Correlated theoretical uncertainties for the one-jet inclusive cross section

Fredrick I. Olness[1,*] and Davison E. Soper[2,†]

[1]*Department of Physics, Southern Methodist University, Dallas, Texas 75275-0175, USA*
[2]*Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403-5203, USA*

We discuss the correlated systematic theoretical uncertainties that may be ascribed to the next-to-leading order QCD theory used to predict the one-jet inclusive cross section in hadron collisions. We estimate the magnitude of these errors as functions of the jet transverse momentum and rapidity. The total theoretical error is decomposed into a set of functions of transverse momentum and rapidity that give a model for statistically independent contributions to the error. This representation can be used to include the systematic theoretical errors in fits to the experimental data.

I. INTRODUCTION

Predictions of the standard model are typically made with the aid of next-to-leading order (NLO) perturbative calculations (or sometimes with NNLO calculations). Evidently, these predictions are not exactly equal to what one should measure if the standard model is correct. If we have an NLO calculation, we leave out NNLO and N³LO contributions, etc. We also leave out contributions that are suppressed by a power of the large momentum scale of the problem. Of course, we do not know exactly how big these contributions are: if we could calculate them, we would include them in the prediction. Nevertheless, we can estimate the size of the corrections. They then constitute "theory errors" in the prediction, which are quite similar to experimental systematic errors in the measurement.

In this paper we distinguish between errors associated with higher order contributions and power suppressed contributions to the cross section, which we call theory errors, and errors associated with our imperfect knowledge of the parton distribution functions needed for the prediction. Estimated theory errors are needed in two contexts. First, if an experiment does not agree with the theoretical prediction within the experimental statistical and systematic errors, then we need to see if there is agreement within the combined experimental and theory errors and the errors from the parton distributions used in the prediction. In the case that the disagreement is outside of the combined errors, then we have a signal for new physics.

The second context in which we need estimated theory errors is in the determination of parton distribution functions from experimental measurements. The theory errors give a contribution to the errors that we associate with the parton distribution functions that emerge from a fit to the data. Evidently, if we do not include theory errors, the resulting errors in the parton distribution functions will be too small. Additionally, if for one kind of process the theory errors are large while for another kind of process the theory errors are small, then we will give the large-error process too much weight in the fit.

In this paper, we provide an estimate of the theory error for the one-jet inclusive cross section $d^2\sigma/dP_T dy$ in hadron-hadron collisions, where P_T is the transverse momentum or "transverse energy" of the jet, and y is the rapidity of the jet. There is good data for this process from the CDF and D0 experiments at Fermilab, including careful estimates of the experimental systematic errors. Estimates of the theory errors are needed to accompany the estimates of the experimental systematic errors.

We warn that there is no unique method to estimate theory errors. Thus our task is to provide a method that is defensible if not necessarily optimal. We seek to provide an estimate in a form that includes the correlations from one $\{P_T, y\}$ point to another.

II. GENERAL SETUP

We treat theory errors in a fashion that is similar to that used for correlated systematic errors in the experimental results. We use next-to-leading order quantum chromodynamics (QCD) theory to make predictions for the one-jet inclusive cross section[1]

$$\frac{d\sigma}{dP_T dy} = \int dx_1 \int dx_2 f_{a/A}(x_1, \mu) f_{b/B}(x_2, \mu) \frac{d\hat{\sigma}_{ab\to\text{jet}}}{dP_T dy}.$$

In the calculation, one uses Monte Carlo integration so that there is a random statistical error for each point $\{P_T, y\}$. We do not include these statistical errors in the analysis here since they are typically quite small (say 2%) and one can reduce them by running the program for a longer time. If we wished to include the errors from fluctuations in the Monte Carlo integrations, that task would be straightfor-

_____
*olness@smu.edu
†soper@uoregon.edu

_____
[1]Specifically, we use the program of Ref. [1], although there are other programs that can give the same results. The code is available at http://zebu.uoregon.edu/~soper/EKSJets/jet.html

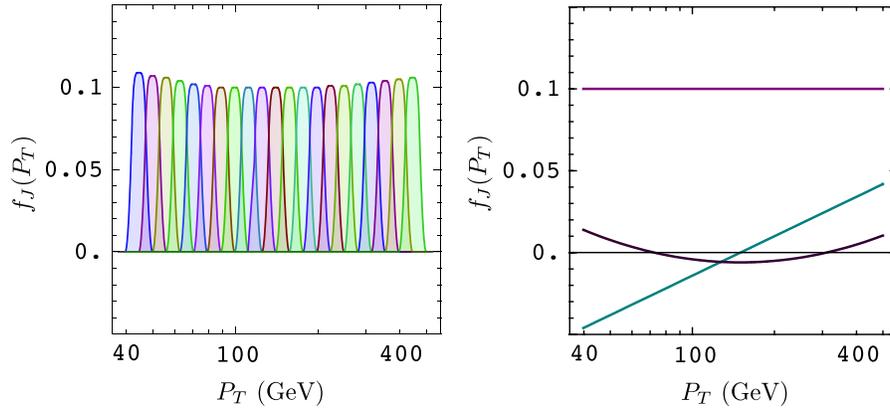FIG. 1 (color online).   illustration of (a) uncorrelated and (b) correlated theoretical errors. In (a), the total error is about 10% for all P_T, but the error at any P_T is not correlated with the error at nearby points. In (b), there are just three functions $f_J(P_T)$ giving, again, about a 10% total error at any one P_T. Because the $f_J(P_T)$ are smooth functions, the theoretical error at a given P_T will be smoothly related to the error at other P_T values.

ward because the statistical nature of these fluctuations is known.

We will start our investigation by studying jet production corresponding to the Tevatron Run 2, with $\sqrt{s} = 1960$ GeV, as a function of P_T and y. We will display the results for $y = \{0, 1, 2\}$ as functions of P_T; we also present formulas for the P_T and y dependence, from which estimated errors for the specific kinematic ranges used by CDF and D0 can be inferred.

We need estimated errors that can be used in a statistical analysis. However, we do not have at hand a statistical ensemble of worlds in which terms beyond those included in the NLO theory vary. Thus we make estimates that we hope are reasonable but that can and should be subject to debate.

We formulate the treatment of theory errors as follows. We let

$$\frac{d\sigma}{dP_T dy} = \left[\frac{d\sigma}{dP_T dy}\right]_{\text{NLO}} \left\{1 + \sum_J \lambda_J f_J(P_T, y)\right\}. \quad (1)$$

Here the functions $f_J(P_T, y)$ are definite functions, while the λ_J are unknown parameters. Thus $\lambda_J f_J(P_T, y)$ represents an unknown theoretical contribution that might modify the NLO theory. We treat the λ_J as Gaussian random variables with variance 1. That is, the size of the uncertainty with label J is represented by how big $f_J(P_T, y)$ is. If one thinks of this as representing an imaginary ensemble of worlds in which theory calculations come out differently, then these worlds all have the same f_J but the λ_J vary.

