quark compositeness implementation in PYTHIA [23] to form the ratio of QCD+substructure/QCD, then multiply this ratio by the NLO QCD calculation for each iteration of parton distribution as the PDF parameters are being fit. As expected, we find larger cross sections with the constructive interference choice $A = -1$, approximately a factor of 10 larger. However, the ratio of the first two eta bins used by DØ is slightly larger with the destructive interference term $A = 1$, and this is what is used for the current best limit of $\Lambda > 2.7$ TeV [21]. Therefore we have used the destructive interference choice for illustration in the rest of this paper.

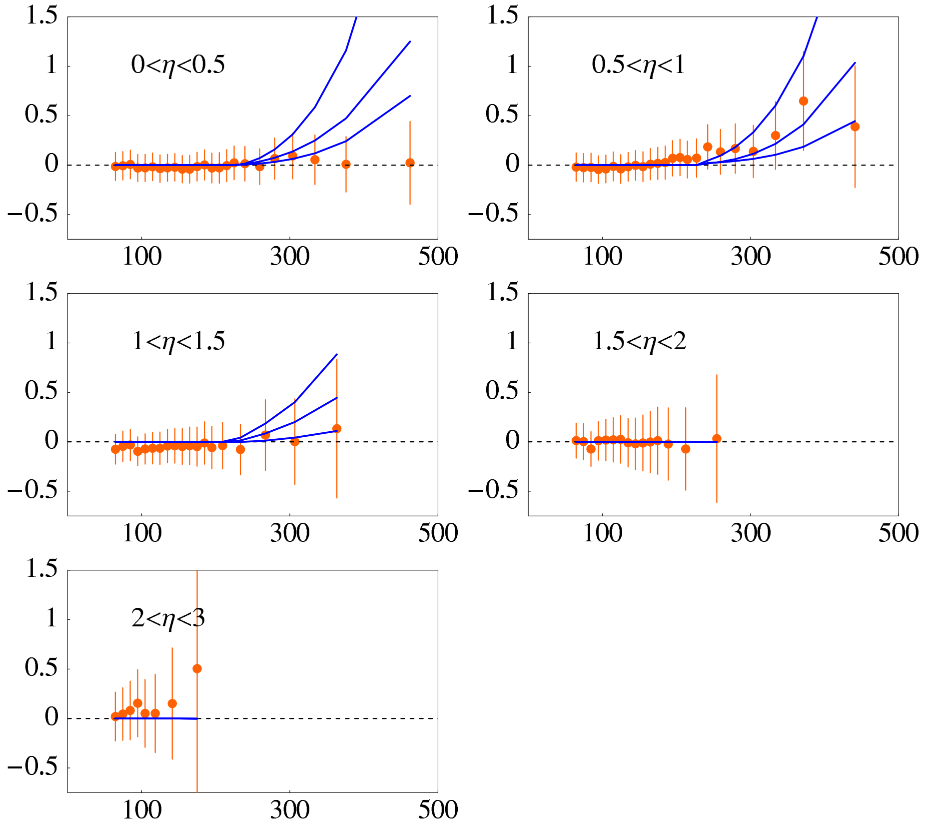

Figure 23: Comparison of three models in which a hypothetical contact interaction contributes to jet production. The ordinate is the fractional difference between the model and the standard theory; the abscissa is the jet E_T in GeV. The contact interaction has $\Lambda = 1.6$, 2.0, and 2.4 TeV for the three curves. In all cases the PDF's are the standard CTEQ6.1M. The DØ data are superimposed for comparison.

Figure 23 shows the cross section for jet production including a contact interaction, for three values of Λ: 1.6, 2.0 and 2.4 TeV. The cross section is plotted as a ratio to the pure QCD prediction with the CTEQ6.1M parton distributions. The range of kinematic parameters are the same as the DØ data, and the DØ data is superimposed on the plots with error bars of statistical and systematic errors combined in quadrature.

If the effect of the contact interaction is smaller than the PDF uncertainty of the QCD prediction then the data cannot rule out the compositeness model; any difference between theory and data could be explained either as a PDF effect or as a sign of new physics. However, if the compositeness model disagrees with the data by an amount that is larger than the PDF uncertainty, the model is ruled out. The PDF uncertainty of the QCD prediction is comparable to the experimental error bars (see figure 15), so we may use those error bars to judge whether the effect of the contact interaction is smaller than the PDF uncertainty. The contact interaction is seen to be ruled out for $\Lambda = 1.6$ TeV, primarily from the disagreement with the central rapidity interval, $0 < \eta < 0.5$. The cross section at large rapidity ($\eta > 1.5$) is not sensitive to the contact interaction.

