

REVIEW ARTICLE

# Hard Interactions of Quarks and Gluons: a Primer for LHC Physics

J. M. Campbell  
Department of Physics and Astronomy  
University of Glasgow  
Glasgow G12 8QQ  
United Kingdom

J. W. Huston  
Department of Physics and Astronomy  
Michigan State University  
East Lansing, MI 48840  
USA

W. J. Stirling  
Institute for Particle Physics Phenomenology  
University of Durham  
Durham DH1 3LE  
United Kingdom

**Abstract.** In this review article, we will develop the perturbative framework for the calculation of hard scattering processes. We will undertake to provide both a reasonably rigorous development of the formalism of hard scattering of quarks and gluons as well as an intuitive understanding of the physics behind the scattering. We will emphasize the role of logarithmic corrections as well as power counting in  $\alpha_s$  in order to understand the behaviors of hard scattering processes. We will include “rules of thumb” as well as “official recommendations”, and where possible will seek to dispel some myths. We will also discuss the impact of soft processes on the measurements of hard scattering processes. Experiences that have been gained at the Fermilab Tevatron will be recounted and, where appropriate, extrapolated to the LHC.

## 1. Introduction

Scattering processes at high energy hadron colliders can be classified as either hard or soft. Quantum Chromodynamics (QCD) is the underlying theory for all such processes, but the approach and level of understanding is very different for the two cases. For hard processes, e.g. Higgs or high  $p_T$  jet production, the rates and event properties

can be predicted with good precision using perturbation theory. For soft processes, e.g. the total cross section, the underlying event etc., the rates and properties are dominated by non-perturbative QCD effects, which are less well understood. For many hard processes, soft interactions are occurring along with the hard interactions and their effects must be understood for comparisons to be made to perturbative predictions. An understanding of the rates and characteristics of predictions for hard processes, both signals and backgrounds, using perturbative QCD (pQCD) is crucial for both the Tevatron and LHC.

In this review article, we will develop the perturbative framework for the calculation of hard scattering processes. We will undertake to provide both a reasonably rigorous development of the formalism of hard scattering of quarks and gluons as well as an intuitive understanding of the physics behind the scattering. We will emphasize the role of logarithmic corrections as well as power counting in  $\alpha_s$  in order to understand the behaviors of hard scattering processes. We will include “rules of thumb” as well as “official recommendations”, and where possible will seek to dispel some myths. We will also discuss the impact of soft processes on the measurements of hard scattering processes.

In Section 2, we introduce the hard scattering formalism and the QCD factorization theorem. In Section 3, we apply this formalism to some basic processes at leading order, next-to-leading order and next-to-next-to-leading order. Section 4 provides a detailed discussion of parton distribution functions and global pdf fits and in Section 5 we compare the predictions of pQCD to measurements at the Tevatron. Lastly, in Section 6, we provide some benchmarks and predictions for measurements to be performed at the LHC.

As a bonus feature, this paper comes along with a free companion website ‡, where updates and more detailed discussions than are possible in this limited space will be available.

## 2. Hard Scattering Formalism and the QCD Factorization Theorem

### 2.1. Introduction

In this section we will discuss in more detail how the QCD factorisation theorem can be used to calculate a wide variety of hard scattering cross sections in hadron-hadron collisions. For simplicity we will restrict our attention to leading-order processes and calculations; the extension of the formalism to more complicated processes and to include higher-order perturbative contributions will be discussed in later sections.

We begin with a brief review of the factorisation theorem. It was first pointed out by Drell and Yan [1] more than 30 years ago that parton model ideas developed for deep inelastic scattering could be extended to certain processes in hadron-hadron collisions. The paradigm process was the production of a massive lepton pair by quark-antiquark

‡ [www.pa.msu.edu/~huston/Les\\_Houches\\_2005/Les\\_Houches\\_SM.html](http://www.pa.msu.edu/~huston/Les_Houches_2005/Les_Houches_SM.html)

annihilation — the Drell–Yan process — and it was postulated that the hadronic cross section  $\sigma(AB \rightarrow \mu^+\mu^- + X)$  could be obtained by weighting the subprocess cross section  $\hat{\sigma}$  for  $q\bar{q} \rightarrow \mu^+\mu^-$  with the parton distribution functions (pdf's)  $f_{q/A}(x)$  extracted from deep inelastic scattering:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a) f_{b/B}(x_b) \hat{\sigma}_{ab \rightarrow X} , \quad (1)$$

where for the Drell–Yan process,  $X = l^+l^-$  and  $ab = q\bar{q}, \bar{q}q$ . The domain of validity is the asymptotic ‘scaling’ limit (the analogue of the Bjorken scaling limit in deep inelastic scattering)  $M_X \equiv M_{l^+l^-}, s \rightarrow \infty, \tau = M_{l^+l^-}^2/s$  fixed. The good agreement between theoretical predictions and the measured cross sections provided confirmation of the parton model formalism, and allowed for the first time a rigorous, quantitative treatment of certain hadronic cross sections. Studies were extended to other ‘hard scattering’ processes, for example the production of hadrons and photons with large transverse momentum, with equally successful results. Problems, however, appeared to arise when perturbative corrections from real and virtual gluon emission were calculated. Large logarithms from gluons emitted collinear with the incoming quarks appeared to spoil the convergence of the perturbative expansion. It was subsequently realised that these logarithms were the same as those that arise in deep inelastic scattering structure function calculations, and could therefore be absorbed, via the DGLAP equations, in the definition of the parton distributions, giving rise to logarithmic violations of scaling. The key point was that *all* logarithms appearing in the Drell–Yan corrections could be factored into renormalized parton distributions in this way, and *factorization theorems* which showed that this was a general feature of hard scattering processes were derived [2]. Taking into account the leading logarithm corrections, (1) simply becomes:

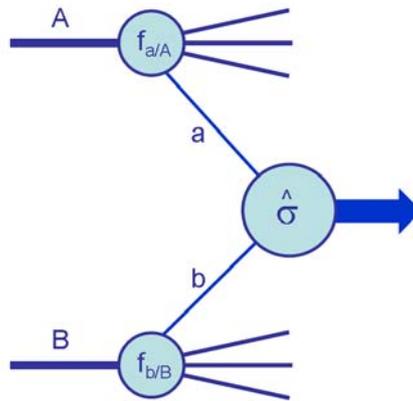
$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab \rightarrow X} . \quad (2)$$

corresponding to the structure depicted in Figure 1. The  $Q^2$  that appears in the parton distribution functions (pdfs) is a large momentum scale that characterizes the hard scattering, e.g.  $M_{l^+l^-}^2, p_T^2, \dots$ . Changes to the  $Q^2$  scale of  $\mathcal{O}(1)$ , e.g.  $Q^2 = 2M_{l^+l^-}^2, M_{l^+l^-}^2/2$  are equivalent in this leading logarithm approximation.

The final step in the theoretical development was the recognition that the *finite* corrections left behind after the logarithms had been factored were not universal and had to be calculated separately for each process, giving rise to perturbative  $\mathcal{O}(\alpha_S^n)$  corrections to the leading logarithm cross section of (2). Schematically

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times [\hat{\sigma}_0 + \alpha_S(\mu_R^2) \hat{\sigma}_1 + \dots]_{ab \rightarrow X} . \quad (3)$$

Here  $\mu_F$  is the *factorization scale*, which can be thought of as the scale that separates the long- and short-distance physics, and  $\mu_R$  is the *renormalization scale* for the QCD running coupling. Formally, the cross section calculated to all orders in perturbation theory is invariant under changes in these parameters, the  $\mu_F^2$  and  $\mu_R^2$  dependence of the coefficients, e.g.  $\hat{\sigma}_1$ , exactly compensating the explicit scale dependence of the parton distributions and the coupling constant. This compensation becomes more exact as



**Figure 1.** Diagrammatic structure of a generic hard scattering process.

more terms are included in the perturbation series, as will be discussed in more detail in the following section. In the absence of a complete set of higher order corrections, it is necessary to make a specific choice for the two scales in order to make cross section predictions. Different choices will yield different (numerical) results, a reflection of the uncertainty in the prediction due to unknown higher order corrections, see Section 3. To avoid unnaturally large logarithms reappearing in the perturbation series it is sensible to choose  $\mu_F$  and  $\mu_R$  values of the order of the typical momentum scales of the hard scattering process, and  $\mu_F = \mu_R$  is also often assumed. For the Drell-Yan process, for example, the standard choice is  $\mu_F = \mu_R = M$ , the mass of the lepton pair.

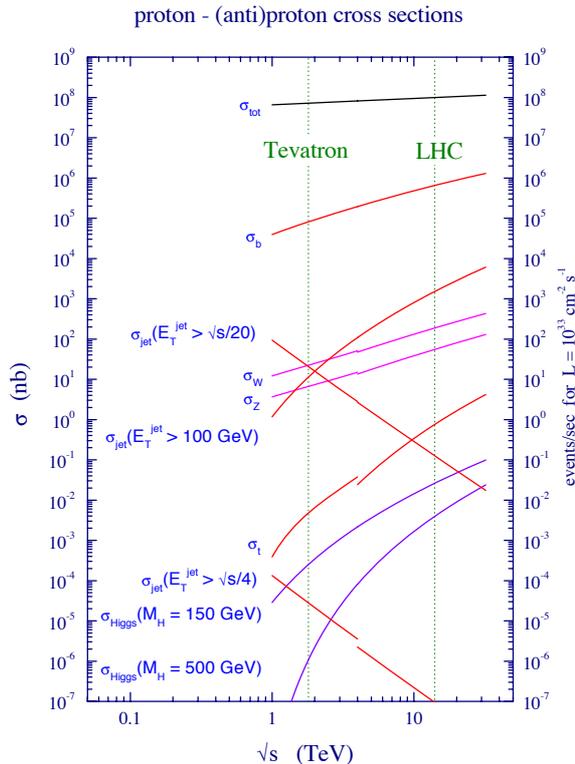
The recipe for using the above leading-order formalism to calculate a cross section for a given (inclusive) final state  $X$ +anything is very simple: (i) identify the leading-order partonic process that contributes to  $X$ , (ii) calculate the corresponding  $\hat{\sigma}_0$ , (iii) combine with an appropriate combination (or combinations) of pdfs for the initial-state partons  $a$  and  $b$ , (iv) make a specific choice for the scales  $\mu_F$  and  $\mu_R$ , and (v) perform a numerical integration over the variables  $x_a$ ,  $x_b$  and any other phase-space variables associated with the final state  $X$ . Some simple examples are

|                            |  |
|----------------------------|--|
| Z production               | $q\bar{q} \rightarrow Z$   |
| top quark production       | $q\bar{q} \rightarrow t\bar{t}$ , $gg \rightarrow t\bar{t}$          |
| large $E_T$ jet production | $gg \rightarrow gg$ , $qg \rightarrow qg$ , $qq \rightarrow qq$ etc. |

where appropriate scale choices are  $M_Z$ ,  $m_t$ ,  $E_T$  respectively. Expressions for the corresponding subprocess cross sections  $\hat{\sigma}_0$  are widely available in the literature, see for example [3].

The parton distributions used in these hard-scattering calculations are solutions of the DGLAP equations [4].

$$\frac{\partial q_i(x, \mu^2)}{\partial \log \mu^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} \left\{ P_{q_i q_j}(y, \alpha_S) q_j\left(\frac{x}{y}, \mu^2\right) + P_{q_i g}(y, \alpha_S) g\left(\frac{x}{y}, \mu^2\right) \right\}$$



**Figure 2.** Standard Model cross sections at the Tevatron and LHC colliders.

$$\frac{\partial g(x, \mu^2)}{\partial \log \mu^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} \left\{ P_{gq_j}(y, \alpha_S) q_j\left(\frac{x}{y}, \mu^2\right) + P_{gg}(y, \alpha_S) g\left(\frac{x}{y}, \mu^2\right) \right\} \quad (4)$$

where the splitting functions have perturbative expansions:

$$P_{ab}(x, \alpha_S) = P_{ab}^{(0)}(x) + \frac{\alpha_S}{2\pi} P_{ab}^{(1)}(x) + \dots \quad (5)$$

Expressions for the LO and NLO splitting functions can be found in [3]. The DGLAP equations determine the  $Q^2$  dependence of the pdfs. The  $x$  dependence, on the other hand, has to be obtained from fitting deep inelastic and other hard-scattering data. This will be discussed in more detail in Section 4. Note that for consistency, the order of the expansion of the splitting functions should be the same as that of the subprocess cross section, see (3). Thus, for example, a full NLO calculation will include both the  $\hat{\sigma}_1$  term in (3) and the  $P_{ab}^{(1)}$  terms in the determination of the pdfs via (4) and (5).

Figure 2 shows the predictions for some important Standard Model cross sections at  $p\bar{p}$  and  $pp$  colliders, calculated using the above formalism (at next-to-leading order in perturbation theory, i.e. including also the  $\hat{\sigma}_1$  term in (3)).

We have already mentioned that the Drell–Yan process is the paradigm hadron–collider hard scattering process, and so we will discuss this in some detail in what follows. Many of the remarks apply also to other processes, in particular those shown in Figure 2, although of course the higher–order corrections and the initial–state parton combinations are process dependent.

## 2.2. The Drell–Yan process

The Drell–Yan process is the production of a lepton pair ( $e^+e^-$  or  $\mu^+\mu^-$  in practice) of large invariant mass  $M$  in hadron-hadron collisions by the mechanism of quark–antiquark annihilation [1]. In the basic Drell–Yan mechanism, a quark and antiquark annihilate to produce a virtual photon,  $q\bar{q} \rightarrow \gamma^* \rightarrow l^+l^-$ . At high-energy colliders, such as the Tevatron and LHC, there is of course sufficient centre-of-mass energy for the production of on-shell  $W$  and  $Z$  bosons as well. The cross section for quark-antiquark annihilation to a lepton pair via an intermediate massive photon is easily obtained from the fundamental QED  $e^+e^- \rightarrow \mu^+\mu^-$  cross section, with the addition of the appropriate colour and charge factors.

$$\hat{\sigma}(q\bar{q} \rightarrow e^+e^-) = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{N} Q_q^2, \quad (6)$$

where  $Q_q$  is the quark charge:  $Q_u = +2/3$ ,  $Q_d = -1/3$  etc. The overall colour factor of  $1/N = 1/3$  is due to the fact that only when the colour of the quark matches with the colour of the antiquark can annihilation into a colour-singlet final state take place.