We will propose to use just a few functions f_J. We offer the following defense of this strategy. Consider a simplified case of a cross section that is a function of just one variable, P_T. If we were to believe that the uncertainty in the prediction of this cross section is of order, say, 10%, but we have no idea of what the shape of the true cross section is within a 10% band about the prediction, then we would choose many functions $f_J(P_T)$, each of size 0.10, but with

each being nonzero only in a very tiny range of P_T. This approach is illustrated in Fig. 1(a); such a view seems to us unreasonable.

Experience with various perturbative and nonperturbative contributions teaches that they are smooth functions of the relevant variables, P_T in this case. This arguably more reasonable scenario is illustrated in Fig. 1(b). As illustrated by the three curves,[2] one contribution beyond NLO could be flat, amounting to a constant "K factor," another might be a smoothly increasing function of P_T, while yet another might be positive at high and low P_T and negative in between. However, we judge it unlikely that a currently uncalculated contribution would have multiple maxima between low and high P_T.

Thus we seek a few functions $f_J(P_T, y)$ that have some dependence on $\{P_T, y\}$ and represent, as best we can determine, our understanding of the character of uncalculated contributions. In the following sections, we analyze several sources of theory errors and associate them with functions $f_J(P_T, y)$.

III. PERTURBATIVE UNCERTAINTY

The main source of uncertainty at large jet transverse momentum, at least in our estimation, is the fact that we have calculated only at NLO, leaving contributions from higher orders of perturbation uncalculated. We estimate this uncertainty using the dependence of the computed cross section on the renormalization and factorization scales. We present this estimate in this section. In the

---

[2]Specifically, in this figure we use the functions $f_1(P_T) = 0.1$, $f_2(P_T) = 0.08 \log(P_T/M)$, and $f_3(P_T) = 0.06\{[\log(P_T/M)]^2 - 0.1\}$ where $M = 150$ GeV. These curves are for illustrative purposes only, and the $f_J(P_T)$ functions differ from the set $f_J(P_T, y)$ we will use to parametrize the correlated systematic uncertainties.

following section, we check this estimate using an independent method involving threshold effects.

A. Error estimate from scale dependence

The first ingredient in our estimation of theory errors is based on the traditional method in which one evaluates the dependence of the computed NLO cross section on two scales: the renormalization scale μ_R and the factorization scale μ_F. One often makes a standard choice for these scales: $\mu_R = \mu_F = P_T/2$. We will take this choice as our central value and define

$$x_1 = \log_2\left(\frac{\mu_R}{P_T/2}\right), \qquad x_2 = \log_2\left(\frac{\mu_F}{P_T/2}\right). \quad (2)$$

We compute the cross section near $x_1 = x_2 = 0$, that is near the scale choice $\mu_R = \mu_F = P_T/2$. Then $\{x_1, x_2\}$ measures (logarithmically) the distance from this central value. We then fit the cross section to a quadratic polynomial $P(\vec{x})$ in $\vec{x}$-space,

$$\left[\frac{d\sigma(x_1, x_2)}{dP_T}\right]_{\text{NLO}} \approx \left[\frac{d\sigma(0, 0)}{dP_T}\right]_{\text{NLO}}[1 + P(\vec{x})], \quad (3)$$

where

$$P(\vec{x}) = \sum_J x_J A_J + \sum_{J,K} x_J M_{JK} x_K, \quad (4)$$

with $\vec{x} = (x_1, x_2)$ and $J, K = \{1, 2\}$.

We know that if we had an NNLO calculation, the dependence of the cross section on $\vec{x}$ would be canceled to order α_s^2. Thus the coefficients A_J and M_{JK} carry information about the perturbative coefficients beyond NLO. For this reason, we use the coefficients A_J and M_{JK} to provide an estimate of the error induced by truncating the perturbative expansion at one-loop order. We define a simple recipe for this purpose. We define an estimated error[3] $\mathcal{E}_{\text{scale}}$ as the root-mean-square average of $P(\vec{x})$ over a circle with a certain radius $|\vec{x}|$,

$$\mathcal{E}_{\text{scale}}^2 = \frac{1}{2\pi} \int_0^{2\pi} d\theta P(|\vec{x}|\cos\theta, |\vec{x}|\sin\theta)^2. \quad (5)$$

We need to select a value of $|\vec{x}|$, and we make the choice

$$|\vec{x}| = 2. \quad (6)$$

In the most common method of estimating errors from scale variation, we would vary $(2\mu_R/P_T, 2\mu_F/P_T)$ between $(1, 1)$ and $(2, 2)$ and between $(1, 1)$ and $(1/2, 1/2)$. This amounts to changing $\vec{x}$ from 0 to a vector of length $|\vec{x}| = \sqrt{2}$ in a particular direction that corresponds to something close to the direction of strongest variation. The choice $|\vec{x}| = 2$ is somewhat larger than this standard choice. For instance, $|\vec{x}| = 2$ in the direction $\vec{x} \propto (1, 1)$

---

corresponds to

$$\left(\frac{\mu_R}{P_T/2}, \frac{\mu_F}{P_T/2}\right) = (2^{\sqrt{2}}, 2^{\sqrt{2}}) \approx (2.7, 2.7). \quad (7)$$

We average over the directions of $\vec{x}$ instead of taking a particular direction. For this reason, the value of Eq. (6) gives results that are similar to the method that is often used. While varying the μ-scales along the (1, 1) direction will often work, our averaging technique provides a general method that seems sensible even when the one of the directions of slowest variation happens to align with the (1, 1) direction.

A straightforward calculation shows that, with the definition (5),

$$\mathcal{E}_{\text{scale}}^2 = \frac{|\vec{x}|^2}{2}\vec{A}^2 + \frac{|\vec{x}|^4}{8}[(\text{Tr}M)^2 + 2\,\text{Tr}M^2]. \quad (8)$$

We determine the coefficients A_J and M_{JK} by calculating the one-jet inclusive cross section for a given value of P_T and rapidity. We use nine points in $\vec{x}$-space, obtained by setting each $\{\mu_R, \mu_F\}$ scale to $\{\frac{1}{4}P_T, \frac{1}{2}P_T, P_T\}$ and fit the results to the form given in Eqs. (3) and (4).

B. Contour plots

We illustrate this procedure for estimating the theoretical error from this source in Fig. 2, where we display contour plots of $1 + P(\vec{x})$ corresponding to the jet cross section at the Tevatron with $P_T = 100$ GeV for $y = 0$ and for $y = 2$. For both values of y, we find a saddle point in the vicinity of $\{x_1, x_2\} = \{0, 0\}$ which corresponds to $\{\mu_R, \mu_F\} = \{P_T/2, P_T/2\}$. This location of the saddle point is a general feature that holds throughout much of the kinematic range; it motivates the choice $\{\mu_R, \mu_F\} = \{P_T/2, P_T/2\}$ as our central values.