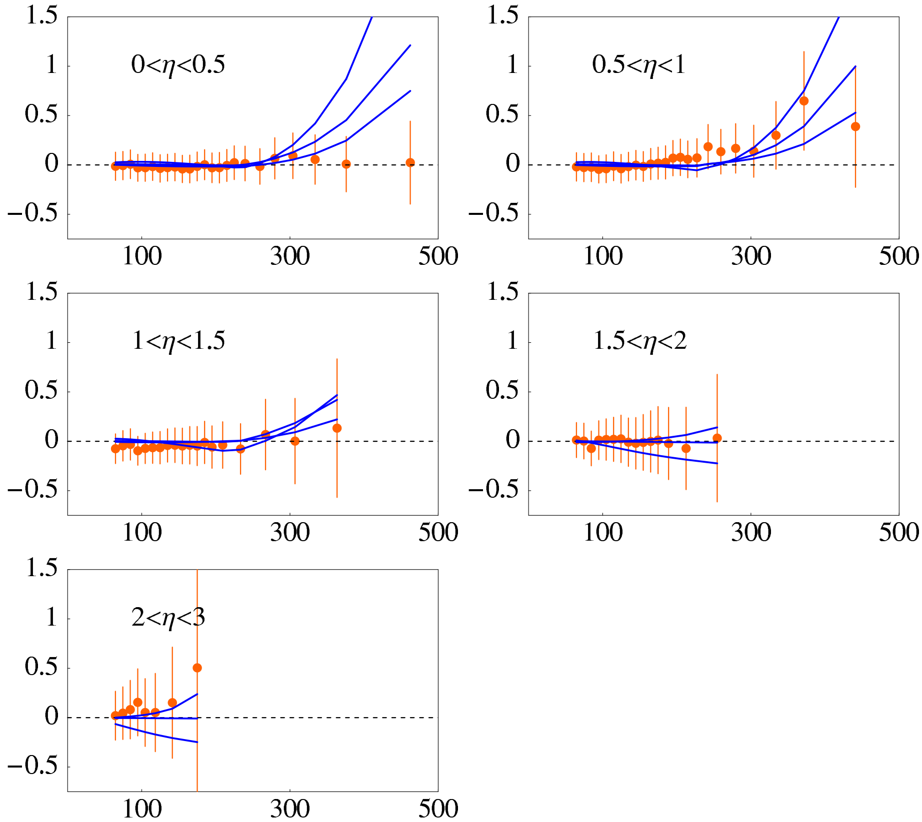

Figure 24: Comparison of three models in which a hypothetical contact interaction contributes to jet production, with refitting of the PDF's. The curves have the same meaning as in figure 23. The PDF's for the three models have been refit to the global data set to give the best fit including the contact interaction in jet production.

The largest two Λ values, $\Lambda = 2.0, 2.4\,\text{TeV}$, are not inconsistent with the DØ data on inclusive jet production because the effect of the contact interaction is within the PDF uncertainty. The global fit for $\Lambda = 2.4\,\text{TeV}$ is even slightly better than for the pure QCD model, although the difference in the overall quality of the fit ($\Delta\chi^2_{\text{global}} = 9$) is much less than our standard tolerance for PDF uncertainty ($\Delta\chi^2_{\text{global}} \lesssim 100$).

If there is a contact interaction, then the CTEQ6.1M PDF's would be inaccurate because the data were fit to a pure QCD model. Therefore a complete analysis of the contact interaction requires that the PDF's be refit, comparing the data to the theory with compositeness. Figure 24 shows the final result of this analysis. The cross section for jet production with the DØ kinematic parameters is again shown for three values of the Λ parameter (1.6, 2.0 and 2.4 TeV) as a ratio to the pure QCD model with CTEQ6.1M. In figure 24 the PDF's for each Λ value come from a global fit that includes the contact interaction in the theory of the jet cross section. The smallest value of Λ is now more consistent with the data, but still ruled out.

The limit derived from this analysis, $\Lambda > 1.6\,\text{TeV}$, is not as strong as those from either the CDF or DØ analyses that use angular distribution information [19]. This is a result of the remaining uncertainty in the PDFs, and the fact that in this model of new

Λ [TeV]	PDF set	global χ^2	DØ χ^2/N	CDF χ^2/N
1.6	CTEQ6.1M	2111	0.939	5.45
2.0	CTEQ6.1M	1966	0.566	2.07
2.4	CTEQ6.1M	1948	0.493	1.73
∞	CTEQ6.1M	1957	0.590	1.68
1.6	best refit	2079	1.230	3.63
2.0	best refit	1965	0.576	2.00
2.4	best refit	1947	0.493	1.73

Table 2: χ^2's for three theoretical models with a contact interaction.

physics the angular distributions are quite different than QCD. Other deviations from the Standard Model might change the absolute cross sections while maintaining a more QCD-like angular distribution. In this case this analysis is the most in-depth study so far of allowed deviations in jet cross sections due to new physics.

The limit derived from this analysis is also less than that obtained by the DØ collaboration when using the ratio of the dijet cross sections in the two rapidity bins ($0.0-0.5$ and $0.5-1.0$). However, it should be noted that at least part of the higher limit from DØ may be due to the dijet cross section for $0.5 < \eta < 1.0$ being somewhat larger than the standard model prediction, while the dijet cross section for $0.0 < \eta < 0.5$ is in better agreement with the SM prediction. The larger the ratio of the $0.5-1.0$ to $0.0-0.5$ rapidity intervals, the stronger the compositeness limit. The detailed information about the absolute values of the cross sections in each rapidity bin has not been used.

Table 2 shows how the χ^2's are affected by the contact interaction.

5. Predictions for run 2

The increase in the center-of-mass energy (from 1.8 to 1.96 TeV) and the increased luminosity expected in Run 2 (from approximately $100\,\mathrm{pb}^{-1}$ to $2-15\,\mathrm{fb}^{-1}$ will result in a dramatically larger kinematic range for measuring jet production, and thus for searching for possible new physics as well. Also, in Run 2, the new CDF endplug calorimeter will allow for an extension of measurements of the inclusive jet cross section to the forward rapidity region, similar to the analysis that has been performed by DØ in Run 1b.

The predictions for the Run 2 jet cross section as a function of transverse momentum (using the DØ rapidity intervals) are shown in figure 25. With $2\,\mathrm{fb}^{-1}$ (Run 2a) a measurement of the inclusive jet cross section out to a transverse energy of 600 GeV is possible, while with $15\,\mathrm{fb}^{-1}$ (Run 2b) the measurement can be extended to almost 650 GeV. The predictions are displayed on the left graph in the form of the curves for the 40 eigenvector basis sets of CTEQ6.1M. The central prediction and overall uncertainty band from the Master Equation (2.6) are shown separately on the right graph. Figure 26 shows the uncertainty range of the Run 2 cross section by displaying the ratios of the eigenvector basis sets to the central prediction. Near the kinematic limit, the PDF uncertainties allow the cross section predictions to be as much as a factor of 2 larger than those from the central fit (CTEQ6.1M), again due to eigenvector 15.