In general, the incoming quark and antiquark will have a spectrum of centre-of-mass energies  $\sqrt{\hat{s}}$ , and so it is more appropriate to consider the differential mass distribution:

$$\frac{d\hat{\sigma}}{dM^2} = \frac{\hat{\sigma}_0}{N} Q_q^2 \delta(\hat{s} - M^2), \quad \hat{\sigma}_0 = \frac{4\pi\alpha^2}{3M^2}, \quad (7)$$

where  $M$  is the mass of the lepton pair. In the centre-of-mass frame of the two hadrons, the components of momenta of the incoming partons may be written as

$$\begin{aligned} p_1^\mu &= \frac{\sqrt{\hat{s}}}{2} (x_1, 0, 0, x_1) \\ p_2^\mu &= \frac{\sqrt{\hat{s}}}{2} (x_2, 0, 0, -x_2). \end{aligned} \quad (8)$$

The square of the parton centre-of-mass energy  $\hat{s}$  is related to the corresponding hadronic quantity by  $\hat{s} = x_1x_2s$ . Folding in the pdfs for the initial state quarks and antiquarks in the colliding beams gives the hadronic cross section:

$$\begin{aligned} \frac{d\sigma}{dM^2} &= \frac{\hat{\sigma}_0}{N} \int_0^1 dx_1 dx_2 \delta(x_1x_2s - M^2) \\ &\times \left[ \sum_k Q_k^2 (q_k(x_1, M^2)\bar{q}_k(x_2, M^2) + [1 \leftrightarrow 2]) \right]. \end{aligned} \quad (9)$$

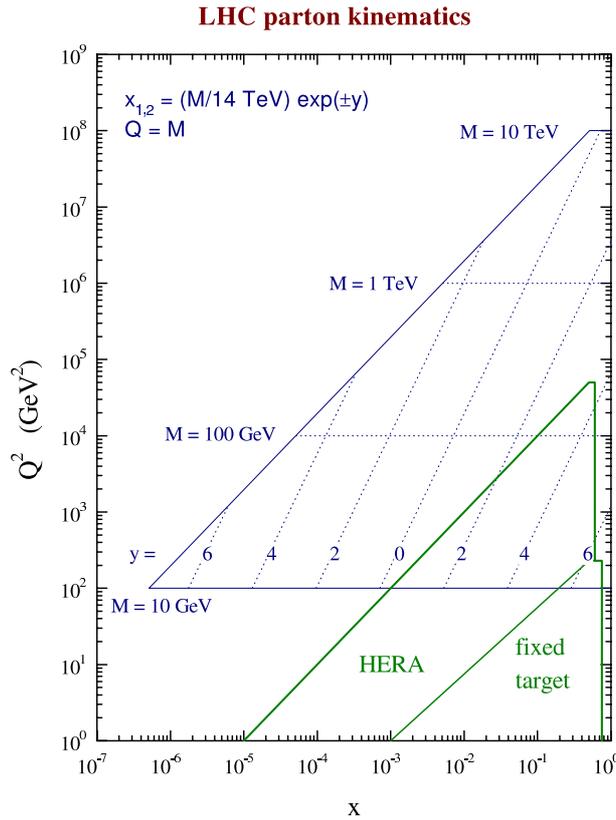
From (8), the rapidity of the produced lepton pair is found to be  $y = 1/2 \ln(x_1/x_2)$ , and hence

$$x_1 = \sqrt{\tau} e^y, \quad x_2 = \sqrt{\tau} e^{-y}. \quad (10)$$

The double-differential cross section is therefore

$$\frac{d\sigma}{dM^2 dy} = \frac{\hat{\sigma}_0}{N_S} \left[ \sum_k Q_k^2 (q_k(x_1, M^2)\bar{q}_k(x_2, M^2) + [1 \leftrightarrow 2]) \right]. \quad (11)$$

with  $x_1$  and  $x_2$  given by (10). Thus different values of  $M$  and  $y$  probe different values of the parton  $x$  of the colliding beams. The formulae relating  $x_1$  and  $x_2$  to  $M$  and  $y$



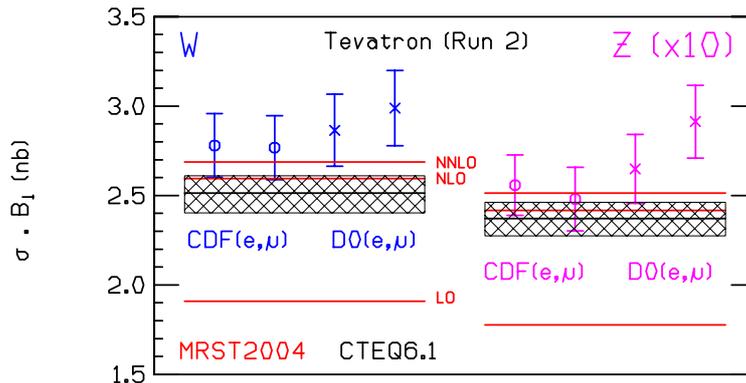
**Figure 3.** Graphical representation of the relationship between parton  $(x, Q^2)$  variables and the kinematic variables corresponding to a final state of mass  $M$  produced with rapidity  $y$  at the LHC collider with  $\sqrt{s} = 14$  TeV.

of course also apply to the production of *any* final state with this mass and rapidity. Assuming the factorisation scale ( $Q$ ) is equal to  $M$ , the mass of the final state, the relationship between the parton  $(x, Q^2)$  values and the kinematic variables  $M$  and  $y$  is illustrated pictorially in Figure 3, for the LHC collision energy  $\sqrt{s} = 14$  TeV. For a given rapidity  $y$  there are two (dashed) lines, corresponding to the values of  $x_1$  and  $x_2$ . For  $y = 0$ ,  $x_1 = x_2 = M/\sqrt{s}$ .

In analogy with the Drell–Yan cross section derived above, the subprocess cross sections for (on-shell)  $W$  and  $Z$  production are readily calculated to be

$$\begin{aligned}\hat{\sigma}^{q\bar{q}'\rightarrow W} &= \frac{\pi}{3}\sqrt{2}G_F M_W^2 |V_{qq'}|^2 \delta(\hat{s} - M_W^2), \\ \hat{\sigma}^{q\bar{q}\rightarrow Z} &= \frac{\pi}{3}\sqrt{2}G_F M_Z^2 (v_q^2 + a_q^2) \delta(\hat{s} - M_Z^2),\end{aligned}\quad (12)$$

where  $V_{qq'}$  is the appropriate Cabibbo–Kobayashi–Maskawa matrix element, and  $v_q$  ( $a_q$ ) is the vector (axial vector) coupling of the  $Z$  to the quarks. These formulae are valid in the *narrow width production* in which the decay width of the gauge boson is neglected. The resulting cross sections can then be multiplied by the branching ratio for any particular hadronic or leptonic final state of interest.



**Figure 4.** Predictions for the  $W$  and  $Z$  total cross sections at the Tevatron and LHC, using MRST2004 [5] and CTEQ6.1 pdf's [6], compared with recent data from CDF and D0. The MRST predictions are shown at LO, NLO and NNLO. The CTEQ6.1 NLO predictions and the accompanying pdf error bands are also shown.

High-precision measurements of  $W$  and  $Z$  production cross sections from the Fermilab Tevatron  $p\bar{p}$  collider are available and allow the above formalism to be tested quantitatively. Thus Figure 4 shows the cross sections for  $W^\pm$  and  $Z^0$  production and decay into various leptonic final states from the CDF [7] and D0 [8] collaborations at the Tevatron. The theoretical predictions are calculated at LO (i.e. using (12)), NLO and NNLO in perturbation theory using the  $\overline{\text{MS}}$  scheme MRST parton distributions of [5], with renormalization and factorization scales  $\mu_F = \mu_R = M_W, M_Z$ . The net effect of the NLO and NNLO corrections, which will be discussed in more detail in the following Section, is to increase the lowest-order cross section by about 25% and 5% respectively.

Perhaps the most important point to note from Figure 4 is that, aside from unknown (and presumably small)  $O(\alpha_s^3)$  corrections, there is virtually no theoretical uncertainty associated with the predictions – the parton distributions are being probed in a range of  $x \sim M_W/\sqrt{s}$  where they are constrained from deep inelastic scattering, see Figure 3, and the scale dependence is weak [5]. This overall agreement with experiment, therefore, provides a powerful test of the whole theoretical edifice that goes into the calculation. Figure 4 also illustrates the importance of higher-order perturbative corrections when making detailed comparisons of data and theory.

### 2.3. Transverse momentum distributions

Like Drell-Yan lepton pairs, most  $W$  and  $Z$  bosons (here collectively denoted by  $V$ ) are produced with relatively little transverse momentum, i.e.  $p_T \ll M_V$ . In the leading-order model discussed above, in which the colliding partons are assumed to be exactly collinear with the colliding beam particles, the gauge bosons are produced with zero transverse momentum. This approach does not take account of the intrinsic (non-perturbative) transverse motion of the quarks and gluons inside the colliding

hadrons, nor of the possibility of generating large transverse momentum by recoil against additional energetic partons produced in the hard scattering.

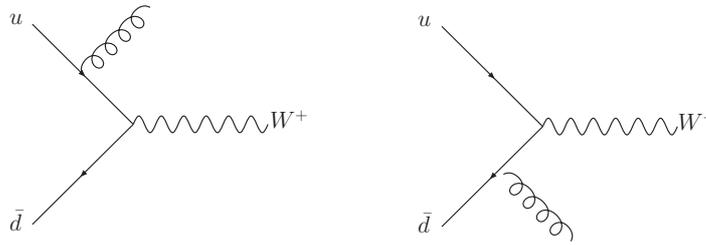
At very small  $p_T$ , the intrinsic transverse motion of the quarks and gluons inside the colliding hadrons,  $k_T \sim \Lambda_{\text{QCD}}$ , cannot be neglected. Indeed the measured  $p_T$  distribution of Drell–Yan lepton pairs produced in fixed-target  $pN$  collisions is well parametrised by assuming a gaussian distribution for the intrinsic transverse momentum with  $\langle k_T \rangle \sim 700$  MeV, see for example [3]. However the data on the  $p_T$  distribution also show clear evidence of a hard, power-law tail, and it is natural to attribute this to the (perturbative) emission of one or more hard partons, i.e.  $q\bar{q} \rightarrow Vg$ ,  $qg \rightarrow Vq$  etc. The Feynman diagrams for these processes are identical to those for large  $p_T$  direct photon production, and the corresponding annihilation and Compton matrix elements are, for  $W$  production,

$$\begin{aligned} \sum |\mathcal{M}^{q\bar{q} \rightarrow Wg}|^2 &= \pi\alpha_S \sqrt{2} G_F M_W^2 |V_{q\bar{q}}|^2 \frac{8}{9} \frac{t^2 + u^2 + 2M_W^2 s}{tu}, \\ \sum |\mathcal{M}^{gq \rightarrow Wq'}|^2 &= \pi\alpha_S \sqrt{2} G_F M_W^2 |V_{qq'}|^2 \frac{1}{3} \frac{s^2 + u^2 + 2tM_W^2}{-su}, \end{aligned} \quad (13)$$

with similar results for the  $Z$  boson and for Drell–Yan lepton pairs. The sum is over colours and spins in the final and initial states, with appropriate averaging factors for the former. The transverse momentum distribution  $d\sigma/dp_T^2$  is then obtained by convoluting these matrix elements with parton distributions in the usual way. In principle, one can combine the hard (perturbative) and intrinsic (non-perturbative) contributions, for example using a convolution integral in transverse momentum space, to give a theoretical prediction valid for all  $p_T$ . A more refined prediction would then include next-to-leading-order ( $O(\alpha_S^2)$ ) perturbative corrections, for example from processes like  $q\bar{q} \rightarrow Vgg$ , to the high  $p_T$  tail. Some fraction of the  $O(\alpha_S)$  and  $O(\alpha_S^2)$  contributions could be expected to correspond to distinct  $V + 1$  jet and  $V + 2$  jet final states respectively.

However, a major problem in carrying out the above procedure is that the  $2 \rightarrow 2$  matrix elements are singular when the final-state partons become soft and/or are emitted collinear with the initial-state partons. These singularities are related to the poles at  $t = 0$  and  $u = 0$  in the above matrix elements. In addition, processes like  $q\bar{q} \rightarrow Vgg$  are singular when the two final-state gluons become collinear. This means in practice that the lowest-order perturbative contribution to the  $p_T$  distribution is singular as  $p_T \rightarrow 0$ , and that higher-order contributions from processes like  $q\bar{q} \rightarrow Vgg$  are singular for any  $p_T$ .

The fact that the predictions of perturbative QCD *are* in fact finite for physical processes is due to a number of deep and powerful theorems (applicable to any quantum field theory) that guarantee that for suitably defined cross sections and distributions, the singularities arising from real and virtual parton emissions at intermediate stages of the calculation cancel when all contributions are included. We have already seen an example of this in the discussion above. The  $O(\alpha_S)$  contribution the total  $W$  cross section from the process  $q\bar{q} \rightarrow Wg$  is singular when  $p_T(W) = 0$ , but this singularity is exactly cancelled by a  $O(\alpha_S)$  contribution from a virtual gluon loop correction to



**Figure 5.** Lowest order diagrams for the production of a  $W$  and one jet at hadron colliders.

$q\bar{q} \rightarrow W$ . The net result is the finite NLO contribution to the cross section displayed in Figure 4. The details of how and under what circumstances these cancellations take place will be discussed in the following section.