The estimated scale dependence error, $\mathcal{E}_{\text{scale}}$, is then obtained by averaging the deviation of the cross section at a given radius in $\vec{x}$-space. As discussed above, we choose a radius of $|\vec{x}| = 2$, as indicated by the circle in Fig. 2. The slope of the $\{x_1, x_2\}$ surface is steeper for the $y = 2$ case as compared with the $y = 0$ case. Consequently, we find a larger $\mathcal{E}_{\text{scale}}$ for $y = 2$ ($\sim 18\%$) as compared to $y = 0$ ($\sim 9\%$).

C. Comment on the range of scale choices

In the above analysis, we estimate the theoretical uncertainty by varying the μ scales by a factor about a central value. This is a conventional choice, but is it reasonable? To examine this question, one can look at cases in which NNLO calculations exist. Here, we choose one typical case as an example. In Fig. 3, we show the NNLO cross section for Higgs production at the Large Hadron Collider (LHC) as a function of the P_T^{veto} parameter as calculated by Ref. [2]. Here, the renormalization and factorization scales are varied by a factor of 2, $\{\mu_R, \mu_F\} \in [M_h/2, 2M_h]$.

---

[3]We shall use $\mathcal{E}_{\text{scale}}$ to denote the theoretical systematic error due to scale dependence only, and $\mathcal{E}$ (no subscript) to denote the total theoretical systematic error.
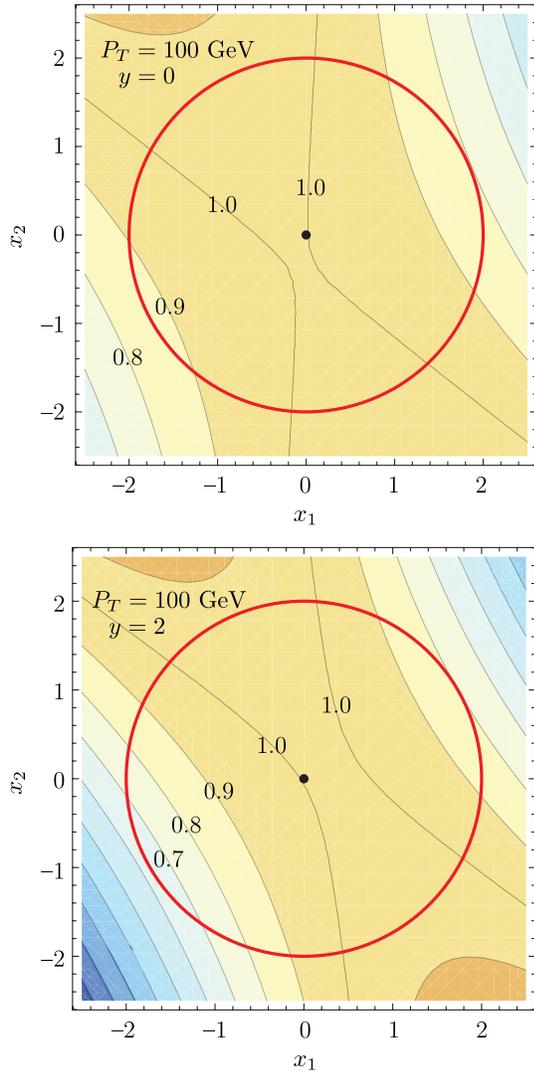
FIG. 2 (color online). Contour plot of the jet cross section in the $\{x_1, x_2\}$ plane for the Tevatron ($\sqrt{s} = 1960$ GeV) with $P_T = 100$ GeV and (a) central rapidity $y = 0$ and (b) forward rapidity $y = 2$. We plot the ratio of the cross section compared to the central value at $\{x_1, x_2\} = \{0, 0\}$. Contour lines are drawn at intervals of 0.10. The (red) circle is at radius $|x| = 2$.

Consider, for example, P_T^{veto} near 80 GeV. To simplify our argument, let us suppose that the exact QCD result is known and that it lies in the middle of the NNLO error band. We then ask whether the estimated NLO error band was reasonable, now that we know the exact answer. To do a real statistical analysis, we should have at hand many NLO calculations of separate and independent quantities, each with its error estimate. For each such quantity, a NNLO calculation that we can regard as nearly "exact" should be available. We would then plot the distribution of the differences between the NLO central value and the true answer in units of the NLO 1σ error estimate. If the error estimates are reliable, this distribution should be a Gaussian distribution with width 1. We cannot do that with just one datum. However, we can say that if the
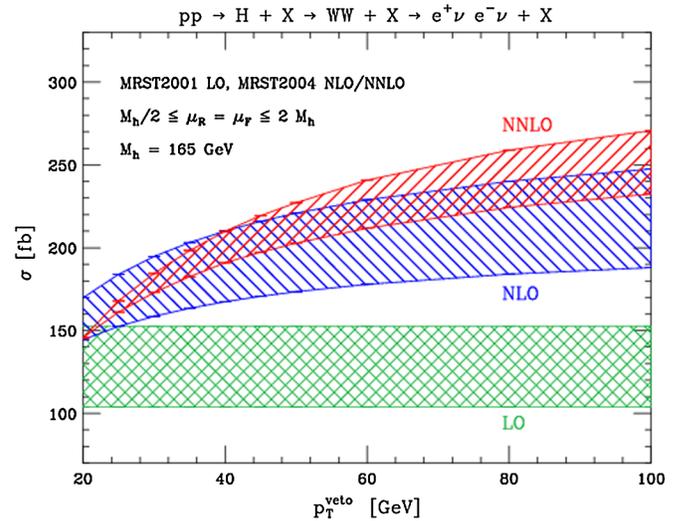

FIG. 3 (color online). The cross section for Higgs production at the LHC for LO, NLO, and NNLO calculations as taken from Ref. [2]. The computed cross section vetos jets ($P_T^{\text{jet}} > P_T^{\text{veto}}$) in the central region $|\eta| < 2.5$.

NLO estimate is reasonable then the central NNLO value in the one case that we have should be roughly 1σ away from the NLO central value. If it is 3σ away, then it seems likely that the NLO error was underestimated. If it is 0.1σ away, then seems likely that the NLO error was overestimated. In the case at hand, the difference is about 1σ, so we have some evidence that the error was correctly estimated.

D. Scale dependence total uncertainty

Implementing the procedure outlined above, we find the theoretical systematic error estimated from scale dependence, $\mathcal{E}_{\text{scale}}$; this is displayed in Fig. 4. The (blue) points are $\mathcal{E}_{\text{scale}}$ computed as described above from the NLO cross section [1] and the (red) curve is a smooth fit to these points.

We see that $\mathcal{E}_{\text{scale}}(P_T, y)$, is a slowly rising function of P_T. For the rapidity $y = 0$ at the Tevatron ($\sqrt{s} = 1960$ GeV), we find that $\mathcal{E}_{\text{scale}}(P_T, y)$ varies from 9% to 11%. For $y = 1$, the uncertainty ranges from 9% to 20%, and for $y = 2$ the uncertainty increases even more, ranging from 12% to 25% over a more limited P_T range.