JHEP10(2003)046

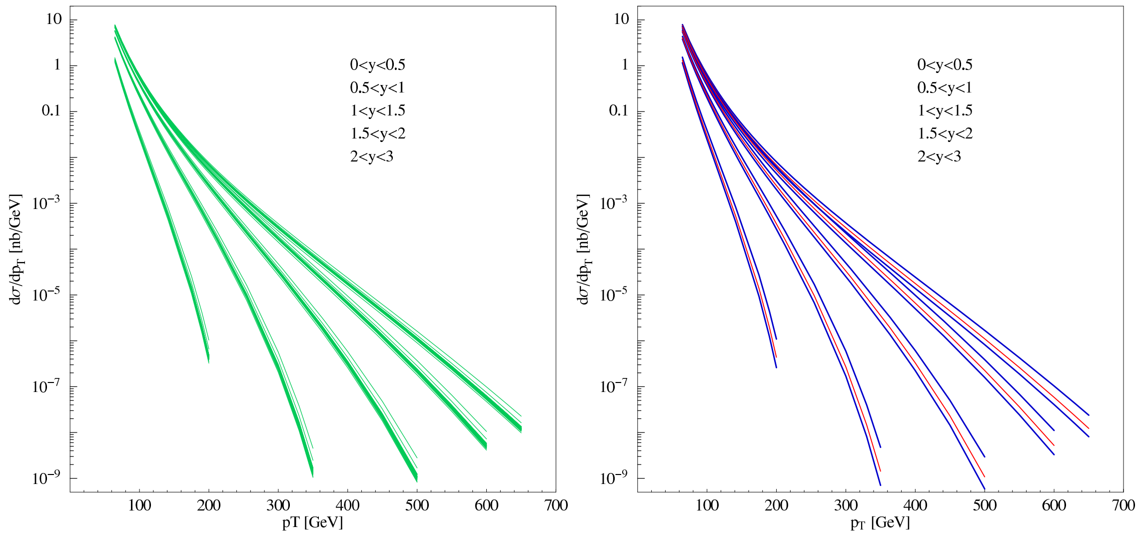

Figure 25: Run 2 cross section for the DØ rapidity bins. Left: The 40 eigenvector basis sets. Right: The central prediction and overall uncertainty band.

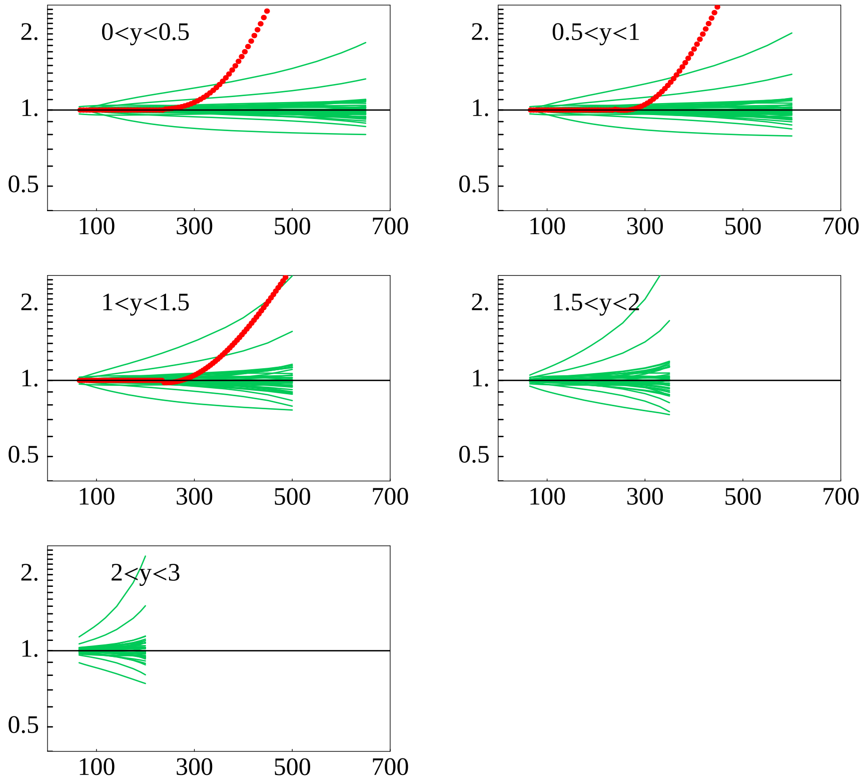

Figure 26: Uncertainty range of the Run 2 cross section for the DØ rapidity bins. The curves show the ratios of the 40 eigenvector basis sets to the central (CTEQ6.1M) prediction (ordinate) versus p_T in GeV (abscissa). The red curve is the effect of a compositeness scale of 2.4 TeV.