### 3. Partonic cross sections

At the heart of the prediction of any hadron collider observable lies the calculation of the relevant hard scattering process. In this section we will outline the perturbative approaches that are employed to calculate these processes and describe some of their features and limitations. In addition, we will describe how the partonic calculations can be used to make predictions for an exclusive hadronic final state.

#### 3.1. Lowest order calculations

The simplest predictions can be obtained by calculating the lowest order in the perturbative expansion of the observable, as discussed in the previous section. This is performed by calculating the squared matrix element represented by tree-level Feynman diagrams and integrating this over the appropriate phase space. For the simplest cases and for certain observables only, the phase space integration can be performed analytically. For example, in Section 2, we calculated the lowest order cross section for Drell-Yan production. However, to obtain fully differential predictions in general, the integration must be carried out numerically. For most calculations, it is necessary to impose restrictions on the phase space in order that divergences in the matrix elements are screened. This can best be understood by consideration of one of the simplest such cases,  $W + 1$  jet production at a hadron collider.

*3.1.1.  $W + 1$  jet production* In Figure 5, we have extended the LO diagrams for Drell-Yan production (for the specific initial state  $u\bar{d}$ ) by adding a final state gluon to each of the initial state quark legs. This is one of the subprocesses responsible for  $W + 1$  jet production, with the other being  $gq \rightarrow Wq$ . After application of the Feynman rules,

the squared matrix elements obtained from the sum of the diagrams take the form:

$$|\mathcal{M}^{u\bar{d}\rightarrow W+g}|^2 \sim \left( \frac{\hat{t}^2 + \hat{u}^2 + 2Q^2 \hat{s}}{\hat{t}\hat{u}} \right), \quad (14)$$

where  $Q^2$  is the virtuality of the  $W$  boson,  $\hat{s} = s_{u\bar{d}}$ ,  $\hat{t} = s_{ug}$ , and  $\hat{u} = s_{\bar{d}g}$ , c.f. (13) of Section 2. This expression diverges in the limit where the gluon is unresolved – either it is collinear to one of the quarks ( $\hat{t} \rightarrow 0$  or  $\hat{u} \rightarrow 0$ ), or it is soft (both invariants vanish, so  $E_g \rightarrow 0$ ). Let us consider the impact of these divergences on the calculation of this cross section. In order to turn the matrix elements into a cross section, one must convolute with PDF's and perform the integration over the appropriate phase space,

$$\sigma = \int dx_1 dx_2 f_u(x_1, Q^2) f_{\bar{d}}(x_2, Q^2) \frac{|\mathcal{M}|^2}{32\pi^2 \hat{s}} \frac{d^3 p_W}{E_W} \frac{d^3 p_g}{E_g} \delta(p_u + p_{\bar{d}} - p_W - p_g), \quad (15)$$

where  $x_1, x_2$  are the momentum fractions of the  $u$  and  $\bar{d}$  quarks. These momentum fractions are of course related to the centre of mass energy of the collider  $s$  by the relation,  $\hat{s} = x_1 x_2 s$ .

After suitable manipulations, this can be transformed into a cross-section that is differential in  $Q^2$  and the transverse momentum ( $p_T$ ) and rapidity ( $y$ ) of the  $W$  boson [34],

$$\frac{d\sigma}{dQ^2 dy dp_T^2} \sim \frac{1}{s} \int dy_g f_u(x_1, Q^2) f_{\bar{d}}(x_2, Q^2) \frac{|\mathcal{M}|^2}{\hat{s}} \quad (16)$$

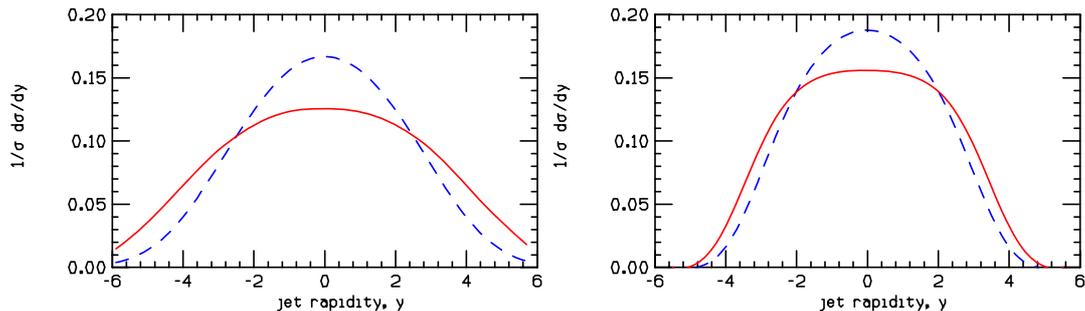
The remaining integral to be done is over the rapidity of the gluon,  $y_g$ . Note that the  $p_T$  of the gluon is related to the invariants of the process by  $p_T^2 = \hat{t}\hat{u}/\hat{s}$ , so that the leading divergence represented by the third term of (14) corresponds to the limit  $p_T \rightarrow 0$ . Furthermore, in this limit  $\hat{s} \rightarrow Q^2$ , so that the behaviour of the cross-section becomes,

$$\frac{d\sigma}{dQ^2 dy dp_T^2} \sim \frac{2}{s} \frac{1}{p_T^2} \int dy_g f_u(x_1, Q^2) f_{\bar{d}}(x_2, Q^2) + (\text{sub-leading in } p_T^2). \quad (17)$$

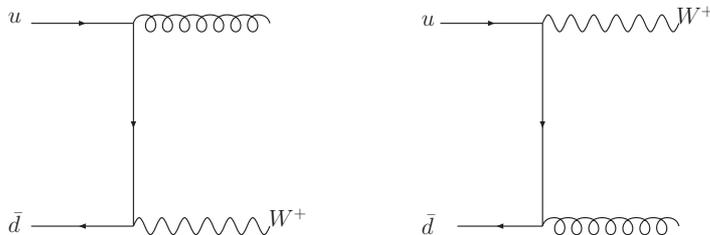
As the  $p_T$  of the  $W$  boson becomes small, the limits on the  $y_g$  integration are given by  $\pm \log(\sqrt{s}/p_T)$ . Under the assumption that the rest of the integrand is approximately constant, the integral can be simply performed. This yields,

$$\frac{d\sigma}{dQ^2 dy dp_T^2} \sim \frac{\log(s/p_T^2)}{p_T^2}, \quad (18)$$

so that the differential cross section contains a logarithmic dependence on  $p_T$ . If no cut is applied on the gluon  $p_T$  we see that this result makes no sense – only after applying a minimum value of the  $p_T$  do we obtain a finite result. However, applying a cutoff at  $p_T = p_{T,\min}$  and then performing the integration yields a result proportional to  $\log^2(s/p_{T,\min}^2)$ . This is typical of a fixed order expansion – it is not merely an expansion in  $\alpha_s$ , but in  $\alpha_s \log(\dots)$ , where the argument of the logarithm depends upon the process and cuts under consideration. As we shall discuss later, these logarithms may be systematically accounted for in various all-orders treatments.

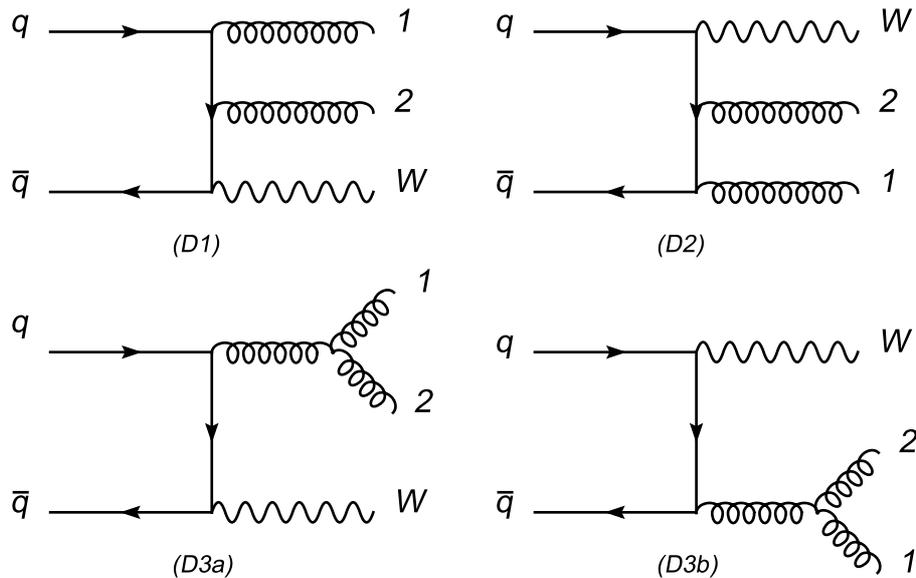


**Figure 6.** The rapidity distribution of the final state parton found in a lowest order calculation of the  $W + 1$  jet cross section at the LHC. The parton is required to have a  $p_T$  larger than 2 GeV (left) or 50 GeV(right). Contributions from  $q\bar{q}$  annihilation (solid red line) and the  $qg$  process (dashed blue line) are shown separately.



**Figure 7.** An alternative way of drawing the diagrams of Figure 5.

In Figure 6 we show the the rapidity distribution of the jet, calculated using this lowest order process. In the calculation, a sum over all species of quarks has been performed and the contribution from the quark-gluon process included. The rapidity distribution is shown for two different choices of minimum jet transverse momentum, which is the cut-off used to regulate the collinear divergences discussed above. For very small values of  $p_T$ , we can view the radiated gluon as being emitted from the quark line at an early time, typically termed “initial-state radiation”. From the left-hand plot, this radiation is indeed produced quite often at large rapidities, although it is also emitted centrally with a large probability. The canonical “wisdom” is that initial-state radiation is primarily found in the forward region. There is indeed a collinear pole in the matrix element so that a fixed energy gluon tends to be emitted close to the original parton direction. However, we are interested not in fixed energy but rather in fixed transverse momentum. When using a higher  $p_T$  cut-off the gluon is emitted less often at large rapidities and is more central, as shown by the plot on the right-hand side. In this case, one can instead think of the diagrams as a  $2 \rightarrow 2$  scattering as depicted in Figure 7. Of course, the manner in which such Feynman diagrams are drawn is purely a matter of convention. The diagrams are exactly the same as in Figure 5, but re-drawing them in this way is suggestive of the different kinematic region that is probed with a gluon at



**Figure 8.** The 4 diagrams that contribute to the matrix elements for the production of  $W + 2$  gluons when gluon 1 is soft.

high  $p_T$ .

There is also a collinear pole involved for the emission of gluons from final state partons. Thus, the gluons will be emitted preferentially near the direction of the emitting parton. In fact, it is just such emissions that give rise to the finite size of the jet arising from a single final state parton originating from the hard scatter. Much of the jet structure is determined by the hardest gluon emission; thus NLO theory, in which a jet consists of at most 2 partons, provides a good description of the jet shape observed in data [9].

*3.1.2.  $W+2$  jet production* By adding a further parton, one can simulate the production of a  $W + 2$  jet final state. Many different partonic processes contribute in general, so for the sake of illustration we just consider the production of a  $W$  boson in association with two gluons.

First, we shall study the singularity structure of the matrix elements in more detail. In the limit that one of the gluons,  $p_1$  is soft, the singularities in the matrix elements occur in 4 diagrams only. These diagrams, in which gluon  $p_1$  is radiated from an external line, are depicted in Figure 8. The remaining diagrams, in which gluon  $p_1$  is attached to an internal line, do not give rise to singularities because the adjacent propagator does not vanish in this limit.

This is the first of our examples in which the matrix elements contain non-trivial colour structure. Denoting the colour labels of gluons  $p_1$  and  $p_2$  by  $t^A$  and  $t^B$  respectively, diagram (1) is proportional to  $t^B t^A$ , whilst (2) is proportional to  $t^A t^B$ . The final two diagrams, (3a) and (3b) are each proportional to  $f^{ABC} t^C$ , which can of course be written as  $(t^A t^B - t^B t^A)$ . Using this identity, the amplitude (in the limit that  $p_1$  is soft) can be

written in a form in which the dependence on the colour matrices is factored out,

$$\mathcal{M}^{q\bar{q}\rightarrow Wgg} = t^A t^B (D_2 + D_3) + t^B t^A (D_1 - D_3) \quad (19)$$

so that the kinematic structures obtained from the Feynman rules are collected in the functions  $D_1$ ,  $D_2$  (for diagrams (1) and (2)) and  $D_3$  (the sum of diagrams (3a) and (3b)). The combinations that appear in (19) are often referred to as colour-ordered amplitudes.