E. Scale dependence correlated uncertainty

As described in Sec. II, we decompose the total scale dependence uncertainty, $\mathcal{E}_{\text{scale}}$, into a (small) number of functions $f_J(P_T, y)$ which then combine to form the total uncertainty $\mathcal{E}_{\text{scale}}$.

Since the $f_J(P_T, y)$ functions represent independent sources of uncertainty, $\mathcal{E}_{\text{scale}}$ is the quadrature sum

$$\mathcal{E}_{\text{scale}}(P_T, y) \equiv \sqrt{\sum f_J(P_T, y)^2}. \tag{9}$$

We chose a set of functions $f_J(P_T, y)$ that satisfies Eq. (9).
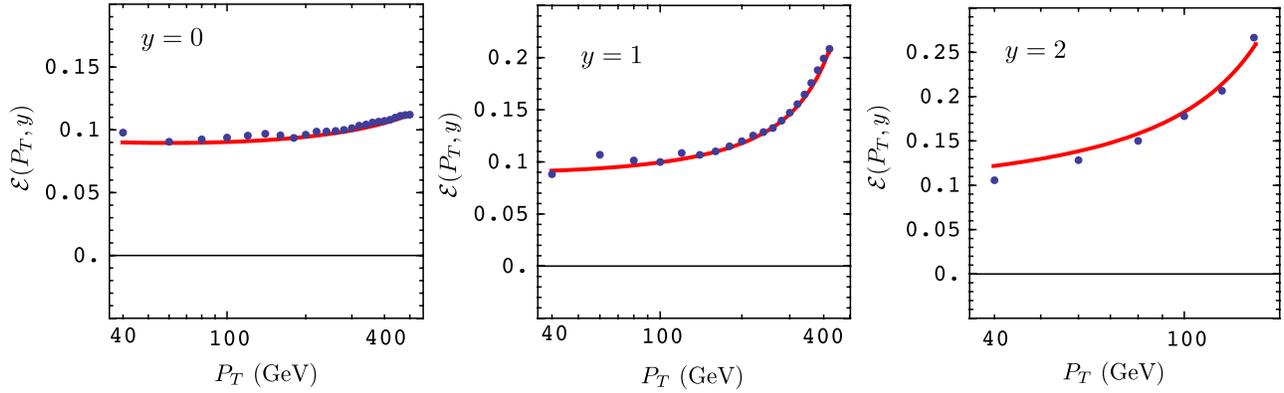
FIG. 4 (color online).   The estimate of the uncertainty $\mathcal{E}(P_T, y) = \mathcal{E}_{\text{scale}}$ due to the scale variation as given in Eq. (8) for the Tevatron ($\sqrt{s} = 1960$ GeV) with $y = \{0, 1, 2\}$. The calculation from the jet code is represented by the (blue) points, and the fit based on Eq. (9) is shown with the solid (red) curve.
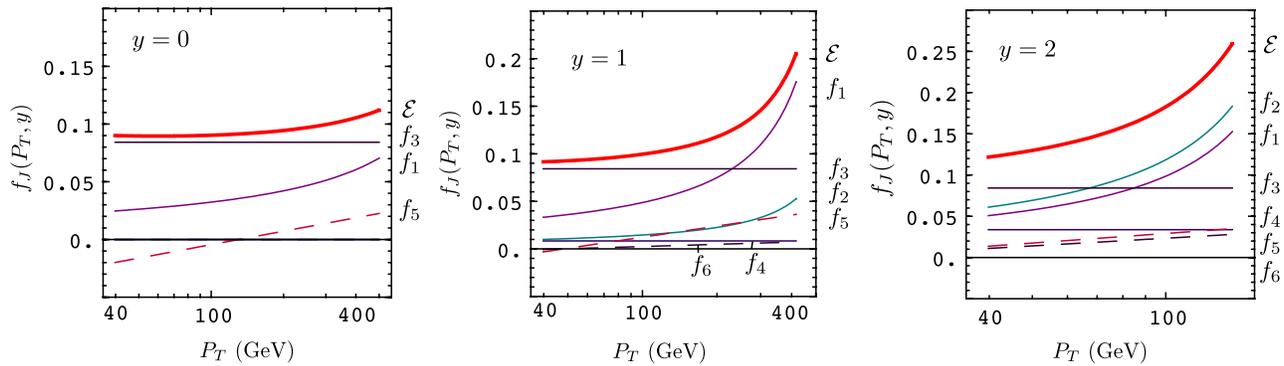

FIG. 5 (color online).   The estimate of the uncertainty $\mathcal{E}_{\text{scale}}$ due to the scale variation as given in Eq. (8) for the Tevatron ($\sqrt{s} = 1960$ GeV) with $y = \{0, 1, 2\}$. The combined uncertainty $\mathcal{E}_{\text{scale}}$ is shown as the upper thick (red) curve, and the individual functions $f_J(P_T, y)$ are indicated below.

We take the $f_J(P_T, y)$ to depend on y and on the ratio of P_T to the quantity[4]

$$M(y) = \sqrt{s}\, e^{-y}. \qquad (10)$$

For the set of $f_J(P_T, y)$ functions we choose

$$
\begin{aligned}
f_1(P_T, y) &= \frac{9.62 \times 10^{-2}}{\log(M(y)/P_T)}, \\
f_2(P_T, y) &= \frac{2.89 \times 10^{-2} y^2}{\log(M(y)/P_T)}, \\
f_3(P_T, y) &= 8.42 \times 10^{-2}, \\
f_4(P_T, y) &= 0.842 \times 10^{-2} y^2, \\
f_5(P_T, y) &= 1.68 \times 10^{-2} \log\!\left(\frac{15 P_T}{M(y)}\right), \\
f_6(P_T, y) &= 0.336 \times 10^{-2} y^2 \log\!\left(\frac{15 P_T}{M(y)}\right).
\end{aligned}
\qquad (11)
$$

These functions are illustrated in Fig. 5. The first two terms are singular as $P_T \to M(y)$. The first controls the singular behavior near $y = 0$ while the second modifies the singular behavior for large y. The remaining terms constitute a polynomial in $\log(P_T)$ and y^2. Thus, we parametrize the y-dependence with the set of functions $\{1, y^2\}$, and the P_T-dependence with the set of functions $\{1/L, 1, L\}$ where L represents a logarithmic function of P_T. We believe that the parametrization in terms of these $2 \times 3 = 6$ functions is sufficient to reasonably describe the theoretical uncertainties.

Note that the coefficients of f_3 and f_4 are in the ratio 10:1 and the coefficients of f_5 and f_6 are in the ratio 5:1. While we could find an excellent fit without f_4 and f_6, we retain these terms to provide flexibility when one tries to fit the λ_J coefficients to actual data.

We can perform a similar exercise for the LHC as well; these results will be compiled and presented in Sec. VII.

---

[4]We scale P_T by $M(y)$ to make the argument of the logarithms dimensionless. This quantity provides a simple scaling, and roughly corresponds to scaling by the maximum P_T, $P_T^{\text{max}} \sim \sqrt{s}/(2\cosh(y))$, for large y.