JHEP10(2003)046

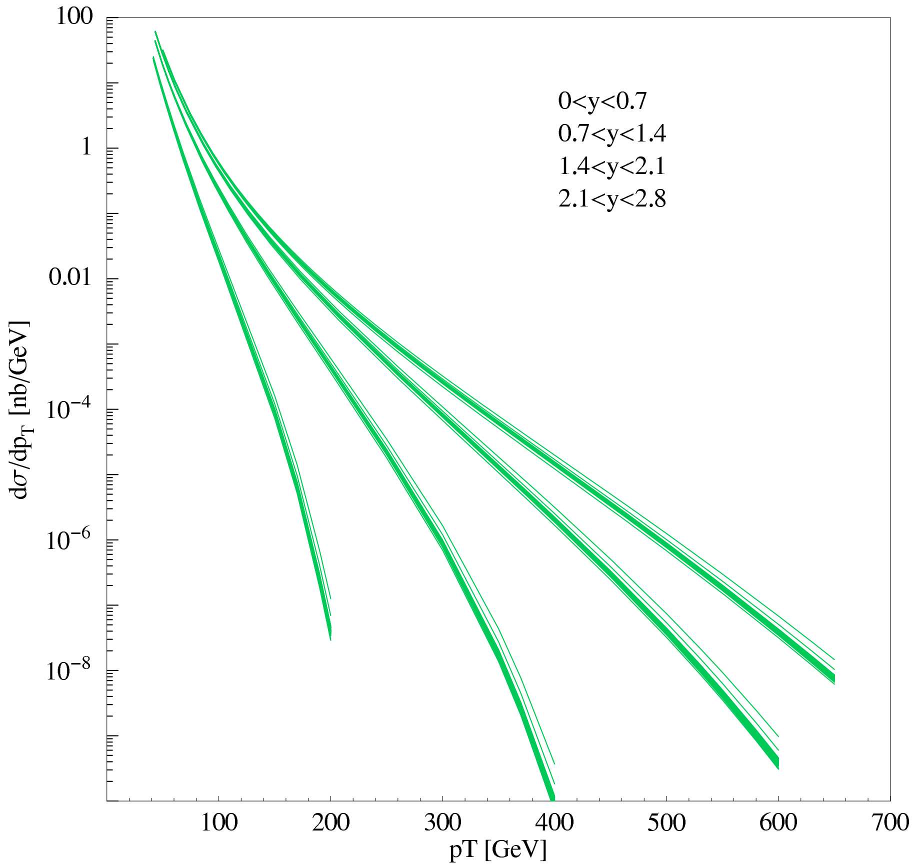

Figure 27: Run 2 cross section for the CDF rapidity bins.

Also shown in figure 26 is the effect of a compositeness scale of 2.4 TeV (with destructive interference). In the central rapidity region, the effect of the compositeness term is significantly greater than the PDF uncertainty in the highest transverse energy region.

For the Run 2 jet cross section predictions, we have used the Run 2 cone jet algorithm [24, 25] which uses 4-vector kinematics both to define the jet and to specify its transverse momentum and rapidity. This algorithm has been adopted by both the CDF and DØ collaborations for their inclusive Run 2 jet results. Note that this algorithm specifies the transverse momentum p_T rather than the transverse energy E_T.[8]

For completeness we also show the plots of the predictions of the inclusive jet cross section at the Tevatron Run 2 for the CDF choice of rapidity bins. Figure 27 shows the cross section as a function of p_T on a log scale. Figure 28 shows the uncertainty band in the form of the ratios of $d\sigma/dp_T$ for the 40 eigenvector basis sets compared to the central prediction.

While the statistical errors in Runs 2a and 2b will be small, so that the jet cross section can be extended to high p_T, the experience from Run 1b shows that the systematic errors will have a dominant effect on the significance of the results. However, we can anticipate

[8]The Run 2 jet algorithm results in cross section predictions for CDF that are $\sim 7\%$ lower than those using the Run 1 algorithm, over most of the kinematic range.

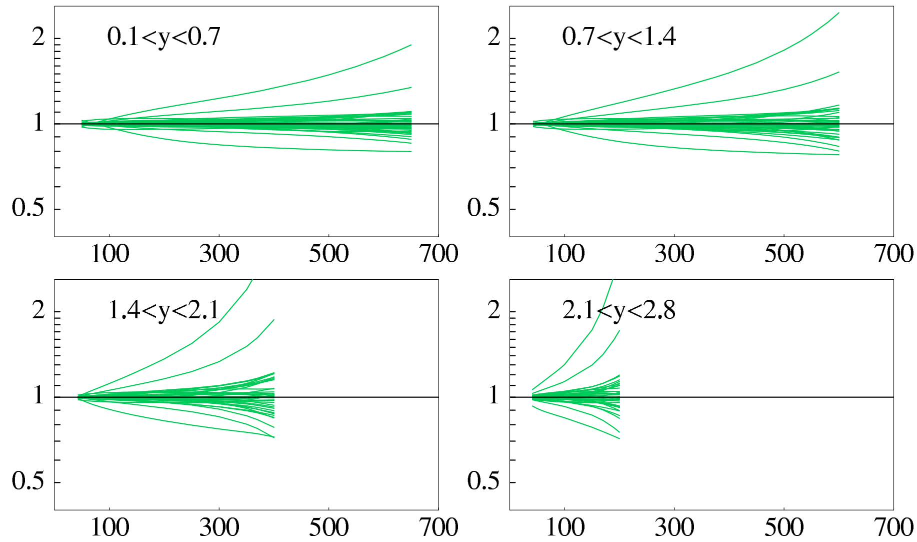

Figure 28: Uncertainty range of the Run 2 cross section for the CDF rapidity bins. The curves show the ratios of the 40 eigenvector basis sets compared to the central (CTEQ6.1M) prediction (ordinate) versus p_T in GeV (abscissa). The dotted curves are the effect of a compositeness scale of 2.4 TeV.

that the PDF's, especially the gluon distribution, will be pinned down at large x by these data. The experimental collaborations will make the data most useful for global analysis by providing detailed information on systematic uncertainties.

5.1 The ratio of run 2 to run 1b cross sections

As mentioned previously, the Run 2 center-of-mass energy $\sqrt{s}$ is higher than that from Run 1b. The increase in $\sqrt{s}$ will have only a modest impact on most physics cross sections. However, it will lead to a rather large increase in the production of high-E_T jets. It may be useful to examine the ratio of jet production (for the same E_T values) at the two different energies. In such a ratio, many of the theoretical (and experimental) errors will cancel. The Run 2 to Run 1b ratio for the 40 sets of PDF's is shown in figure 29 for the central rapidity range of CDF ($0.1 < |\eta| < 0.7$). As can be observed the ratio has a rather narrow theoretical error band from PDF uncertainties, i.e., smaller uncertainty than that of the absolute prediction of the inclusive cross section.[9] Indeed, in the region $E_T < 100$ GeV, the PDF errors are smaller than the Monte Carlo errors, which show up as small fluctuations in the predictions.

6. Inclusive jet production at the LHC

The increase of the center-of-mass energy to 14 TeV at the LHC will result in a dramatically larger accessible kinematic range. Inclusive jet cross sections will be measured out to

[9]Here we have multiplied our Run 2 predictions by a factor of 1.07 so that the comparison effectively uses the Run 1 algorithm for both the numerator and denominator. This is done for convenience since CDF will present its first results for the Run 2 jet cross section using the Run 1 algorithm.