With the colour factors stripped out, it is straightforward to square the amplitude in (19) using the identities  $\text{tr}(t^A t^B t^B t^A) = N C_F^2$  and  $\text{tr}(t^A t^B t^A t^B) = -C_F/2$ ,

$$\begin{aligned} |\mathcal{M}^{q\bar{q}\rightarrow Wgg}|^2 &= N C_F^2 \left[ |D_2 + D_3|^2 + |D_1 - D_3|^2 \right] - C_F (D_2 + D_3)(D_1 - D_3)^\dagger \\ &= \frac{C_F N^2}{2} \left[ |D_2 + D_3|^2 + |D_1 - D_3|^2 - \frac{1}{N^2} |D_1 + D_2|^2 \right]. \end{aligned} \quad (20)$$

Moreover, these colour-ordered amplitudes possess special factorization properties in the limit that gluon  $p_1$  is soft. In this limit they can be written as the product of an eikonal term and the matrix elements containing only one gluon,

$$\begin{aligned} D_2 + D_3 &\longrightarrow \epsilon_\mu \left( \frac{q^\mu}{p_1 \cdot q} - \frac{p_2^\mu}{p_1 \cdot p_2} \right) \mathcal{M}_{q\bar{q}\rightarrow Wg} \\ D_1 - D_3 &\longrightarrow \epsilon_\mu \left( \frac{p_2^\mu}{p_1 \cdot p_2} - \frac{\bar{q}^\mu}{p_1 \cdot \bar{q}} \right) \mathcal{M}_{q\bar{q}\rightarrow Wg} \end{aligned} \quad (21)$$

where  $\epsilon_\mu$  is the polarization vector for gluon  $p_1$ . The squares of these eikonal terms are easily computed using the replacement  $\epsilon_\mu \epsilon_\nu^* \rightarrow -g_{\mu\nu}$  to sum over the gluon polarizations. This yields terms of the form,

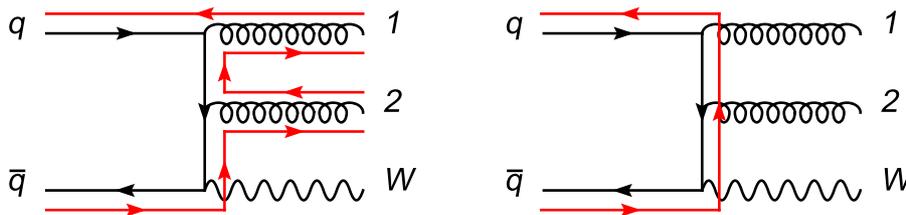
$$\frac{a \cdot b}{p_1 \cdot a p_1 \cdot b} \equiv [a \ b], \quad (22)$$

so that the final result is,

$$|\mathcal{M}^{q\bar{q}\rightarrow Wgg}|^2 \xrightarrow{\text{soft}} \frac{C_F N^2}{2} \left[ [q \ p_2] + [p_2 \ \bar{q}] - \frac{1}{N^2} [q \ \bar{q}] \right] \mathcal{M}^{q\bar{q}\rightarrow Wg}. \quad (23)$$

Inspecting this equation, one can see that the leading term (in the number of colours) contains singularities along two lines of colour flow – one connecting the gluon  $p_2$  to the quark, the other connecting it to the anti-quark. On the other hand, the sub-leading term has singularities along the line connecting the quark and anti-quark. It is these lines of colour flow that indicate the preferred directions for the emission of additional gluons. In the sub-leading term the colour flow does not relate the gluon colour to the parent quarks at all. The matrix elements are in fact the same as those for the emission of two photons from a quark line (apart from overall coupling factors) with no unique assignment to either diagram 1 or diagram 2, unlike the leading term. For this reason only the information about the leading colour flow is used by parton shower Monte Carlos such as HERWIG [10] and PYTHIA [11]. As a further example, the lines of colour flow in a  $W + 2$  jet event are shown in Figure 9.

Since all the partons are massless, it is trivial to re-write the eikonal factor of (22) in terms of the energy of the radiated gluon,  $E$  and the angle it makes with the hard



**Figure 9.** Two examples of colour flow in a  $W + 2$  jet event, shown in red. In the left hand diagram, a leading colour flow is shown. The right-hand diagram depicts the sub-leading colour flow resulting from interference.

partons,  $\theta_a, \theta_b$ . It can then be combined with the phase space for the emitted gluon to yield a contribution such as,

$$[a\ b] dPS_{\text{gluon}} = \frac{1}{E^2} \frac{1}{1 - \cos \theta_a} E dE d\cos \theta_a . \quad (24)$$

In this form, it is clear that the cross-section diverges as either  $\cos \theta_a \rightarrow 1$  (the gluon is emitted collinear to parton  $a$ ) or  $E \rightarrow 0$  (for any angle of radiation). Moreover, each divergence is logarithmic and regulating the divergence, by providing a fixed cutoff (either in angle or energy) will produce a single logarithm from collinear configurations and another from soft ones – just as we found when considering the specific case of  $W + 1$  jet production in the previous subsection.

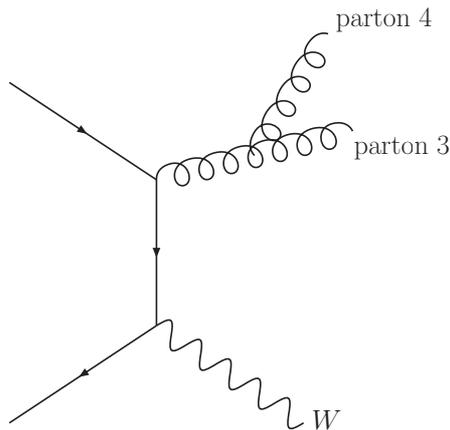
This argument can be applied at successively higher orders of perturbation theory. Each gluon that is added yields an additional power of  $\alpha_s$  and, via the eikonal factorization outlined above, can produce an additional two logarithms. This means that we can write the  $W + 1$  jet cross section schematically as a sum of contributions,

$$d\sigma = \sigma_0(W + 1 \text{ jet}) \left[ 1 + \alpha_s(c_{12}L^2 + c_{11}L + c_{10}) + \alpha_s^2(c_{24}L^4 + c_{23}L^3 + c_{22}L^2 + c_{21}L + c_{20}) + \dots \right] \quad (25)$$

where  $L$  represents the logarithm controlling the divergence, either soft or collinear. The size of  $L$  depends upon the criteria used to define the jets – the minimum transverse energy of a jet and the jet cone size. The coefficients  $c_{ij}$  in front of the logarithms depend upon colour factors. Note that the addition of each gluon results not just in an additional factor of  $\alpha_s$ , but in a factor of  $\alpha_s$  times logarithms. For many important kinematic configurations, the logs can be large, leading to an enhanced probability for additional gluon emissions to occur. For inclusive quantities, where the same cuts are applied to every jet, the logs tends to be small, and counting powers of  $\alpha_s$  becomes a valid estimator for the rate of production of additional jets.

Noticing that the factor  $(\alpha_s L)$  appears throughout (25), it is useful to re-write the expansion in brackets as,

$$\begin{aligned} [\dots] &= 1 + \alpha_s L^2 c_{12} + (\alpha_s L^2)^2 c_{24} + \alpha_s L c_{11} (1 + \alpha_s L^2 c_{23}/c_{11} + \dots) + \dots \\ &= \exp [c_{12} \alpha_s L^2 + c_{11} \alpha_s L] , \end{aligned} \quad (26)$$



**Figure 10.** A final state configuration containing a  $W$  and 2 partons. After the jet definition has been applied, either zero, one or two jets may be reconstructed.

where the infinite series have been resummed into an exponential form. The first term in the exponent is commonly referred to as the leading logarithmic term, with the second being required in order to reproduce next-to-leading logarithms. This reorganization of the perturbative expansion is especially useful when the product  $\alpha_s L$  is large, for instance when the logarithm is a ratio of two physical scales that are very different such as  $\log(m_H/m_b)$ . This exponential form is the basis of all orders predictions and can be interpreted in terms of Sudakov probabilities, both subjects that we will return to in later discussions.

It is instructive to recast the discussion of the total  $W$  cross section in these terms, where the calculation is decomposed into components that each contain a given number of jets:

$$\sigma_W = \sigma_{W+0j} + \sigma_{W+1j} + \sigma_{W+2j} + \sigma_{W+3j} + \dots \quad (27)$$

Now, as in (25), we can further write out each contribution as an expansion in powers of  $\alpha_s$  and logarithms,

$$\begin{aligned} \sigma_{W+0j} &= a_0 + \alpha_s(a_{12}L^2 + a_{11}L + a_{10}) \\ &\quad + \alpha_s^2(a_{24}L^4 + a_{23}L^3 + a_{22}L^2 + a_{21}L + a_{20}) + \dots \\ \sigma_{W+1j} &= \alpha_s(b_{12}L^2 + b_{11}L + b_{10}) \\ &\quad + \alpha_s^2(b_{24}L^4 + b_{23}L^3 + b_{22}L^2 + b_{21}L + b_{20}) + \dots \\ \sigma_{W+2j} &= \dots \end{aligned} \quad (28)$$

As the jet definitions change, the size of the logarithms shuffle the contributions from one term to another. For example, as the jet cone size is decreased the logarithm  $L$  increases. As a result, the average jet multiplicity goes up and terms in (27) that represent relatively higher multiplicities will become more important.

This is illustrated in Figure 10. Such a configuration may be reconstructed as an event containing up to two jets, depending upon the jet definition and the momenta

of the partons. The matrix elements for this process contain terms proportional to  $\alpha_s \log(p_{T,3}/p_{T,4})$  and  $\alpha_s \log(1/\Delta R_{34})$  which is the reason that minimum values for the transverse energy and separation must be imposed. We shall see later that this is not the case at next-to-leading order.

Finally, we note that although the decomposition in (27) introduces quantities which are dependent upon the jet definition, we can recover results that are independent of these parameters by simply summing up the terms in the expansion that enter at the same order of perturbation theory. In this way, we just recover the perturbative expansion of the total cross section,

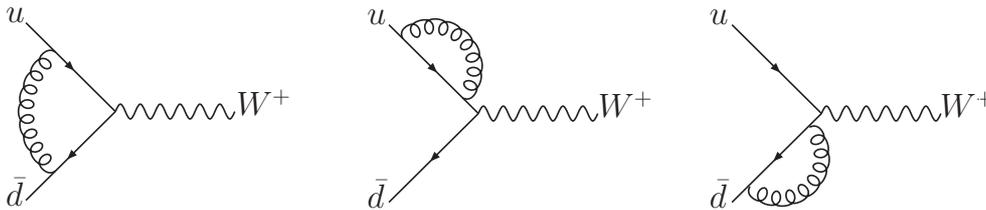
$$\begin{aligned}\sigma_{LO} &= a_0 \\ \sigma_{NLO} &= \alpha_s \left( a_{12}L^2 + a_{11}L + a_{10} + b_{12}L^2 + b_{11}L + b_{10} \right) .\end{aligned}$$

*3.1.3. Leading order tools* Once suitable cuts have been applied, as we have discussed extensively above, leading order cross sections can be calculated using a number of computer programs.

There is a wide range of programs available, most notably ALPGEN [12, 13], the COMPHEP package [14, 15] and MADGRAPH [16, 17]. All of these programs implement the calculation of the diagrams numerically and provide a suitable phase space over which they can be integrated. ALPGEN uses an approach which is not based on a traditional Feynman diagram evaluation [18], whereas the other two programs rely on more conventional methods such as the helicity amplitudes evaluation of HELAS [19] in MADGRAPH.

Although in principle these programs can be used to calculate any tree-level prediction, in practice the complexity of the process that may be studied is limited by the number of particles that is produced in the final state. This is largely due to the factorial growth in the number of Feynman diagrams that must be calculated. Even in approaches which do not rely directly on the Feynman diagrams, the growth is still as a power of the number of particles. For processes which involve a large number of quarks and gluons, as is the case when attempting to describe a multi-jet final state at a hadron collider such as the Tevatron or the LHC, an additional concern is the calculation of colour matrices which appear as coefficients in the amplitudes [20].

In many cases, such as in the calculation of amplitudes representing multiple gluon scattering, the final result is remarkably simple. Motivated by such results, the last couple of years has seen remarkable progress in the development of new approaches to QCD tree-level calculations. Some of the structure behind the amplitudes can be understood by transforming to “twistor space” [21], in which amplitudes are represented by intersecting lines. This idea can be taken further with the introduction of “MHV” rules [22], which use the simplest (maximally helicity-violating, or MHV) amplitudes as the building blocks of more complicated ones. Although these rules at first only applied to amplitudes containing gluons, they were soon extended to cases of more general interest at hadron colliders [23, 24, 25, 26, 27, 28]. Even more recently,



**Figure 11.** Virtual diagrams included in the next-to-leading order corrections to Drell-Yan production of a  $W$  at hadron colliders.

further simplification of amplitudes has been obtained by using “on-shell recursion relations” [29, 30]. As well as providing very compact expressions, this approach has the advantage of being both easily proven and readily extendible to processes involving fermions and vector bosons.

### 3.2. Next-to-leading order calculations

Although lowest order calculations can in general describe broad features of a particular process and provide the first estimate of its cross section, in many cases this approximation is insufficient. The inherent uncertainty in a lowest order calculation derives from its dependence on the unphysical renormalization and factorization scales, which is often large. In addition, some processes may contain large logarithms that need to be resummed or extra partonic processes may contribute only when going beyond the first approximation. Thus, in order to compare with predictions that have smaller theoretical uncertainties, next-to-leading order (NLO) calculations are imperative for experimental analyses in Run II of the Tevatron and at the LHC.

*3.2.1. Virtual and real radiation contributions* A next-to-leading order QCD calculation requires the consideration of all diagrams that contribute an additional strong coupling factor,  $\alpha_s$ . These diagrams are obtained from the lowest order ones by adding additional quarks and gluons and they can be divided into two categories, virtual (or loop) contributions and the real radiation component. We shall illustrate this by considering the next-to-leading order corrections to Drell-Yan production at a hadron collider. The virtual diagrams for this process are shown in Figure 11 whilst the real diagrams are exactly the ones that enter the  $W + 1$  jet calculation (in Figure 5).

Let us first consider the virtual contributions. In order to evaluate the diagrams in Figure 11, it is necessary to introduce an additional loop momentum  $\ell$  which circulates around the loop in each diagram and is unconstrained. To complete the evaluation of these diagrams, one must therefore integrate over the momentum  $\ell$ . However, the resulting contribution is not finite but contains infrared divergences – in the same way that the diagrams of Figure 5 contain infrared (soft and collinear) singularities. By

isolating the singularities appropriately, one can see that the divergences that appear in each contribution are equal, but of opposite sign. The fact that the sum is finite is a demonstration of the theorems of Bloch and Nordsieck [31] and Kinoshita, Lee and Nauenberg [32, 33], which guarantee that this is the case at all orders in perturbation theory and for any number of final state particles.

The real contribution consists of the diagrams in Figure 5, together with a quark-gluon scattering piece that can be obtained from these diagrams by interchanging the gluon in the final state with a quark (or antiquark) in the initial state. As discussed in the previous section, the quark-antiquark matrix elements contain a singularity as the gluon transverse momentum vanishes.