IV. SUMMATION OF THRESHOLD LOGS

For parton-parton scattering near the threshold for the production of a jet with a given P_T, there is restricted phase space for real gluon emission. Thus, there is an incomplete
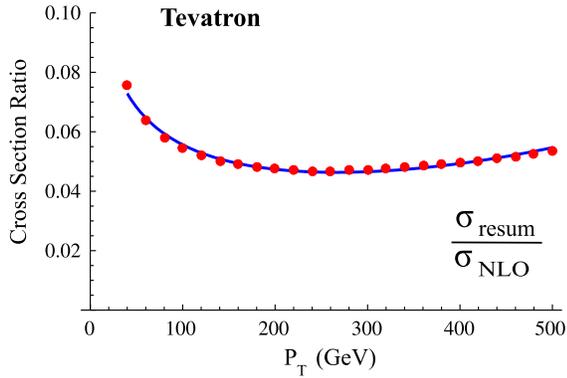
FIG. 6 (color online). The ratio of the two-loop threshold resummation contributions for jet production compared to the total NLO cross section $\sigma_{\text{resum}}/\sigma_{\text{NLO}}$ at the Tevatron ($\sqrt{s} = 1960$ GeV) vs P_T in GeV. We have set the scales to $\mu_F = \mu_R = P_T/2$, and used $y = 0$. The points are computed using the implementation of the 2-loop threshold resummation by Kidonakis and Owens [9].

cancellation of infrared divergences between real and virtual graphs, resulting in large logarithms L inside the integration over parton momentum fractions. At nth order in α_s these logarithms enter the cross section in the general form $\alpha_s^n L^{2n}$. The leading logarithms can be summed to all orders in α_s. We make use of the numerical results from Ref. [3], which has been implemented in the FASTNLO program [4].

Figure 6 displays the size of the threshold correction for Tevatron jet measurements at $y = 0$. The curve is presented for the scale choice $\mu = P_T/2$; we note that for this scale choice, the threshold correction is generally smaller than with other scale choices.[5]

We find the threshold corrections in this kinematic regime to be less than those discussed in the previous section (Sec. III) and shown in Fig. 5. As the threshold corrections also arise from uncomputed higher-order terms, these corrections are, in a sense, already accommodated by the larger uncertainty that we estimated from scale variation in Eq. (11). Indeed, the functions f_J for $J = 1$ and $J = 2$ contain singularities for $P_T \rightarrow M(y)$ that are meant to incorporate the threshold singularities. For this reason, we will not add a separate $f_J(P_T, y)$ function in the expression for the total uncertainty $\mathcal{E}$ to represent the effects of threshold logarithms.

---

[5]We do not present curves for $y = 1$ and $y = 2$ because these curves show a rise of the correction as P_T decreases from 200 GeV, even though decreasing P_T puts us farther from the threshold. This rise is more pronounced for large y than we see for $y = 0$ in Fig. 6. We suspect that this behavior is an artifact of kinematic choices in the algorithm for summing threshold logarithms, rather than being a real physical effect.

V. UNDERLYING EVENT AND HADRONIZATION

A separate source of uncertainties in jet measurements comes from what is colloquially known as "splash-in" and "splash-out" corrections. "Splash-in" corrections arise from the underlying event, which can deposit additional energy into the jet cone; we will refer to these more formally as underlying event (UE) corrections. "Splash-out" corrections come from the hadronization process of the jet which may move some of the jet energy outside the defined jet cone. We will refer to these as hadronization corrections (HC).

In either case, the correction is modeled as adding an amount δP_T to the observed transverse momentum (or transverse energy) of the jet. We denote the average over many events of δP_T by $\langle \delta P_T \rangle$. A complete analysis of the UE and HC contributions was performed by Cacciari, Dasgupta, Magnea, Salam in Refs. [5–7]. We find this to be an entirely suitable method for our estimate of $\langle \delta P_T \rangle$, and we adapt their results in the following.

A. Underlying event (UE)

We can parametrize the effect of the underlying event corrections on the apparent P_T of the jet as

$$\langle \delta P_T \rangle_{\text{UE}} = \Lambda_{\text{UE}} \tfrac{1}{2} R^2 \quad , \tag{12}$$

where R is the cone radius of the jet and Λ_{UE} is the average transverse energy per unit rapidity in the underlying event. Because we model the "splash-in" energy as random and uncorrelated with how the jet develops, the contribution from the underlying event will scale as the area of the jet cone—hence the factor of R^2 in Eq. (12). At Tevatron energies, Ref. [5] finds

$$\Lambda_{\text{UE}}(1960 \text{ GeV}) \approx 3 \pm 1 \text{ GeV}. \tag{13}$$

Thus, the $\langle P_T \rangle$ shift from the underlying event corrections is given by

$$\langle \delta P_T \rangle_{\text{UE}} \approx +0.7 \text{ GeV} \pm 0.3 \text{ GeV}, \tag{14}$$

for a jet cone with $R = 0.7$.

B. Hadronization correction (HC)

The R dependence of hadronization correction is very different from that of the underlying event correction [5–7]. The smaller the jet cone is, the more likely it is that hadronization will spray hadrons out of the cone. Hence, we will parametrize these corrections as proportional to $1/R$. Following Ref. [5], we write the hadronization correction as

$$\langle \delta P_T^i \rangle_{\text{HC}} = -C_i \frac{2}{R} \mathcal{A}(\mu_I) \quad , \tag{15}$$

where $\mathcal{A}(\mu_I)$ parametrizes the soft gluon radiation. Reference [5] takes $\mu_I = 2$ GeV, and finds $\mathcal{A}(2 \text{ GeV}) \approx 0.2$ GeV. In Eq. (15), C_i is a color factor that depends on

whether the jet is initiated by a quark, for which $C_i = C_F = 4/3$, or by a gluon, for which $C_i = C_A = 3$. We thus need an estimate of the fraction of jets that are gluon jets. Using calculations from the literature [8], we estimate that, for the Tevatron in the low P_T region, the fractions of quark and gluon jets are approximately

$$f_q \approx \tfrac{2}{3}, \qquad f_g \approx \tfrac{1}{3}.$$

Using these fractions, we can form a weighted average of the quark and gluon terms to obtain

$$\langle \delta P_T \rangle_{\mathrm{HC}} = f_q \langle \delta P_T^q \rangle_{\mathrm{HC}} + f_g \langle \delta P_T^g \rangle_{\mathrm{HC}}$$
$$= -f_q \frac{2C_F}{R} \mathcal{A}(\mu_I) - f_g \frac{2C_A}{R} \mathcal{A}(\mu_I)$$
$$\approx -1 \text{ GeV} \pm 0.5 \text{ GeV}. \tag{16}$$

Here, we have used a typical cone radius of $R = 0.7$ and taken a conservative choice for the uncertainty of 50% of the correction.