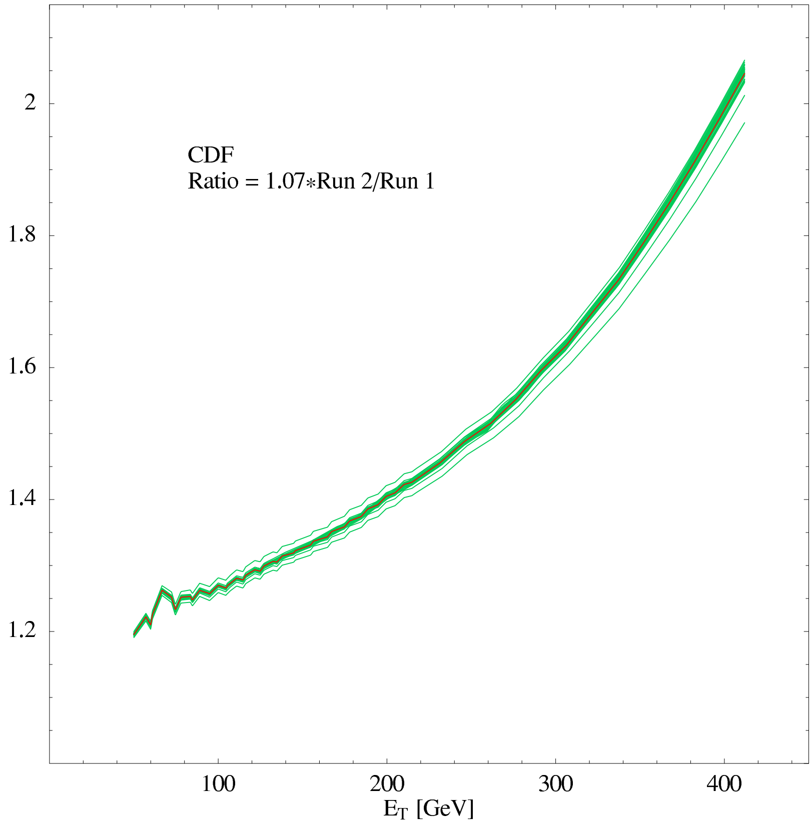

Figure 29: The ratio of the Run 2 to Run 1b cross sections for the central rapidity range ($0.1 < |\eta| < 0.7$). The Run 2 cross section has been multiplied by 1.07 to take into account the difference in the jet definitions for the Run 2 and Run 1b analyses. (The small fluctuations are due to the Monte Carlo errors in the calculations of the K-factors (NLO/LO) at each energy.)

transverse momentum of $5\,\mathrm{TeV}/c$ in the central rapidity range and out to $1.5\,\mathrm{TeV}/c$ in the forward rapidity region. Figure 30 shows the predictions with uncertainty ranges for three rapidity intervals. Figure 31 shows the ratios of the 40 eigenvector basis set predictions to the central prediction. The cross section uncertainties near the kinematic limit at the LHC due to PDF uncertainties are similar in magnitude to those obtained for Run 2 at the Tevatron. Again the extremes of the predictions are provided by the eigenvector basis sets $+15$ and -15, which correspond to extremes of the gluon distribution at large x.

The jet cross section at the LHC will be a discovery mode for new physics. Supersymmetry would produce an enhanced cross section from production of gluino and squark jets. Extra dimensions would show up from production of Kaluza-Klein modes of the graviton. So the pure QCD prediction will be compared to the data and any difference will be the first evidence for these or other new physics scenarios.

7. Conclusions

Jet production serves as a probe of the highest Q^2 scales accessible at a hadron-hadron collider. NLO QCD provides a good description of the inclusive jet cross sections measured

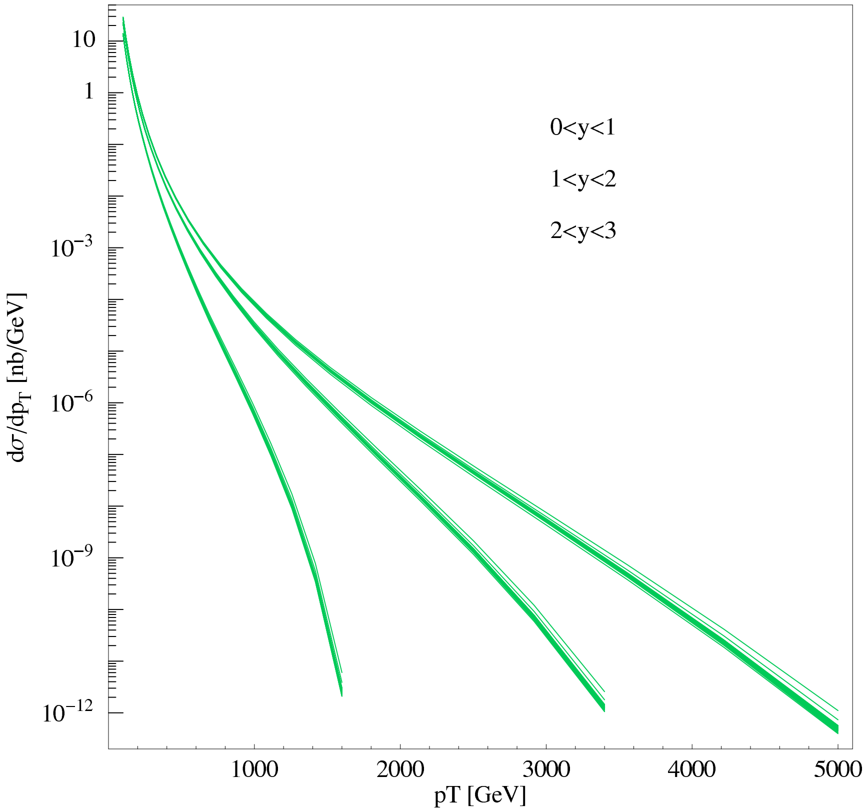

Figure 30: The inclusive jet cross section as a function of p_T for three rapidity bins at the LHC. The three rapidity ranges are $(0, 1)$, $(1, 2)$ and $(2, 3)$. Predictions of all 40 eigenvector basis sets are superimposed.