In our NLO calculation we want to carefully regulate and then isolate these singularities in order to extend the treatment down to zero transverse momentum. The most common method to regulate the singularities is dimensional regularization. In this approach, the number of dimensions is continued to  $D = 4 - 2\epsilon$  so that in intermediate stages the singularities appear as single and double poles in  $\epsilon$ . After they have cancelled, the limit  $D \rightarrow 4$  can be safely taken. Within this scheme, the cancellation of divergences between real and virtual terms can be seen schematically by consideration of a toy calculation [35],

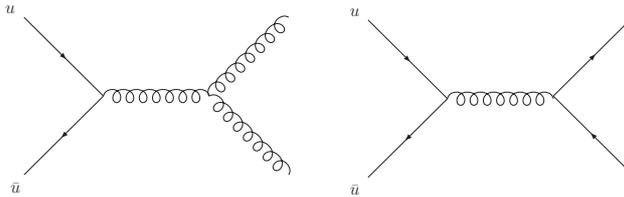
$$\mathcal{I} = \lim_{\epsilon \rightarrow 0} \left( \int_0^1 \frac{dx}{x} x^\epsilon \mathcal{M}(x) - \frac{1}{\epsilon} \mathcal{M}(0) \right). \quad (29)$$

Here,  $\mathcal{M}(x)$  represents the real radiation matrix elements which are integrated over the extra phase space of the gluon emission, which contains a factor  $x^\epsilon$ .  $x$  represents a kinematic invariant which vanishes as the gluon becomes unresolved. The second term is representative of the virtual contribution, which contains an explicit pole,  $1/\epsilon$ , multiplying the lowest order matrix elements,  $\mathcal{M}(0)$ .

Two main techniques have been developed for isolating the singularities, which are commonly referred to as the subtraction method [36, 37, 38, 39] and phase-space slicing [40, 41]. For the sake of illustration, we shall consider only the subtraction method. In this approach, one explicitly adds and subtracts a divergent term such that the new real radiation integral is manifestly finite. In the toy integral this corresponds to,

$$\begin{aligned} \mathcal{I} &= \lim_{\epsilon \rightarrow 0} \left( \int_0^1 \frac{dx}{x} x^\epsilon [\mathcal{M}(x) - \mathcal{M}(0)] + \mathcal{M}(0) \int_0^1 \frac{dx}{x} x^\epsilon - \frac{1}{\epsilon} \mathcal{M}(0) \right) \\ &= \int_0^1 \frac{dx}{x} [\mathcal{M}(x) - \mathcal{M}(0)]. \end{aligned} \quad (30)$$

This idea can be generalized in order to render finite the real radiation contribution to any process, with a separate counter-term for each singular region of phase space. Processes with a complicated phase space, such as  $W+2$  jet production, can end up with a large number of counterterms. NLO calculations are often set up to generate cross sections by histogramming “events” generated with the relevant matrix elements. Such events can not be directly interfaced to parton shower programs as the presence



**Figure 12.** The leading order diagrams representing inclusive jet production from a quark anti-quark initial state.

of virtual corrections means that many of the events will have (often large) negative weights. Only the total sum of events over all relevant subprocesses will lead to a physically meaningful cross section.

The inclusion of real radiation diagrams in a NLO calculation extends the range of predictions that may be described by a lowest order calculation. For instance, in the example above the  $W$  boson is produced with zero transverse momentum at lowest order and only acquires a finite  $p_T$  at NLO. Even then, the  $W$  transverse momentum is exactly balanced by that of a single parton. In a real event, the  $W$   $p_T$  is balanced by the sum of several jet transverse momenta. In a fixed order calculation, these contributions would be included by moving to even higher orders so that, for instance, configurations where the  $W$  transverse momentum is balanced by two jets enter at NNLO. Although this feature is clear for the  $p_T$  distribution of the  $W$ , the same argument applies for other distributions and for more complex processes.

*3.2.2. Scale dependence* One of the benefits of performing a calculation to higher order in perturbation theory is the reduction of the dependence of related predictions on the unphysical renormalization ( $\mu_R$ ) and factorization scales ( $\mu_F$ ). This can be demonstrated by considering inclusive jet production from a quark anti-quark initial state [6], which is represented by the lowest order diagrams shown in Figure 12. This is a simplification of the full calculation, but is the dominant contribution when the typical jet transverse momentum is large.

For this process, we can write the lowest order prediction for the single jet inclusive distribution as,

$$\frac{d\sigma}{dE_T} = \alpha_s^2(\mu_R) \sigma_0 \otimes f_q(\mu_F) \otimes f_{\bar{q}}(\mu_F), \quad (31)$$

where  $\sigma_0$  represents the lowest order partonic cross section calculated from the diagrams of Figure 12 and  $f_i(\mu_F)$  is the parton distribution function for a parton  $i$ . Similarly, after including the next-to-leading order corrections, the prediction can be written as,

$$\frac{d\sigma}{dE_T} = \left[ \alpha_s^2(\mu_R) \sigma_0 + \alpha_s^3(\mu_R) \left( \sigma_1 + 2b_0 \log(\mu_R/E_T) \sigma_0 - 2P_{qq} \log(\mu_F/E_T) \sigma_0 \right) \right] \otimes f_q(\mu_F) \otimes f_{\bar{q}}(\mu_F). \quad (32)$$

In this expression the logarithms that explicitly involve the renormalization and factorization scales have been exposed. The remainder of the  $\mathcal{O}(\alpha_s^3)$  corrections lie in the function  $\sigma_1$ .

From this expression, the sensitivity of the distribution to the renormalization scale is easily calculated using,

$$\mu_R \frac{\partial \alpha_s(\mu_R)}{\partial \mu_R} = -b_0 \alpha_s^2(\mu_R) - b_1 \alpha_s^3(\mu_R) + \mathcal{O}(\alpha_s^4), \quad (33)$$

where the two leading coefficients in the beta-function,  $b_0$  and  $b_1$ , are given by  $b_0 = (33 - 2n_f)/6\pi$ ,  $b_1 = (102 - 38n_f/3)/8\pi^2$ . The contributions from the first and third terms in (32) cancel and the result vanishes, up to  $\mathcal{O}(\alpha_s^4)$ .

In a similar fashion, the factorization scale dependence can be calculated using the non-singlet DGLAP equation,

$$\mu_F \frac{\partial f_i(\mu_F)}{\partial \mu_F} = \alpha_s(\mu_F) P_{qq} \otimes f_i(\mu_F). \quad (34)$$

This time, the partial derivative of each parton distribution function, multiplied by the first term in (32) cancels with the final term. Thus, once again, the only remaining terms are of order  $\alpha_s^4$ .

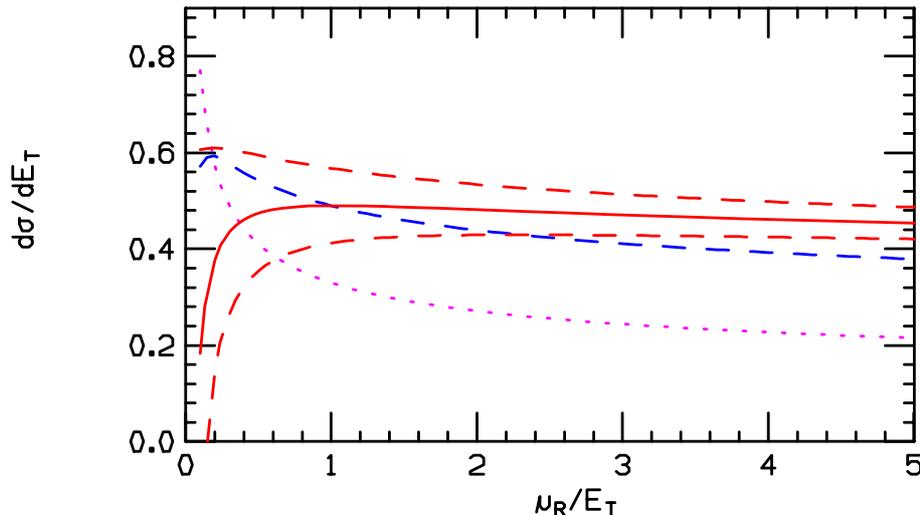
This is a generic feature of a next-to-leading order calculation. An observable that is predicted to order  $\alpha_s^n$  is independent of the choice of either renormalization or factorization scale, up to the next higher order in  $\alpha_s$ .

This discussion can be made more concrete by inserting numerical results into the formulae indicated above. For simplicity, we will consider only the renormalization scaled dependence, with the factorization scale held fixed at  $\mu_F = E_T$ . In this case it is simple to extend (32) one higher order in  $\alpha_s$  [46],

$$\begin{aligned} \frac{d\sigma}{dE_T} = & \left[ \alpha_s^2(\mu_R) \sigma_0 + \alpha_s^3(\mu_R) (\sigma_1 + 2b_0 L \sigma_0) \right. \\ & \left. + \alpha_s^4(\mu_R) (\sigma_2 + 3b_0 L \sigma_1 + (3b_0^2 L^2 + 2b_1 L) \sigma_0) \right] \otimes f_q(\mu_F) \otimes f_{\bar{q}}(\mu_F), \end{aligned}$$

where the logarithm is abbreviated as  $L \equiv \log(\mu_R/E_T)$ . For a realistic example at the Tevatron Run I,  $\sigma_0 = 24.4$  and  $\sigma_1 = 101.5$ . With these values the LO and NLO scale dependence can be calculated; the result is shown in Figure 13. At the moment the value of  $\sigma_2$  is unknown (see the NNLO section below). However, a range of predictions based on plausible values that it could take are also shown in the figure,  $\sigma_2 = 0$  (solid) and  $\sigma_2 = \pm \sigma_1^2/\sigma_0$  (dashed). It is clear that the renormalization scale dependence is reduced when going from LO and NLO and will become smaller still at NNLO.

Although Figure 13 is representative of the situation found at NLO, the exact details depend upon the kinematics of the process under study and on choices such as the running of  $\alpha_s$  and the PDF's used. Of particular interest are the positions on the NLO curve which correspond to often-used scale choices. Due to the structure of (32) there will always be a peak in the NLO curve, around which the scale dependence is minimized. The scale at which this peak occurs is often favoured as a choice. For



**Figure 13.** The single jet inclusive distribution at  $E_T = 100$  GeV, appropriate for Run I of the Tevatron. Theoretical predictions are shown at LO (dotted magenta), NLO (dashed blue) and NNLO (red). Since the full NNLO calculation is not complete, three plausible possibilities are shown.

example, for inclusive jet production at the Tevatron, a scale of  $E_T^{jet}/2$  is usually chosen. This is near the peak of the NLO cross section for many kinematic regions. It is also usually near the scale at which the LO and NLO curves cross, leading to a  $K$ -factor (see below) near unity. A different strategy is to pick the scale at which the LO and NLO predictions coincide. By definition, the  $K$ -factor is unity with this scheme. In the case of inclusive jet production, one can see that these two scales are rather different – around 10 GeV and 200 GeV respectively. Finally, a rather different motivation comes from the consideration of a “physical” scale for the process. For instance, in the case of  $W$  production, one might think that a natural scale is the  $W$  mass. Clearly, these three typical methods for choosing the scale at which cross sections should be calculated do not in general agree. If they do, one may view it as a sign that the perturbative expansion is well-behaved. If they do not agree then the range of predictions provided by the different choices can be ascribed to the “theoretical error” on the calculation.

*3.2.3. The NLO  $K$ -factor* The  $K$ -factor for a given process is a useful shorthand which encapsulates the strength of the NLO corrections to the lowest order cross section. It is calculated by simply taking the ratio of the NLO to the LO cross section. In principle, the  $K$ -factor may be very different for various kinematic regions of the same process. In practice, the  $K$ -factor often varies slowly and may be approximated as one number.

However, when referring to a given  $K$ -factor one must take care to consider the cross section predictions that entered its calculation. For instance, the ratio can depend quite strongly on the PDF’s that were used in both the LO and NLO evaluations. It is by now standard practice to use a NLO PDF (for instance, the CTEQ6M set) in evaluating the NLO cross section and a LO PDF (such as CTEQ6L) in the lowest order

**Table 1.**  $K$ -factors for various processes at the Tevatron and the LHC, calculated using a selection of input parameters. In all cases, the CTEQ6M PDF set is used at NLO.  $\mathcal{K}$  uses the CTEQ6L1 set at leading order, whilst  $\mathcal{K}'$  uses the same set, CTEQ6M, as at NLO. Jets satisfy the requirements  $p_T > 15$  GeV and  $|\eta| < 2.5$  (5.0) at the Tevatron (LHC). In the  $W + 2$  jet process the jets are separated by  $\Delta R > 0.52$ , whilst the weak boson fusion (WBF) calculations are performed for a Higgs of mass 120 GeV.