C. $\langle \delta P_T \rangle$ from the UE and HC

Combining the underlying event of Eq. (14) and the hadronization corrections of Eq. (16), the net P_T shift is

$$\langle \delta P_T \rangle \approx -0.3 \text{ GeV} \pm 0.6 \text{ GeV}, \tag{17}$$

where we have added the separate uncertainties in quadrature.

The individual underlying event and hadronization results for $\langle \delta P_T \rangle$ are displayed in Fig. 7 for the Tevatron using the parametrizations of Eq. (14) and (16). The combined result for $\langle \delta P_T \rangle$, including the uncertainty band, is also displayed. The underlying event and hadronization
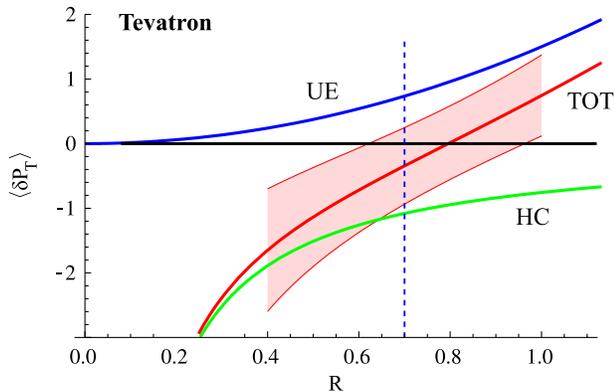

FIG. 7 (color online).   We display the expected P_T shift, $\langle \delta P_T \rangle$, in GeV vs jet cone radius R for the UE, HC, and combined results (TOT) at the Tevatron. The calculation of the HC uses a combination of quark-initiated ($f_q = 2/3$) and gluon-initiated ($f_g = 1/3$) jets. The upper solid (blue) line represents the UE correction, and the lower solid (green) line represents the HC terms. The combination of these corrections (TOT) is represented by the central (red) band including the uncertainties. The vertical line corresponds to $R = 0.7$.

corrections have opposite sign, and we note that for a jet cone radius of $R = 0.7$, the two corrections nearly cancel each other.

D. From $\langle \delta P_T \rangle$ to $\delta \sigma$

The differential jet cross section can be approximated by a power law of the form

$$\frac{d\sigma(P_T)}{dP_T} \approx \frac{\text{const}}{P_T^n}. \tag{18}$$

in the specific P_T range of interest. For jets at the Tevatron in the intermediate P_T range of $\sim[50, 300]$ GeV, we find $n \approx 7$ as illustrated by Fig. 8.

The effect of the underlying event and hadronization corrections is to shift the jet P_T from its value P_T^{pert} at the NLO parton level to a new value

$$P_T = P_T^{\mathrm{pert}} + \langle \delta P_T \rangle \quad,$$

where $\langle \delta P_T \rangle$ is the average change in the transverse jet transverse momentum due to underlying event additions and hadronization subtractions from Eq. (17).

If we write the true differential cross section as a function f,

$$\frac{d\sigma(P_T)}{dP_T} \equiv f(P_T) \quad,$$

then f is related to the perturbatively calculated function f_{pert} by

$$f(P_T) \approx f_{\mathrm{pert}}(P_T^{\mathrm{pert}}) = f_{\mathrm{pert}}(P_T - \langle \delta P_T \rangle).$$

We can perform a Taylor expansion about P_T for small δP_T,
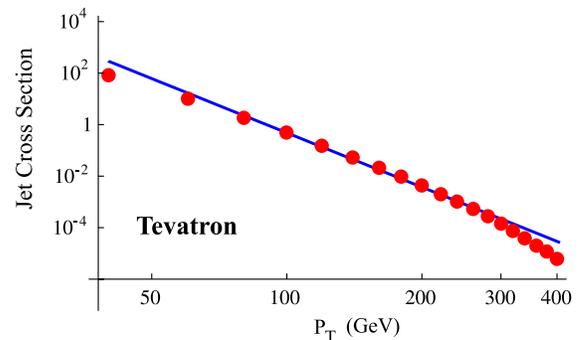

FIG. 8 (color online).   Jet cross section $d^2\sigma/dP_T/dy$ vs P_T in GeV with $y = 0$ at the Tevatron in units of nb/GeV. The line is a power law fit with $n = 7$; this describes the slope of the jet data in the range $P_T \approx [50, 300]$ GeV.

$$f(P_T) \approx f_{\text{pert}}(P_T - \langle \delta P_T \rangle)$$

$$\approx f_{\text{pert}}(P_T) - \langle \delta P_T \rangle \frac{df'_{\text{pert}}(P_T)}{dP_T}$$

$$= f_{\text{pert}}(P_T) \left\{ 1 + n \frac{\langle \delta P_T \rangle}{P_T} \right\}.$$

Here we have used the power law of Eq. (18) to replace $f'(P_T)$ by $-nf(P_T)/P_T$. Thus, to first order we find[6]

$$\frac{d\sigma}{dP_T} \approx \frac{d\sigma_{\text{pert}}}{dP_T} \left[1 + n \frac{\langle \delta P_T \rangle}{P_T} + \cdots \right], \quad (19)$$

so that the fractional correction is $n\langle \delta P_T \rangle/P_T$. Using $n \approx 7$ and the estimate from Eq. (17) of $\langle \delta P_T \rangle$, we find that the fractional correction to the cross section is approximately

$$7 \times \frac{-0.3 \text{ GeV} \pm 0.6 \text{ GeV}}{P_T} \approx -\frac{2 \text{ GeV}}{P_T} \pm \frac{4 \text{ GeV}}{P_T}.$$

Thus we estimate the fractional uncertainty from the underlying event and hadronization to be $4 \text{ GeV}/P_T$.

We account for this source of uncertainty by adding a new function $f_J(P_T, y)$ with $J = 7$,

$$f_7(P_T, y) = \frac{4 \text{ GeV}}{P_T} \quad (20)$$

for Tevatron jets in the P_T range of $\sim[50, 300]$ GeV.

VI. SUMMARY FOR THE TEVATRON

We have described the correlated theoretical systematic uncertainty using a total of seven functions, as summarized in Table I. The net error at any one value of $\{P_T, y\}$ is obtained by adding these seven functions in quadrature

$$\mathcal{E}(P_T, y) \equiv \sqrt{\sum f_J(P_T, y)^2}. \quad (21)$$

We now summarize the complete set of contributions to the uncertainty of the differential jet cross section as a function of $\{P_T, y\}$ for the Tevatron:

$$f_1(P_T, y) = \frac{9.62 \times 10^{-2}}{\log(M(y)/P_T)}, \quad f_2(P_T, y) = \frac{2.89 \times 10^{-2} y^2}{\log(M(y)/P_T)}$$

$$f_3(P_T, y) = 8.42 \times 10^{-2}, \quad f_4(P_T, y) = 0.842 \times 10^{-2} y^2,$$

$$f_5(P_T, y) = 1.68 \times 10^{-2} \log\left(\frac{15 P_T}{M(y)}\right),$$

$$f_6(P_T, y) = 0.336 \times 10^{-2} y^2 \log\left(\frac{15 P_T}{M(y)}\right),$$

$$f_7(P_T, y) = \frac{4 \text{ GeV}}{P_T}. \quad (22)$$

We display these results in Fig. 9. For $P_T \gtrsim 100$ GeV, the perturbative uncertainties are dominant, and slowly rise

[6]Cf., Eq. (5.9) of Dasgupta *et al.* in Ref. [5]

TABLE I. A compilation of the source of uncertainties (f_J) that comprise the total jet cross section uncertainty $\mathcal{E}$. The perturbative uncertainties arise from the higher, uncalculated, orders of perturbation theory and are estimated using the $\{\mu_F, \mu_R\}$ scale variation of the calculated cross section. The nonperturbative uncertainties are an estimate of the underlying event and hadronization corrections.