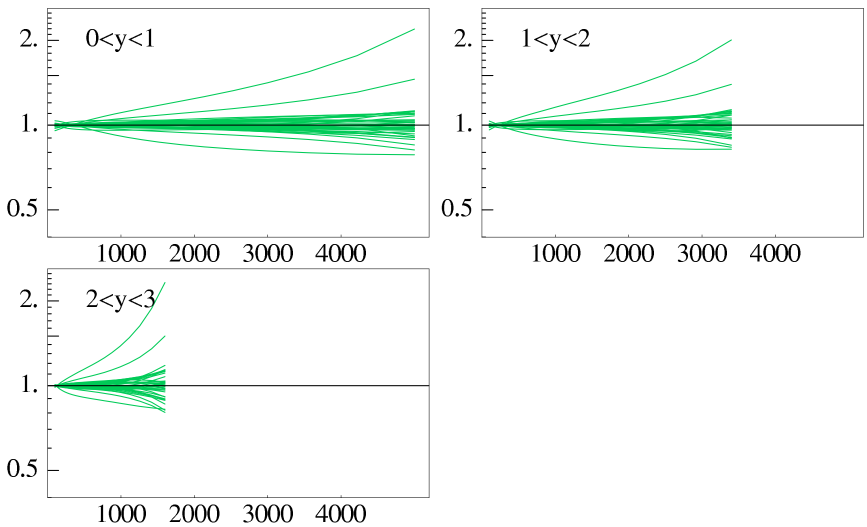

Figure 31: The uncertainty range of the inclusive jet cross section at the LHC. The curves are graphs of the ratios of the cross sections for the 40 eigenvector basis sets compared to the central (CTEQ6.1M) prediction (ordinate) versus p_T in GeV (ordinate).

by CDF and D0 in Run 1 at the Tevatron, provided that the gluon distribution is enhanced at high x relative to earlier estimates. Such an outcome is now the standard consequence of the global fitting procedure due to the statistical power of the Tevatron jet data, and the preference for a larger gluon in all rapidity regions. The uncertainties on the jet cross section predictions still remain large, however, primarily due to the remaining uncertainties on the gluon distribution. In most regions, they are comparable to the experimental systematic uncertainties.

The standard choice for the renormalization and factorization scales in evaluating inclusive jet cross sections has been $E_T^{jet}/2$. We have re-done the global fitting using alternate scales of E_T^{jet} and $2E_T^{jet}$ and have found changes to the gluon distribution that are well within the gluon PDF uncertainty band. The effects of the resummation of threshold logarithms has been calculated for the first time in the forward rapidity regions. As for the central rapidity region, the effects of the threshold logarithms are relatively small.

The increase of the center-of-mass energy to 1.96 TeV (from 1.8 TeV) and the expected greater integrated luminosity will lead to an expanded kinematic range for Run 2 at the Tevatron. Predictions have been made for both the inclusive jet cross section and its uncertainty for Run 2 and for the LHC. The PDF uncertainty on the jet cross section near the limits of the data reach in both cases are similar to those obtained for Run 1.

We have carried out an exercise where we have examined the allowed range for any new physics such as compositeness, given the current scale of uncertainties on the jet cross section predictions. Given the Run 1b jet data, any new physics must have a scale larger than 1.6 TeV. When data is available from Run 2 at the Tevatron, and from the LHC, analogous analyses will be used to identify evidence for new physics or place even stricter limits on new physics.

A. Reduction of scale dependence in the NLO calculation

One component of the theoretical uncertainty of any perturbative QCD calculation is the dependence of the answer on the various renormalization and factorization scales. One virtue of using NLO calculations in global determinations of parton distributions is the reduction in scale dependence usually exhibited by such calculations. The following brief discussion outlines the origin of the reduction of the scale dependence which is often observed with NLO calculations.

Consider a large transverse momentum process such as the single jet inclusive cross section involving only massless partons. Furthermore, in order to simplify the notation, suppose that the transverse momentum is sufficiently large that only the quark distributions need be considered. In the following, a sum over quark flavors is implied. Schematically, one can write the lowest order cross section as

$$E\frac{d^3\sigma}{dp^3} \equiv \sigma = a^2(\mu)\,\hat{\sigma}_B \otimes q(M) \otimes q(M) \tag{A.1}$$

where $a(\mu) = \alpha_s(\mu)/2\pi$ and the lowest order parton-parton scattering cross section is denoted by $\hat{\sigma}_B$. The renormalization and factorization scales are denoted by μ *and* M, respectively. In addition, various overall factors have been absorbed into the definition of $\hat{\sigma}_B$. The symbol $\otimes$ denotes a convolution defined as

$$f \otimes g = \int_x^1 \frac{dy}{y} f\left(\frac{x}{y}\right) g(y)\,. \tag{A.2}$$

When one calculates the $\mathcal{O}(\alpha_s^3)$ contributions to the inclusive cross section, the result can be written as

$$\begin{aligned}\sigma = {} & a^2(\mu)\,\hat{\sigma}_B \otimes q(M) \otimes q(M) + \\ & + 2a^3(\mu)\,b \ln\left(\frac{\mu}{p_T}\right) \hat{\sigma}_B \otimes q(M) \otimes q(M) + \\ & + 2a^3(\mu)\,\ln\left(\frac{p_T}{M}\right) P_{qq} \otimes \hat{\sigma}_B \otimes q(M) \otimes q(M) + \\ & + a^3(\mu)\,K \otimes q(M) \otimes q(M)\,. \end{aligned} \tag{A.3}$$