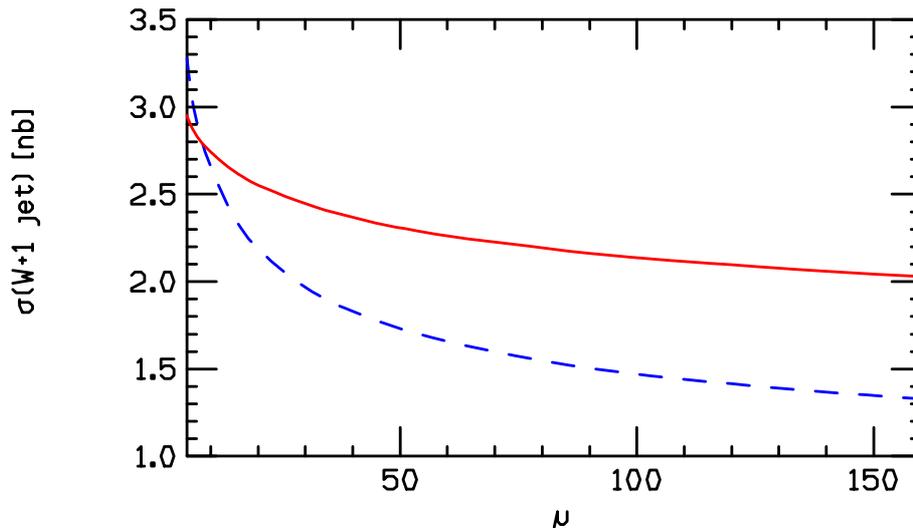
| Process       | Typical scales |                                    | Tevatron K-factor    |                      |                       | LHC K-factor         |                      |                       |
|---------------|----------------|------------------------------------|----------------------|----------------------|-----------------------|----------------------|----------------------|-----------------------|
|               | $\mu_0$        | $\mu_1$                            | $\mathcal{K}(\mu_0)$ | $\mathcal{K}(\mu_1)$ | $\mathcal{K}'(\mu_0)$ | $\mathcal{K}(\mu_0)$ | $\mathcal{K}(\mu_1)$ | $\mathcal{K}'(\mu_0)$ |
| $W$           | $m_W$          | $2m_W$                             | 1.33                 | 1.31                 | 1.21                  | 1.15                 | 1.05                 | 1.15                  |
| $W + 1$ jet   | $m_W$          | $\langle p_T^{\text{jet}} \rangle$ | 1.42                 | 1.20                 | 1.43                  | 1.21                 | 1.32                 | 1.42                  |
| $W + 2$ jets  | $m_W$          | $\langle p_T^{\text{jet}} \rangle$ | 1.16                 | 0.91                 | 1.29                  | 0.89                 | 0.88                 | 1.10                  |
| $t\bar{t}$    | $m_t$          | $2m_t$                             | 1.08                 | 1.31                 | 1.24                  | 1.40                 | 1.59                 | 1.48                  |
| $b\bar{b}$    | $m_b$          | $2m_b$                             | 1.20                 | 1.21                 | 2.10                  | 0.98                 | 0.84                 | 2.51                  |
| Higgs via WBF | $m_H$          | $\langle p_T^{\text{jet}} \rangle$ | 1.07                 | 0.97                 | 1.07                  | 1.23                 | 1.34                 | 1.09                  |

calculation. Sometimes this is not the case, instead the same PDF set may be used for both predictions. Of course, if one wants to estimate the NLO effects on a lowest order cross section, one should take care to match the appropriate  $K$ -factor.

A further complication is caused by the fact that the  $K$ -factor can depend quite strongly on the region of phase space that is being studied. The  $K$ -factor which is appropriate for the total cross section of a given process may be quite different from the one when stringent analysis cuts are applied. For processes in which basic cuts must be applied in order to obtain a finite cross section, the  $K$ -factor again depends upon the values of those cuts. Lastly, of course, as can be seen from Figure 13 the  $K$ -factor depends very strongly upon the renormalization and factorization scales at which it is evaluated. A  $K$ -factor can be less than, equal to, or greater than 1, depending on all of the factors described above.

As examples, in Table 1 we show the  $K$ -factors that have been obtained for a few interesting processes at the Tevatron and the LHC. In each case the value of the  $K$ -factor is compared at two often-used scale choices. For comparison, we also note the  $K$ -factor that is obtained when using the same (CTEQ6M) PDF set at leading order and at NLO. In general, the difference when using CTEQ6L1 and CTEQ6M at leading order is not great. However, for the case of bottom production, the combination of the large difference in  $\alpha_s(m_b)$  and the gluon distribution at small  $x$ , result in very different  $K$ -factors. The values  $\mathcal{K}'$  may, for instance, be useful in performing a NLO normalization of parton shower predictions, as we shall discuss in later sections.

Such  $K$ -factors can be used as estimators for the NLO corrections for the listed processes in situations where only the leading order cross sections are available (for instance, when using a parton shower prediction). Note that, for the case of  $W +$  jet production, we have two relevant scales for the hard scattering process:  $m_W$  and the minimum allowed  $p_T^{\text{jet}}$ . If this threshold is quite low, as is the case in most studies at



**Figure 14.** The LO (dashed) and NLO (solid) scale variation of the  $W + 1$  jet cross section at the Tevatron, using the same inputs as in Table 1.

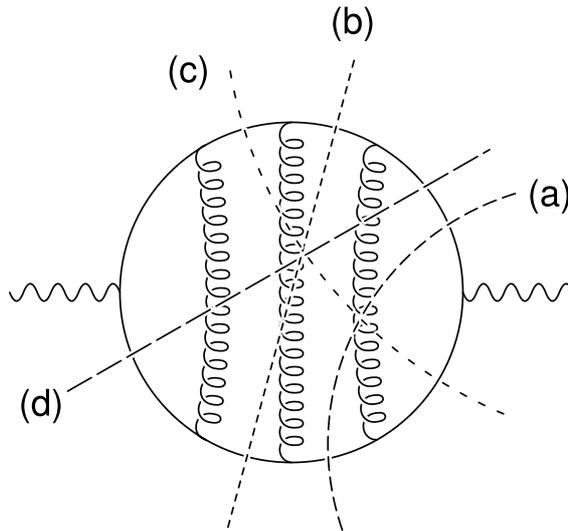
the Tevatron, these scales are quite different. Thus, there can be a fairly large variation in the size of the predicted cross section even at NLO, as illustrated in Figure 14. In the leading order calculation, the cross section varies by about a factor of 2.5 over the range of scales shown. Although this variation is reduced considerably at NLO, the cross section still increases by about 40% when moving from the highest scale shown to the lowest.

### 3.3. NNLO

With all the advantages of NLO, it is only natural to consider going deeper into the perturbative expansion. In the same sense that one only gains a reliable prediction of an observable at NLO, the first meaningful estimate of the theoretical error comes at NNLO. Further reduction of scale uncertainties is expected and, as we shall see, in cases where NLO corrections are large, it is a chance to check the convergence of the perturbative expansion.

With these sorts of justifications in mind, a recent goal of theoretical effort has been the calculation of the 3 jet rate in  $e^+e^-$  annihilation to NNLO. Together with data from LEP and the SLC, this could be used to reduce the error on the measurement of  $\alpha_s(M_Z)$  to a couple of percent. However, the ingredients of a NNLO calculation are both more numerous and more complicated than those entering at NLO. The different contributions can best be understood by considering all possible cuts of a relevant  $\mathcal{O}(\alpha_s^3)$  three-loop diagram, as shown in Figure 15.

The first contribution, represented by cut (a) in Figure 15, corresponds to 2-loop 3-parton diagrams. As the result of much innovative work over the last few years, this contribution is now known (see, for example, references [4]–[6] of [42]). The contribution

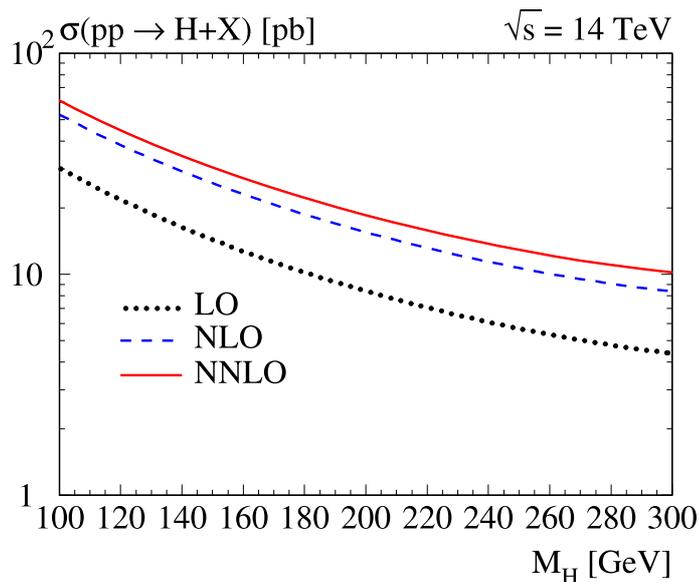


**Figure 15.** A three-loop diagram which, when cut in all possible ways, shows the partonic contributions that must be calculated to perform a NNLO prediction of the  $e^+e^- \rightarrow 3$  jets rate. A description of the contribution represented by each of the cuts (a)-(d) can be found in the text.

labelled by (b) corresponds to the square of the 1-loop 3-parton matrix elements, the same ones which appear (interfered with tree-level) in the NLO calculation [36, 40]. The third contribution (c) also contains 1-loop matrix elements, this time with 4 partons in the final state, one of which is unresolved. As in a NLO calculation, when one parton is unresolved this contribution diverges and a method must be developed to extract all singularities. Both these matrix elements [43, 44, 45, 47] and such methods (for instance, [48] and references therein) have been known for some time. The final contribution (d) involves only tree-level 5-parton matrix elements, but has so far proven the stumbling block to a complete NNLO 3-jet calculation. This piece contains two unresolved partons and, just as before, this gives rise singularities that must be subtracted. However, at present no general procedure for doing this exists and instead calculations can only be performed on a case-by-case basis. Quite recently a method has been developed for  $e^+e^- \rightarrow$  jets calculations which has been used to calculate the doubly-unresolved sub-leading in  $N_c$  contribution to the 3-jet rate [42]. Such progress bodes well for the completion both of this calculation and the closely related 2-jet rate at hadron colliders §.

The calculation that we have described represents the current frontier of NNLO predictions. For slightly simpler  $2 \rightarrow 1$  and  $2 \rightarrow 2$  processes, NNLO results are already available. The total inclusive cross section for the Drell-Yan process, production of a lepton pair by a  $W$  or  $Z$  in a hadronic collision, has long been known to NNLO

§ A consistent NNLO calculation at a hadron collider also requires parton densities evolved at the same order, which is now possible thanks to the calculation of the QCD 3-loop splitting functions [60, 61]. The differences between NLO and NNLO parton densities are reasonably small, though.



**Figure 16.** The inclusive Higgs cross section as a function of the Higgs boson mass.

accuracy [62]. In recent years the inclusive Higgs boson cross section, which is also a one-scale problem in the limit of large  $m_t$ , has also been computed at NNLO [63, 64]. For both these processes, the NLO corrections had already been observed to be large and the inclusion of the NNLO terms only provided a small further increase, thus stabilizing the perturbative expansion of these cross sections. This is illustrated in Figure 16, which shows the inclusive Higgs cross section at the LHC at each order of perturbation theory.

The above calculations have now been extended to include rapidity cuts on the leptons in the Drell-Yan process, in order to be more applicable for studies at the LHC [65]. These calculations extend the method used in [64], which uses an ingenious trick to bypass the problems associated with doubly-unresolved radiation that we have described above. In this approach, the phase space integrals are related to 2-loop integrals that are known and whose calculation can be automated. In this way, NNLO predictions can be provided for simple quantities such as rapidities. Further developments now allow for the introduction of generic cuts, paving the way for more detailed experimental analyses [66].

### 3.4. All orders approaches

Rather than systematically calculating to higher and higher orders in the perturbative expansion of a given observable, a number of different “all-orders” approaches are also commonly used to describe the phenomena observed at high-energy colliders. These alternative descriptions are typically most useful under a different set of conditions than a fixed order approach. The merging of such a description with fixed-order calculations, in order to offer the best of both worlds, is of course highly desirable.

Resummation is one such approach, in which the dominant contributions from each order in perturbation theory are singled out and “resummed” by the use of an evolution equation. Near the boundaries of phase space, fixed order predictions break down due to large logarithmic corrections, as we have seen above. A straightforward example is provided by the production of a vector boson at hadron colliders. In this case, two large logarithms can be generated. One is associated with the production of the vector boson close to threshold ( $\hat{s} = Q^2$ ) and takes the form  $\alpha_s^n \log^{2n-1}(1-z)/(1-z)$ , where  $z = Q^2/\hat{s} - 1$ . The other logarithm, as illustrated earlier, is associated with the recoil of the vector boson at very small transverse momenta  $Q_T$ , so that logarithms appear as  $\alpha_s^n \log^{2n-1}(Q^2/Q_T^2)$ , c.f. (18). Various methods for performing these resummations are available [67, 68, 68, 69, 70, 71, 72, 73, 74], with some techniques including both effects at the same time [75, 76, 77]. As we shall see later, the inclusion of such effects is crucial in order to describe data at the Tevatron and to estimate genuine non-perturbative effects. The ResBos program || [78] is a publically available program that provides NLO resummed predictions for processes such as  $W, Z, \gamma\gamma, Higgs$  production at hadron-hadron colliders. Resummation is of course not restricted to the study of these processes alone, with much progress recently in the resummation of event shape variables at hadron colliders (for a recent review, see [79]).

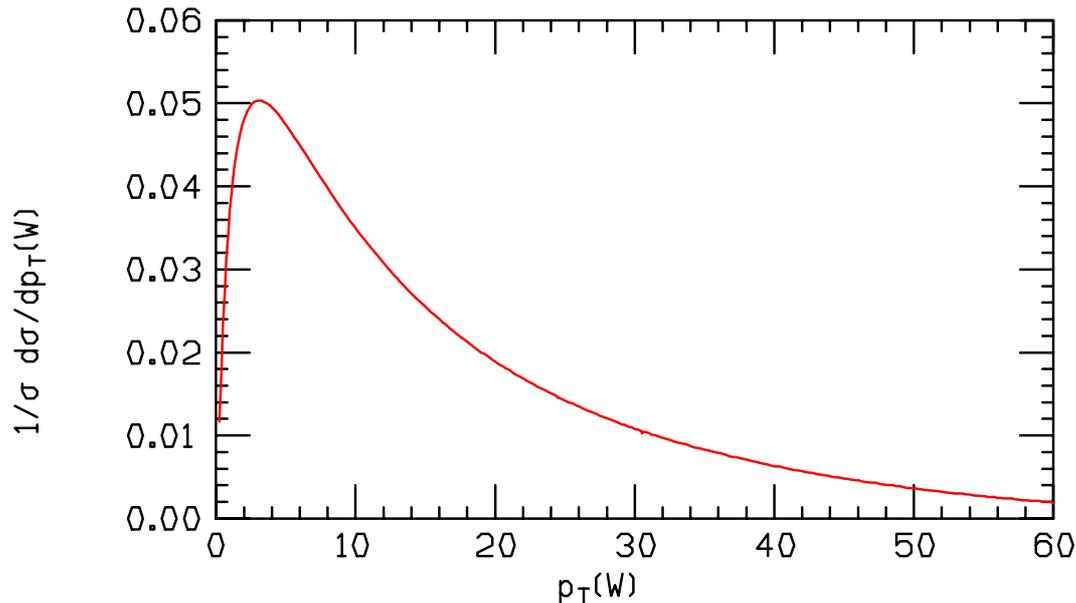
The expression for the  $W$  boson transverse momentum in which the leading logarithms have been resummed to all orders is given by (c.f. (18) and (26)),

$$\frac{d\sigma}{dp_T^2} = \sigma \frac{d}{dp_T^2} \exp\left(-\frac{\alpha_s C_F}{2\pi} \log^2 M_W^2/p_T^2\right). \quad (35)$$

This describes the basic shape for the transverse distribution for  $W$  production, which is shown in Figure 17. Note that in this approximation the  $p_T$  distribution vanishes as  $p_T \rightarrow 0$ , a feature which is not seen experimentally. However this can be explained by the fact that the only configuration included as  $p_T \rightarrow 0$  is the one in which all emitted gluons are soft. In reality (and in a more complete resummed prediction), multiple gluon emissions with a vector sum equal to  $p_T$  contribute and fill in the dip at  $p_T = 0$ .