Uncertainty f_J	Source
$\{f_1, f_2, f_3, f_4, f_5, f_6\}$	perturbative
f_7	nonperturbative

with increasing P_T; this results holds across the full y-range, but the rise with P_T is more pronounced at large y. For $P_T \lesssim 100$ GeV, the uncertainty from the UE and HC terms become increasingly important as P_T decreases.

VII. THEORY ERRORS AT THE LHC

Having demonstrated the method for determining the theoretical systematic uncertainty at the Tevatron, we perform a parallel analysis for the Large Hadron Collider (LHC).

A. Perturbative uncertainty

We again estimate the error from not having calculated beyond NLO by using the dependence of the NLO cross section on the scales $\{\mu_R, \mu_F\}$, just as in the Tevatron case, and this yields the functions $\{f_1, f_2, f_3, f_4, f_5, f_6\}$ summarized in Eq. (27) at the end of this section.

B. Underlying event and hadronization

We proceed as in Sec. V for the Tevatron, accounting for the changed circumstances at the LHC. We first need to estimate the error in the determination of the contribution to the average jet transverse momentum, $\langle \delta P_T \rangle$, arising from the underlying event and from hadronization.

The underlying event contribution to $\langle \delta P_T \rangle$ is determined by the parameter Λ_{UE} in Eq. (12). Consistently with Refs. [5–7], for the LHC we take $\Lambda_{\text{UE}}(14 \text{ TeV}) \approx 10 \pm 4$ GeV, and obtain

$$\langle \delta P_T \rangle_{\text{UE}} \approx +2.5 \text{ GeV} \pm 1 \text{ GeV}. \quad (23)$$

For the contribution to $\langle \delta P_T \rangle$ from hadronization, we use Eq. (16) with $\mathcal{A}(\mu_I) \approx 0.2$ GeV as before. For the fractions f_q and f_g of quark and gluon jets in the relatively low P_T region where the hadronization corrections are significant, we use

$$f_q \approx \tfrac{1}{3}, \quad f_g \approx \tfrac{2}{3}.$$

Using these fractions, we can form a weighted average of the quark and gluon terms and estimate the hadronization contribution to $\langle \delta P_T \rangle$ to be

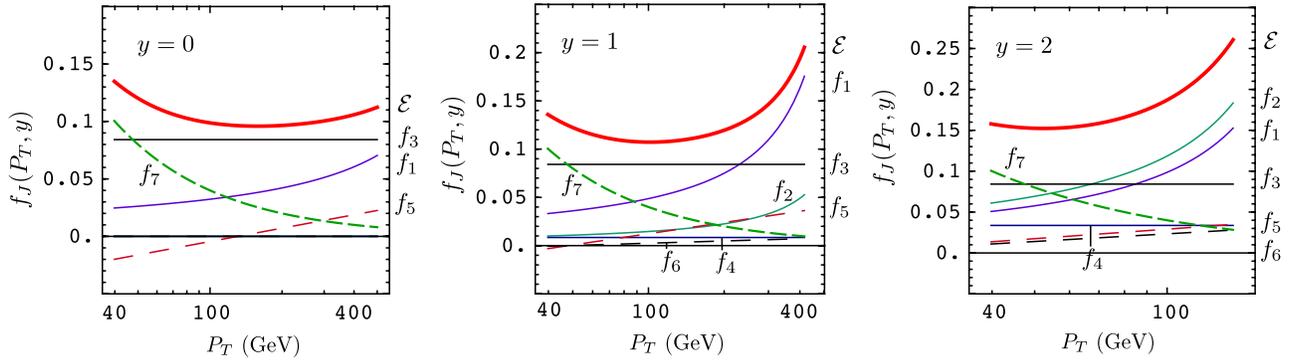$$\langle \delta P_T \rangle_{\text{HC}} = -1.4 \text{ GeV} \pm 0.7 \text{ GeV}. \quad (24)$$

FIG. 9 (color online).   A compilation of the uncertainties for jet production at the Tevatron ($\sqrt{s} = 1960$ GeV) for $y = \{0, 1, 2\}$. The numeric label corresponds to the error components summarized in Eq. (27). The upper thick (red) line is the quadrature sum of the individual errors.

Combining the underlying event and hadronization contributions, we estimate

$$\langle \delta P_T \rangle \approx +1 \text{ GeV} \pm 1.2 \text{ GeV}, \qquad (25)$$

where we have added the separate uncertainties in quadrature.

The results for the underlying event and hadronization contribution to $\langle \delta P_T \rangle$ are displayed in Fig. 10 for the LHC using the parametrizations of Eq. (24) and (23) but with a variable cone size R.

The correction to $\langle \delta P_T \rangle$ determines the correction to the cross section via Eq. (19). For this, we need the power n that describes the approximate power law fall off of the cross section. As illustrated in Fig. 11, a power law with $n \approx 6$ describes the data over the range $P_T \approx [100, 1000]$ GeV. Using $n \approx 6$ and the estimate from Eq. (25) of $\langle \delta P_T \rangle$, we find that the fractional correction to the cross section is approximately

$$6 \times \frac{1 \text{ GeV} \pm 1.2 \text{ GeV}}{P_T} \approx \frac{6 \text{ GeV}}{P_T} \pm \frac{7 \text{ GeV}}{P_T}.$$

Thus we estimate the fractional uncertainty from the underlying event and hadronization to be 7 GeV/P_T. We include this in the estimate of systematic theoretical errors by including a function $f_7(P_T)$ given by

$$f_7(P_T) = \frac{7 \text{ GeV}}{P_T} \qquad (26)$$

for LHC jets in the range $P_T \approx [100, 1000]$ GeV.