In writing eq. (A.3), specific logarithms associated with the running coupling and the scale dependence of the parton distributions have been explicitly displayed; the remaining higher order corrections have been collected in the function K in the last line of eq. (A.3). The μ dependence of the running coupling is obtained from

$$\mu \frac{\partial a(\mu)}{\partial \mu} = \beta(a(\mu)) \tag{A.4}$$

where $\beta = -ba^2(1 + ca)$ with

$$b = \frac{33 - 2f}{6} \qquad \text{and} \qquad c = \frac{153 - 19f}{2(33 - 2f)} \tag{A.5}$$

where f denotes the number of flavors. The scale dependence of the quark distributions is given by the nonsinglet DGLAP equation [26]

$$M \frac{\partial q(x, M)}{\partial M} = a(M) P_{qq} \otimes q(M)\,. \tag{A.6}$$

Now, the NLO expression in eq. (A.3) should be independent of the choice of scale up to corrections of order a^4. That this is indeed the case is easy to see. For convenience, set $M = \mu$ and calculate the logarithmic derivative of σ with respect to μ,

$$\begin{aligned}\mu \frac{\partial \sigma}{\partial \mu} & = 2a(\mu)[-ba^2(\mu)(1 + ca(\mu))]\hat{\sigma} \otimes q(\mu) \otimes q(\mu) + \\ & \quad + 2a^3(\mu) P_{qq} \otimes \hat{\sigma}_B \otimes q(\mu) \otimes q(\mu) + \\ & \quad + 2ba^3(\mu)\hat{\sigma}_B \otimes q(\mu) \otimes q(\mu) - \\ & \quad - 2a^3(\mu) P_{qq} \otimes \hat{\sigma}_B \otimes q(\mu) \otimes q(\mu) + \cdots \\ & = 0 + \mathcal{O}(a^4)\,. \end{aligned} \tag{A.7}$$

The derivative of the lowest order contribution yields the first two lines of eq. (A.7) while the derivative of the logarithms in the second and third lines of eq. (A.3) give rise to the last two lines in eq. (A.7) which cancel the μ dependent terms coming from the derivative of the lowest-order contribution. All other derivatives yield contributions which are of order a^4. Thus, the structure of the NLO calculation is such that the scale dependence is cancelled up to terms of order a^4.

The preceding discussion shows that one might expect a reduced scale dependence for the NLO calculation as opposed to that observed at lowest order. However, the uncancelled $\mathcal{O}(a^4)$ and higher terms will still result in some scale dependence. It is easy to understand the systematics of the remaining dependence for the case at hand. Consider, first, the lowest-order contribution shown in the first line of eq. (A.3). At large values of p_T as $x_T = 2p_T/\sqrt{s}$ tends to one, the scaling violations of the parton distributions lead to a decrease in the cross section as the scale increases. In addition, the running coupling decreases as the scale is increased. Therefore, at fixed p_T the lowest order contribution exhibits a monotonically decreasing behavior as the scale increases.

The second line of eq. (A.3) contains the contribution which partially offsets the scale dependence of the running coupling. Apart from the explicit logarithm, this contribution has much the same scale dependence as the lowest order contribution. Thus, for $\mu < p_T$ this term is negative, becoming positive as the scale is increased beyond p_T.

The third line contains the contribution which partially offsets the scale dependence originating in the parton distributions. In the kinematic region under consideration, the convolution of the splitting function P_{qq} with the lowest contribution gives a negative result, reflecting the fact that the parton distributions decrease with increasing scale at large values of x. Therefore, this contribution is negative when the factorization scale M is chosen to be less than p_T.

Finally, the remaining term on line four of eq. (A.3) has much the same scale dependence as the lowest order term, as the dominant scale dependence has already been exhibited in the terms on lines two and three. Thus, lines one and four give contributions which decrease monotonically with increasing scale while lines two and three start out negative, reach zero when the scales are equal to p_T, and are positive for larger scales. This behavior explains why the typical NLO scale dependence in the high-p_T region yields a result which first increases with increasing scale, reaches a maximum, and then decreases. For many high-p_T processes involving massless partons this maximum occurs near $\mu = M = p_T$.

Next, it is interesting to consider how the location of the maximum changes with p_T. Suppose that one goes to small values of x_T such that the typical values of x in the parton distributions is around .2 or smaller. In this region the scaling violations change sign, corresponding to a growth of the parton distributions with increasing scale, rather than a decrease. The contribution of the third line of eq. (A.7) then starts to compete against that of the second line. Accordingly, one must go to smaller values of μ/p_T before seeing the turnover and decrease with decreasing scale. That is, the position of the maximum at fixed p_T moves to lower values of μ/p_T as p_T decreases. The converse is also clearly true; as p_T increases the position of the maximum moves to higher values of μ/p_T. Thus, if one were to try to find an optimized scale at each value of p_T, the pattern that would emerge would be that the optimal scale would be a fractional multiple of p_T at lower p_T values and would increase to a multiple of p_T greater than one at higher p_T values. This pattern is precisely what is observed when using the single scale version of the Principle of Minimal Sensitivity [27].

JHEP10(2003)046

Another manifestation of this pattern concerns the behavior of the "K factor", defined as the ratio of the NLO to LO result. At a fixed value of p_T, the K factor will first rise and then fall with increasing scale. Suppose that at some p_T value the maximum occurs for the choice $\mu = M = p_T/2$. Then, as one goes to higher p_T values the previous discussion shows that the maximum will move to a higher scale choice such as p_T or even $2p_T$. At the larger p_T values it will then appear that the K factor may increase with increasing scale. This is just the behavior seen in the calculations done for the high rapidity D0 jet data.

In this appendix we have considered only one source of theoretical uncertainty, from the choice of the factorization and renormalization scales. A full understanding of the theoretical errors for a NLO jet cross section would require the calculation of the cross section to NNLO and is thus beyond the scope of the current paper (or, indeed, the field generally at this time.) It is also important to remember that other choices contribute to theoretical uncertainty, e.g., the functional paramatrization and the treatment of heavy quark mass corrections.