A different, but related, approach is provided by parton showers. The numerical implementation of a parton shower, for instance in the programs PYTHIA, HERWIG and SHERPA [80], is a common tool used in many current physics analyses. By the use of the parton showering process, a few partons produced in a hard interaction at a high energy scale can be related to partons at an energy scale close to  $\Lambda_{QCD}$ . At this lower energy scale, a universal non-perturbative model can then be used to provide the transition from partons to the hadrons that are observed experimentally. This is possible because the parton showering follows an evolution equation that can be solved either analytically or numerically. The solution of this evolution equation can be cast in the form of a Sudakov form factor, which indicates the probability of evolving from a higher scale to a lower scale without the emission of a gluon greater than a given value. For the case of parton showers from the initial state, the evolution proceeds backwards

|| <http://hep.pa.msu.edu/people/cao/ResBos-A.html>



**Figure 17.** The resummed (leading log)  $W$  boson transverse momentum distribution.

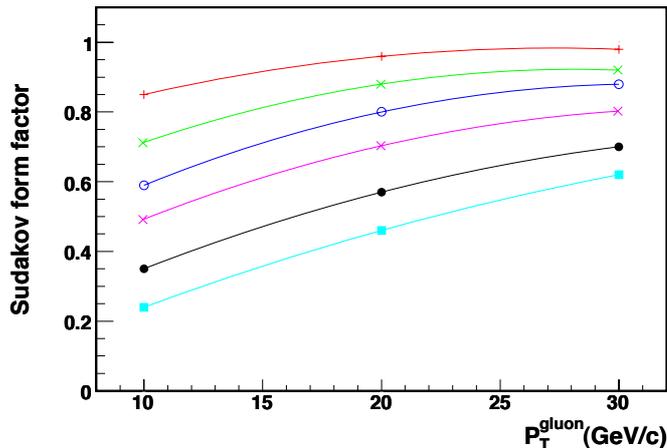
from the hard scale of the process to the cutoff scale, with the Sudakov form factors being weighted by the parton distribution functions at the relevant scales.

In the parton showering process, successive values of an evolution variable  $t$ , a momentum fraction  $z$  and an azimuthal angle  $\phi$  are generated, along with the flavours of the partons emitted during the showering. The evolution variable  $t$  can be the virtuality of the parent parton (as in Pythia versions 6.2 and earlier and in SHERPA),  $E^2(1-\cos\theta)$ , where  $E$  is the energy of the parent parton and  $\theta$  is the opening angle between the two partons (as in Herwig), or the square of the relative transverse momentum of the two partons in the splitting (as in Pythia 6.3). The Herwig evolution variable has angular ordering built in, angular ordering is implicit in the Pythia 6.3 [81] evolution variable, and angular ordering has to be imposed after the fact for the Pythia 6.2 evolution variable. Angular ordering represents an attempt to simulate more precisely those higher order contributions that are enhanced due to soft gluon emission.

Note that with parton showering, we in principle introduce two new scales, one for initial state parton showering and one for the shower in the final state. In the Pythia Monte Carlo, the scale used is most often related to the maximum virtuality in the hard scattering, although a larger scale, such as the total centre-of-mass energy, can also be chosen. The Herwig showering scale is determined by the specific colour flow in the hard process and is related to the mass of the colour string.

We can write an expression for the Sudakov form factor of an initial state parton in the form shown in (36), where  $t$  is the hard scale,  $t_o$  is the cutoff scale and  $P(z)$  is the splitting function for the branching under consideration.

$$\Delta(t) \equiv \exp \left[ - \int_{t_o}^t \frac{dt'}{t'} \int \frac{dz}{z} \frac{\alpha_s}{2\pi} P(z) \frac{f(x/z, t)}{f(x, t)} \right] \quad (36)$$

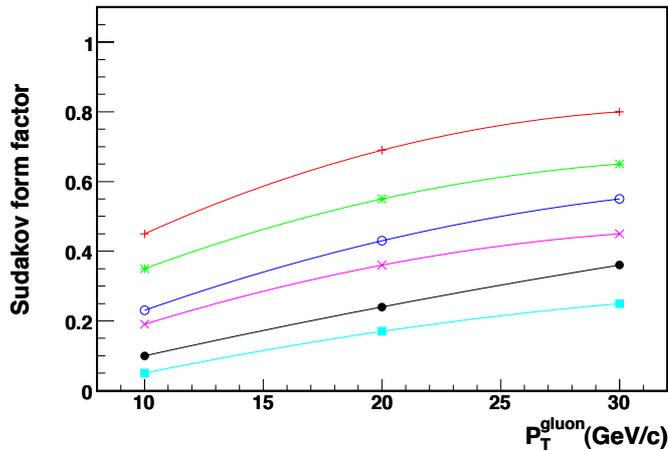


**Figure 18.** The Sudakov form factors for initial state gluons at a hard scale of 100 GeV/c as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton  $x$  values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.

The Sudakov form factor has a similar form for the final state but without the PDF weighting. The introduction of the Sudakov form factor resums all the effects of soft and collinear gluon emission, which leads to well-defined predictions even in these regions. However, this ability comes at a price. Although the soft and collinear regions are logarithmically enhanced and thus the dominant effect, this method omits the non-singular contributions that are typical of large energy, wide angle gluon emission. We shall return to this discussion later.

*3.4.1. Sudakov form factors* As discussed in the previous section, the Sudakov form factor gives the probability for a parton to evolve from a harder scale to a softer scale without emitting a parton harder than some resolution scale, either in the initial state or in the final state. Sudakov form factors form the basis for both parton showering and resummation. Typically, the details of the form factors are buried inside the interior of such programs. It is useful, however, to generate plots of the initial state Sudakov form factors for the kinematic conditions encountered at both the Tevatron and LHC. Such plots indicate the likelihood for the non-radiation of gluons from the initial state partons, and thus conversely for the radiation of at least one such gluon. Thus, they can also serve as a handy reference for the probability of jets from initial state radiation. A Sudakov form factor will depend on: (1) the parton type (quark or gluon), (2) the momentum fraction  $x$  of the initial state parton, (3) the hard and cutoff scales for the process and (4) the resolution scale for the emission. Several examples are discussed below. These plots were generated with the Herwig++ parton shower formalism [82].

In Figure 18 are plotted the Sudakov form factors for the splitting  $g \rightarrow gg$ , at a hard scale of 100 GeV/c, and for several different values of the parton  $x$  value. The form factors are plotted versus the resolution scale for the emitted gluon, which can be

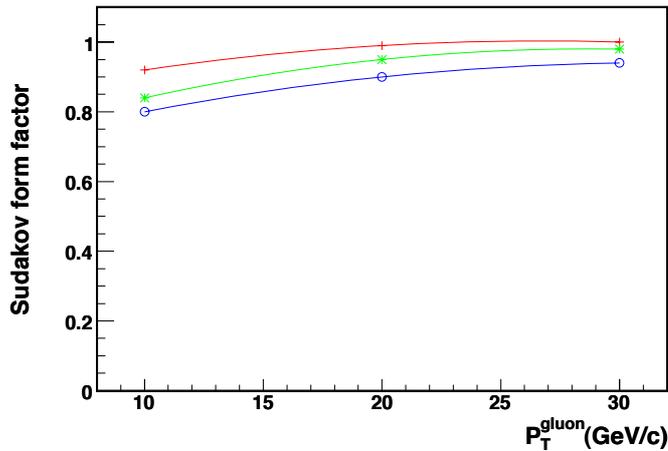


**Figure 19.** The Sudakov form factors for initial state gluons at a hard scale of 500 GeV/c as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton  $x$  values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.

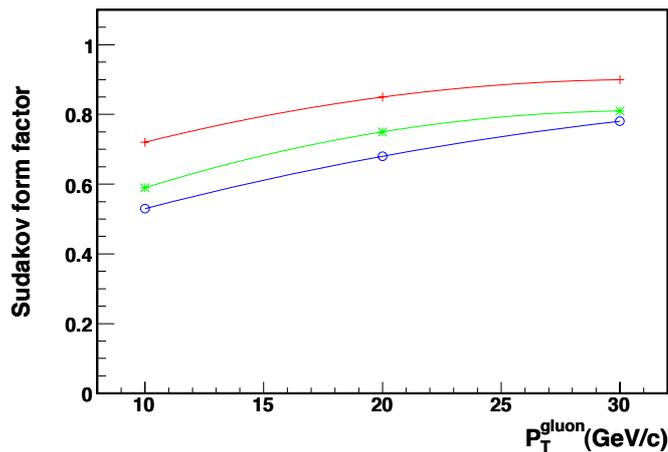
though of roughly as the transverse momentum of the emitted gluon. The probability for no emission decreases as the transverse momentum of the emitted gluon decreases and as the parton  $x$  decreases. The former is fairly obvious; the latter may not be so. The smaller the value of the initial parton momentum fraction, the larger is the possible phase space for gluon emission.

For example, the probability for a gluon with an  $x$  value of 0.03 to evolve from 100 GeV/c down to 10 GeV/c without emitting a gluon of 10 GeV/c or greater can be read off the plot as being 60%; thus the probability for at least one such emission is 40%. This is another example where the probability of emission of a hard gluon is enhanced by a logarithm (in this case the ratio of the hard scale to the resolution scale) compared to the naive expectation of a factor of  $\alpha_s$ . The probability of emission of such a gluon from an initial state gluon on the opposite side of the collision would of course be the same.

In Figure 19, the same Sudakov form factors are plotted but now using a hard scale of 500 GeV/c. The increased probability of a hard gluon emission can be observed. In Figures 20 and 21, the Sudakov form factors are plotted for the hard scales of 100 GeV/c and 500 GeV/c as before, but now for the splitting  $q \rightarrow qg$ . The probability of no emission is larger, due to the smaller colour factor of the initial state quark compared to the gluon. Note that the form factor curves for  $x$  values of less than 0.03 have not been plotted as they would essentially lie on top of the  $x = 0.03$  curve. It is not the smaller colour factor that causes the difference with the gluon but rather the splitting function. The splitting function for  $g \rightarrow gg$  has singularities both as  $z \rightarrow 0$  and as  $z \rightarrow 1$ , while the  $q \rightarrow qg$  has only the  $z \rightarrow 1$  singularity. Thus, for the  $q \rightarrow qg$  splitting, there is not much to gain from decreasing  $x$  on a logarithmic scale, as there is no singularity at  $z = 0$  in the splitting function.



**Figure 20.** The Sudakov form factors for initial state quarks at a hard scale of 100 GeV/c as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton  $x$  values of 0.3, 0.1 and 0.03.



**Figure 21.** The Sudakov form factors for initial state quarks at a hard scale of 500 GeV/c as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton  $x$  values of 0.3, 0.1 and 0.03.

### 3.5. Partons and jet algorithms

In the detectors of experiments at the Tevatron and the LHC, collimated beams of particles are observed. In order to categorize these events, the hadrons are collected into jets using a jet algorithm. To make a comparison with a theoretical calculation of the types we have been discussing, it is necessary to also apply a jet algorithm at the parton level. Ideally, one would like an algorithm which yields similar results at the experimental (hadron) and theoretical (parton) levels. The goal is to characterize the short-distance physics event-by-event, in terms of the jets formed by the algorithm.

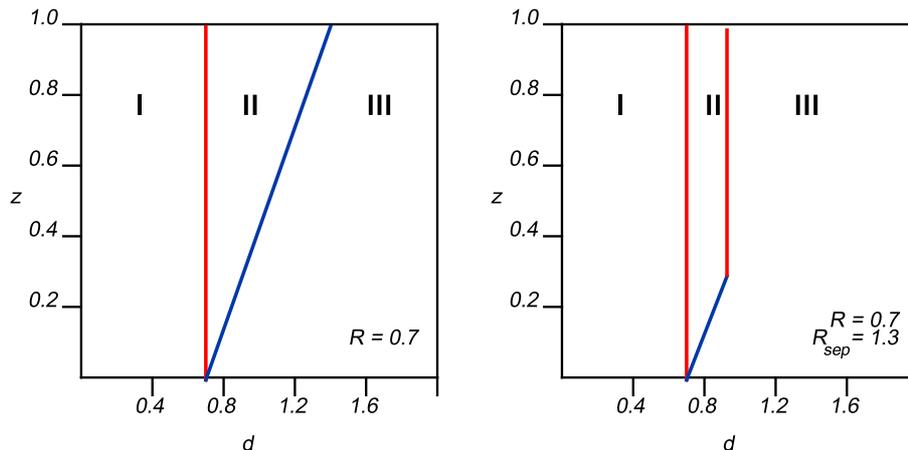
There are two essential stages for any jet algorithm. First, the objects belonging to

a cluster are identified. Second, the kinematic variables defining the jet are calculated from the objects defining the cluster. The two stages are independent. For the latter stage, using the jet algorithms developed for Run 2 at the Tevatron, the jet kinematic properties are defined (using a 4-vector recombination scheme) in terms of:  $p_{jet}$ ,  $p_T^{jet}$ ,  $y^{jet}$  and  $\phi^{jet}$ .