C. Summary: LHC

We now summarize the complete set of contributions to the uncertainty of the differential jet cross section as a function of $\{P_T, y\}$ for the LHC:
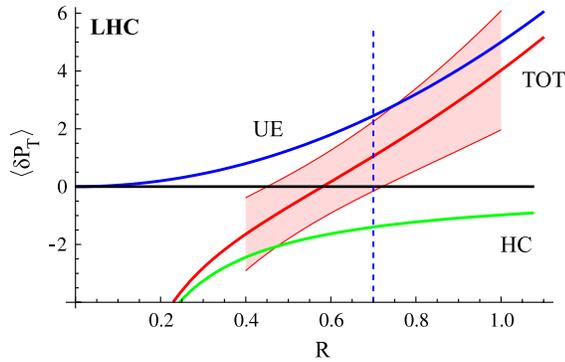

FIG. 10 (color online).   We display the expected P_T shift, $\langle \delta P_T \rangle$, in GeV vs jet cone radius R for the UE, HC, and combined results (TOT) at the LHC. The calculation of the HC uses a combination of quark-initiated ($f_q = 1/3$) and gluon-initiated ($f_g = 2/3$) jets. The upper solid (blue) line represents the UE correction, and the lower solid (green) line represents the HC terms. The combination of these corrections (TOT) is represented by the central (red) band including the uncertainties. The vertical line corresponds to $R = 0.7$.
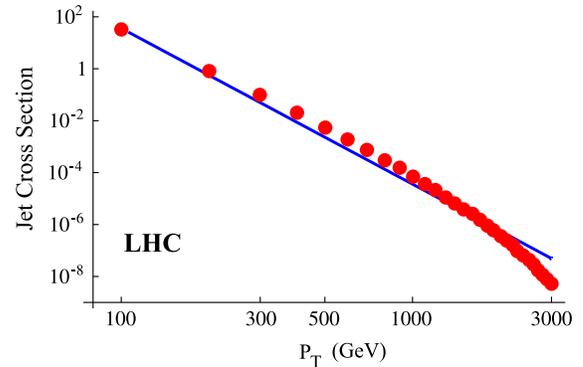

FIG. 11 (color online).   Jet cross section $d^2\sigma/dP_T/dy$ vs P_T in GeV with $y = 0$ at the LHC ($\sqrt{s} = 14$ TeV) in units of nb/GeV. The line is a power law fit with $n = 6$; this describes the slope of the jet data in the range $P_T \approx [100, 1000]$ GeV.
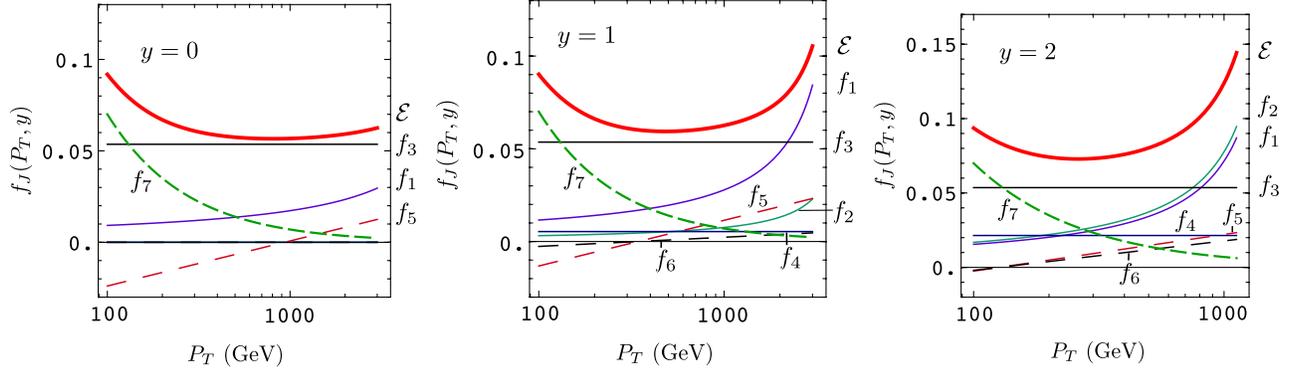
FIG. 12 (color online).   A compilation of the uncertainties for jet production at the LHC ($\sqrt{s} = 14,000$ GeV) for $y = \{0, 1, 2\}$. The numeric label corresponds to the error components summarized in Eq. (27). The upper thick (red) line is the quadrature sum of the individual errors.

$$f_1(P_T, y) = \frac{4.56 \times 10^{-2}}{\log(M(y)/P_T)}, \qquad f_2(P_T, y) = \frac{1.24 \times 10^{-2} y^2}{\log(M(y)/P_T)}$$

$$f_3(P_T, y) = 5.36 \times 10^{-2}, \qquad f_4(P_T, y) = 0.536 \times 10^{-2} y^2,$$

$$f_5(P_T, y) = 1.07 \times 10^{-2} \log\left(\frac{15 P_T}{M(y)}\right),$$

$$f_6(P_T, y) = 0.214 \times 10^{-2} y^2 \log\left(\frac{15 P_T}{M(y)}\right),$$

$$f_7(P_T, y) = \frac{7 \text{ GeV}}{P_T}. \tag{27}$$

We display these results in Fig. 12. In the central rapidity ($y \sim 0$) region for $P_T \gtrsim 500$ GeV the perturbative uncertainties are dominant and slowly rise with increasing P_T, while for $P_T \lesssim 500$ GeV the nonperturbative uncertainties become increasingly important. For $y = 2$, the transition P_T is closer to 300 GeV than 500 GeV.

VIII. CONCLUSIONS

As the LHC prepares to take data, it is important that we be able to determine whether a physics signal is consistent with the standard model. For example, if we observe a signal that is inconsistent with the standard model predic-

tion, but this inconsistency includes *only* experimental errors, we cannot claim this is "new physics" until we demonstrate it is also inconsistent including *both* experimental and theoretical errors. This paper provides a framework to quantitatively make such a determination in the case of jet physics. Similarly, this paper provides a framework to quantitatively fit parton distribution functions to Tevatron and LHC jet data, including estimated errors from the theory.

The framework that we provide involves functions $f_J(P_T, y)$ that represent independent contributions to the theory error. We note that other authors might estimate the errors differently and thus produce different functions $f_J(P_T, y)$. We hope that this will happen and that the merit of different choices will be debated.

[1] Stephen D. Ellis, Zoltan Kunszt, and Davison E. Soper, Phys. Rev. Lett. **64**, 2121 (1990).
[2] Charalampos Anastasiou, Gunther Dissertori, and Fabian Stockli, J. High Energy Phys. 09 (2007) 018.
[3] Nikolaos Kidonakis, arXiv:hep-ph/0110145.
[4] T. Kluge, K. Rabbertz, and M. Wobisch, arXiv:hep-ph/0609285.
[5] Mrinal Dasgupta, Lorenzo Magnea, and Gavin P. Salam, J.

High Energy Phys. 02 (2008) 055.
[6] Matteo Cacciari, Mrinal Dasgupta, Lorenzo Magnea, and Gavin Salam, arXiv:0706.3157.
[7] Mrinal Dasgupta, Lorenzo Magnea, and Gavin Salam, arXiv:0805.2267.
[8] H. L. Lai *et al.*, Phys. Rev. D **55**, 1280 (1997).
[9] Nikolaos Kidonakis and J. F. Owens, Phys. Rev. D **63**, 054019 (2001).