References

[1] CDF collaboration, F. Abe et al., *Inclusive jet cross section in $\bar{P}P$ collisions at $\sqrt{s} = 1.8$ TeV*, *Phys. Rev. Lett.* **77** (1996) 438 [hep-ex/9601008].

[2] J. Huston, E. Kovacs, S. Kuhlmann, H.L. Lai, J.F. Owens, D. Soper and W.K. Tung, *Large transverse momentum jet production and the gluon distribution inside the proton*, *Phys. Rev. Lett.* **77** (1996) 444 [hep-ph/9511386].

[3] H.L. Lai, J. Huston, S. Kuhlmann, F. Olness, J.F. Owens, D. Soper, W.K. Tung and H. Weerts, *Improved parton distributions from global analysis of recent deep inelastic scattering and inclusive jet data*, *Phys. Rev.* **D 55** (1997) 1280 [hep-ph/9606399].

[4] CTEQ collaboration, H.L. Lai et al., *Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions*, *Eur. Phys. J.* **C 12** (2000) 375 [hep-ph/9903282].

[5] CDF collaboration, T. Affolder et al., *Measurement of the inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV*, *Phys. Rev.* **D 64** (2001) 032001 [hep-ph/0102074], erratum *ibid.* **D65** (2002) 039903.

[6] DØ collaboration, B. Abbott et al., *The inclusive jet cross-section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV*, *Phys. Rev. Lett.* **82** (1999) 2451 [hep-ex/9807018]; *High-p_T jets in $\bar{p}p$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV*, *Phys. Rev.* **D 64** (2001) 032003 [hep-ex/0012046].

[7] DØ collaboration, B. Abbott et al., *Inclusive jet production in $p\bar{p}$ collisions*, *Phys. Rev. Lett.* **86** (2001) 1707 [hep-ex/0011036].

[8] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky and W.K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, *J. High Energy Phys.* **07** (2002) 012 [hep-ph/0201195].

[9] M. Klasen and G. Kramer, *Large transverse momentum jet production and DIS distributions of the proton*, *Phys. Lett.* **B 386** (1996) 384 [hep-ph/9605210];
P. Anandam and D.E. Soper, *Factorization scheme dependence of the NLO inclusive jet cross section*, *Phys. Rev.* **D 61** (2000) 094003 [hep-ph/9912327].

[10] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H.L. Lai and W.K. Tung, *Uncertainties of predictions from parton distribution functions, II. The hessian method*, *Phys. Rev.* **D 65** (2002) 014013 [hep-ph/0101032].

[11] J. Pumplin, D.R. Stump and W.K. Tung, *Multivariate fitting and the error matrix in global analysis of data*, *Phys. Rev.* **D 65** (2002) 014011 [hep-ph/0008191].

[12] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, J. Kalk, H.L. Lai and W.K. Tung, *Uncertainties of predictions from parton distribution functions, I. The lagrange multiplier method*, *Phys. Rev.* **D 65** (2002) 014012 [hep-ph/0101051].

[13] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, *Uncertainties of predictions from parton distributions, I. Experimental errors*, *Eur. Phys. J.* **C 28** (2003) 455 [hep-ph/0211080].

[14] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, *MRST2001: partons and α_s from precise deep inelastic scattering and Tevatron jet data*, *Eur. Phys. J.* **C 23** (2002) 73 [hep-ph/0110215].

[15] Z. Sullivan and P.M. Nadolsky, *Heavy-quark parton distribution functions and their uncertainties*, *eConf* **C010630** (2001) P511 [hep-ph/0111358];
P.M. Nadolsky and Z. Sullivan, *PDF uncertainties in W H production at Tevatron*, *eConf* **C010630** (2001) P510 [hep-ph/0110378].

[16] G.C. Blazey and B.L. Flaugher, *Inclusive jet and dijet production at the Tevatron*, *Ann. Rev. Nucl. Part. Sci.* **49** (1999) 633 [hep-ex/9903058].

[17] N. Kidonakis, G. Oderda and G. Sterman, *Threshold resummation for dijet cross sections*, *Nucl. Phys.* **B 525** (1998) 299 [hep-ph/9801268]; *Evolution of color exchange in QCD hard scattering*, *Nucl. Phys.* **B 531** (1998) 365 [hep-ph/9803241].

[18] N. Kidonakis and J.F. Owens, *Effects of higher-order threshold corrections in high-E_T jet production*, *Phys. Rev.* **D 63** (2001) 054019 [hep-ph/0007268].

[19] D0 collaboration, B. Abbott et al., *Measurement of dijet angular distributions and search for quark compositeness*, *Phys. Rev. Lett.* **80** (1998) 666 [hep-ex/9707016];
CDF collaboration, F. Abe et al., *Measurement of dijet angular distributions at CDF*, *Phys. Rev. Lett.* **77** (1996) 5336 [hep-ex/9609011], erratum *ibid.* **78** (1997) 4307.

[20] DØ collaboration, B. Abbott et al., *The inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV*, *Phys. Rev. Lett.* **82** (1999) 2451 [hep-ex/9807018].

[21] DØ collaboration, B. Abbott et al., *The dijet mass spectrum and a search for quark compositeness in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV*, *Phys. Rev. Lett.* **82** (1999) 2457 [hep-ex/9807014].

[22] E. Eichten, K.D. Lane and M.E. Peskin, *New tests for quark and lepton substructure*, *Phys. Rev. Lett.* **50** (1983) 811.

[23] T. Sjostrand, *Pythia 5.6 and Jetset 7.3*, CERN-TH 6488 (1992) and references therein.

[24] See, for example, a detailed discussion in S. Ellis, J. Huston and M. Tonnesmann, *Bulding better cone jet algorithms*, in preparation.

[25] G.C. Blazey et al., *Run II jet physics*, in *Proceedings of the run II QCD and weak boson physics workshop* [hep-ex/0005012].

[26] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, *Nucl. Phys.* **B 126** (1977) 298.

[27] P.M. Stevenson, *Optimized perturbation theory*, *Phys. Rev.* **D 23** (1981) 2916.