At the experimental or simulated data level, jet algorithms cluster together objects such as particles or energies measured in calorimeter cells. At the theoretical level, partons are clustered. The goal of a jet algorithm is to produce similar results no matter the level it is applied. For a  $2 \rightarrow 2$  LO calculation, a jet consists simply of 1 parton and no jet algorithm is necessary. As more partons are added to a calculation, the complexity of a jet grows and approaches the complexity found either in parton shower Monte Carlos or in data. For all situations in which a jet can consist of more than 1 parton, a completely specified jet algorithm is needed. The clustering algorithms rely on the association of these objects based on transverse momentum (the  $k_T$  algorithm) or angles (the cone algorithm), relative to a jet axis.

For NLO calculations, as for example  $W + 2$  jets, a jet can consist of either 1 or 2 partons. Cone jet algorithms as currently used by the Tevatron experiments require the use of seeds (initial directions for jet cones) as the starting points for jet searches. For a partonic level final state, the seeds are the partons themselves. The Run 2 cone algorithm (midpoint) places additional seeds between stable cones having a separation of less than twice the size of the clustering cones; the use of these additional seeds removes problems with infrared instabilities in theoretical calculations. Without a midpoint seed, a jet could be formed/not formed depending on the presence of a soft gluon between two hard partons; this leads to an undesirable logarithmic dependence of the cross section on the energy of this soft gluon.

With a cone algorithm, two partons are nominally included in the same jet if they are within  $R_{cone}$  of the  $p_T$ -weighted jet centroid, and so within  $\Delta R$  of 1.4 if a cone radius of 0.7 is used. However, it was noted that with the experimental jet algorithms used at the Tevatron, that two jets would not be merged into a single jet if they were separated by a distance greater than 1.3 times the cone radius. Thus, a phenomenological parameter  $R_{sep} = 1.3$  was added to the theoretical prediction; two partons would not be merged into a single jet if they were separated by more than  $R_{sep} \times R_{cone}$  from each other. So, in a parton level calculation having at most 3 partons in the final state, two partons are merged into the same jet if they are within  $R_{cone}$  of the  $p_T$ -weighted jet centroid and within  $R_{sep} \times R_{cone}$  of each other; otherwise the two partons are termed separate jets. Thus, for  $W + 2$  jet production at NLO, the final state can consist of either 2 or 3 partons. The 2 parton final state will always be reconstructed as 2 jets; the 3 parton final state may be reconstructed as either 2 or 3 jets depending on whether the 2 lowest  $p_T$  partons satisfy the clustering criteria described above. Note that for some partonic level programs such as JETRAD [83], the clustering is not performed prior to the evaluation of the matrix element. Thus, the individual transverse momenta of the jets are also not known at this time; only the transverse momentum of the highest  $p_T$



**Figure 22.** The parameter space ( $d, Z$ ) for which two partons will be merged into a single jet.

parton, which by momentum conservation must remain unclustered, is known. For this reason, a renormalization/factorization scale of  $p_T^{max}$  (the  $p_T$  of this parton) is used in calculating the cross section for each jet.

A schematic diagram indicating the regions in which two partons will be included in the same jet is shown in Figure 22 [84]. All partons within  $R_{cone}$  of each other will always be clustered in the same jet. This corresponds to the region labeled I. An ideal cone algorithm acting on data would cluster together only the underlying parton level configurations corresponding to region II and not configurations in region III. However, as will be seen in Section 5, the stochastic character of the parton showering + hadronization process makes such a clean division difficult in either real data or with a parton shower Monte Carlo. Because the matrix element for the emission of an additional real parton has both a collinear and soft pole, configurations in regions II and III with two partons having  $\Delta R$  near 0.7 and  $z$  near 0 will be most heavily populated.

The  $k_T$  algorithm is conceptually simpler at all levels. Two partons (or particles, or energies) in calorimeter towers are combined if their relative transverse momentum is less than a given measure. At the parton level, each parton is considered as a proto-jet. The quantities  $k_{T,i}^2 = P_{T,i}^2$  and  $k_{T,(i,j)} = \min(P_{T,i}^2, P_{T,j}^2) \cdot \Delta R_{i,j}^2 / D^2$  are computed for each parton and each pair of partons respectively.  $P_{T,i}$  is the transverse momentum of the  $i^{th}$  parton,  $\Delta R_{i,j}$  is the distance (in  $y, \phi$  space) between each pair of partons, and  $D$  is a parameter that controls the size of the jet. If the smallest of the above quantities is a  $k_{T,i}$  then that parton becomes a jet; if the smallest quantity is a  $k_{T,(i,j)}$ , then the two partons are combined into a single protojet by summing their four-vector components. In a NLO inclusive jet calculation, the two lowest  $p_T$  partons may be combined into a single jet, and thus the final state can consist of either 2 or 3 jets, as was also true for the case of the cone algorithm.

The assumption we are making above is that the jets of hadrons measured in a collider experiment can be represented by the 1 or 2 partons that comprise a jet at

the NLO theoretical level. That is, the 1 or 2 partons present in a NLO jet effectively represent the many partons produced by a parton shower. For example, an equivalent description of the jet shape is provided by the two types of calculation.

This approximation has been borne out in practice with one remaining correction being necessary. Partons whose trajectories lie inside the jet cone eventually produce hadrons, some of which may land outside the cone due to the fragmentation process. The fragmentation correction takes a particularly simple form. For a cone of radius 0.7, each jet loses approximately 1 GeV due to fragmentation, basically independent of the jet transverse energy. The naïve assumption might be that the energy outside the cone due to fragmentation would rise with the jet energy; however, the jet becomes narrower at higher  $E_T$ , leading to a roughly constant amount of energy in the outermost portions of the jet. As will be described in Section 5, corrections also need to be applied to the data or to the theory to take into account the underlying event energy.

### 3.6. Merging parton showers and fixed order

As we have discussed previously, parton showers provide an excellent description in regions which are dominated by soft and collinear gluon emission. On the other hand, matrix element calculations provide a good description of processes where the partons are energetic and widely separated and, in addition, include the effects of interference between amplitudes with the same external partons. But, on the other hand, the matrix element calculations do not take into account the interference effects in soft and collinear gluon emissions which cannot be resolved, and which lead to a Sudakov suppression of such emissions.

Clearly, a description of a hard interaction which combines the two types of calculations would be preferable. For this combination to take place, there first needs to be a universal formalism that allows the matrix element calculation to “talk” to the parton shower Monte Carlo. Such a universal formalism was crafted during the Les Houches Workshop on Collider Physics in 2001 and the resulting “Les Houches Accord” is in common use [85]. The accord specifies an interface between the matrix element and the parton shower program which provides information on the the parton 4-vectors, the mother-daughter relationships, and the spin/helicities and colour flow. It also points to intermediate particles whose mass should be preserved in the parton showering. All of the details are invisible to the casual user and are intended for the matrix element/parton shower authors.

Some care must be taken however, as a straight addition of the two techniques would lead to double-counting in kinematic regions where the two calculations overlap. There have been many examples where matrix element information has been used to correct the first or the hardest emission in a parton shower. There are also more general techniques that allow matrix element calculations and parton showers to each be used in kinematic regions where they provide the best description of the event properties and that avoid double-counting. One such technique is termed CKKW [86].

With the CKKW technique, the matrix element description is used to describe parton branchings at large angle and/or energy, while the parton shower description is used for the smaller angle, lower energy emissions. The phase space for parton emission is thus divided into two regions, matrix element dominated and parton shower dominated, using a resolution parameter  $d_{ini}$ . The argument of  $\alpha_s$  at all of the vertices is chosen to be equal to the resolution parameter  $d_i$  at which the branching has taken place and Sudakov form factors are inserted on all of the quark and gluon lines to represent the lack of any emissions with a scale larger than  $d_{ini}$  between vertices. The  $d_i$  represent a virtuality or energy scale. Parton showering is used to produce additional emissions at scales less than  $d_{ini}$ . A schematic representation of the CKKW scheme is shown in Figure 23 for the case of  $W + \text{jets}$  production at a hadron-hadron collider. A description of a  $W + 2 \text{ jet}$  event in the NLO formalism is also shown for comparison.

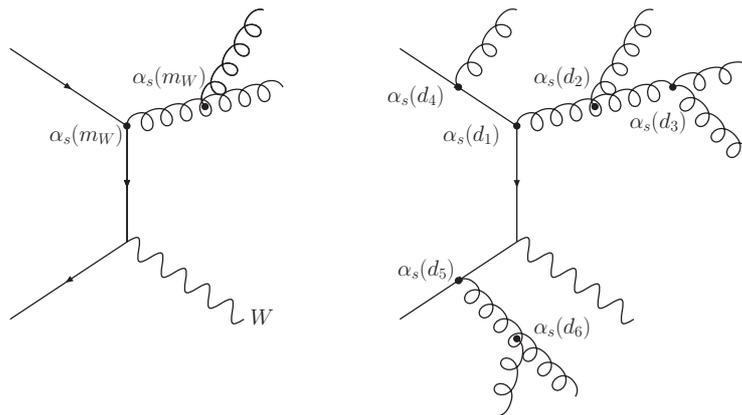
The CKKW procedure provides a matching between the matrix element and parton shower that should be correct to the next-to-leading-logarithm (NLL) level. There are, however, a number of choices that must be made in the matching procedure that do not formally affect the logarithmic behavior but do affect the numerical predictions, on the order of 20-30%. The CKKW procedure gives the right amount of radiation but tends to put some of it in the wrong place with the wrong colour flow. Variations that result from these choices must be considered as part of the systematic error inherent in the CKKW process. This will be discussed further in Section 5.

For the CKKW formalism to work, matrix element information must in principle be available for any number  $n$  of partons in the final state. Practically speaking, having information available for  $n$  up to 4 is sufficient for the description of most events at the Tevatron or LHC. The CKKW formalism is implemented in the parton shower Monte Carlo SHERPA [87] and has also been used for event generation at the Tevatron and LHC using the Mrenna-Richardson formalism [88]. An approximate version of CKKW matching (the mlm approach ¶) is available in ALPGEN 2.0 [13].

### 3.7. Merging NLO calculations and parton showers

A combination of NLO calculations with parton shower Monte Carlos leads to the best of both worlds. The NLO aspect leads to a correct prediction for the rate of the process and also improves the description of the first hard parton emission. The parton shower aspect provides a sensible description of multiple/soft collinear emissions with a final state consisting of hadrons, which can then be input to a detector simulation. In a parton shower interface, a specific subtraction scheme must be implemented to preserve the NLO cross section. As each parton shower Monte Carlo may produce a different real radiation component, the subtraction scheme must necessarily depend on the Monte Carlo program to which the matrix element program is matched. The presence of interference effects with NLO calculations requires that a relatively small fraction ( $\sim 10\%$ ) of events have negative weights (of value  $-1$ ).

¶ ...which one of the authors takes credit for naming.



**Figure 23.** In the NLO formalism, the same scale, proportional to the hardness of the process, is used for each QCD vertex. For the case of the  $W+2$  jet diagram shown above to the left, a scale related to the mass of the  $W$  boson, or to the average transverse momentum of the produced jets, is typically used. The figure to the right shows the results of a simulation using the CKKW formalism. Branchings occur at the vertices with resolution parameters  $d_i$ , where  $d_1 > d_2 \gg d_{ini} > d_3 > d_4 > d_5 > d_6$ . Branchings at the vertices 1-2 are produced with matrix element information while the branchings at vertices 3-6 are produced by the parton shower.

Several groups have worked on the subject to consistently combine partonic NLO calculations with parton showers.

- Collins, Zu [89, 90]
- Frixione, Nason, Webber (MC@NLO) [91, 92, 93]
- Kurihara, Fujimoto, Ishikawa, Kato, Kawabata, Munehisa, Tanaka [94]
- Krämer, Soper [95, 96, 97]
- Nagy, Soper [98, 99]

MC@NLO is the only publically available program that combines NLO calculations with parton showering and hadronization. The Herwig Monte Carlo is used for the latter. The use of a different Monte Carlo, such as Pythia, would require a different subtraction scheme for the NLO matrix elements. The processes included to date are:  $(W, Z, \gamma^*, H, b\bar{b}, t\bar{t}, HW, HZ, WW, WZ, ZZ)$ . Recently, single top hadroproduction has been added to MC@NLO [100]. This is the first implementation of a process that has both initial- and final-state singularities. This allows a more general category of additional processes to be added in the future. Work is proceeding on inclusion of inclusive jet production and WW fusion to Higgs. Adding spin correlations to a process increases the level of difficulty but is important for processes such as single top production.

If, in addition, the CKKW formalism could be used for the description of hard parton emissions, the utility and accuracy of a NLO Monte Carlo could be greatly increased. The merger of these two techniques should be possible in Monte Carlos available by the time of the LHC turn-on.

## 4. Parton distribution functions

### 4.1. Introduction

As discussed previously, the calculation of the production cross sections at hadron colliders for both interesting physics processes and their backgrounds relies upon a knowledge of the distribution of the momentum fraction  $x$  of the partons (quarks and gluons) in a proton in the relevant kinematic range. These parton distribution functions (pdf's) can not be calculated perturbatively but rather are determined by global fits to data from deep inelastic scattering (DIS), Drell-Yan (DY), and jet production at current energy ranges. Two major groups, CTEQ [6] and MRST [5], provide semi-regular updates to the parton distributions when new data and/or theoretical developments become available. In addition, there are also pdf's available from Alekhin [101] and from the two HERA experiments [102, 103, 104, 105]. The newest pdf's, in most cases, provide the most accurate description of the world's data, and should be utilized in preference to older pdf sets.

### 4.2. Processes Involved in Global Analysis Fits

Measurements of deep-inelastic scattering (DIS) structure functions ( $F_2, F_3$ ) in lepton-hadron scattering and of lepton pair production cross sections in hadron-hadron collisions provide the main source on quark distributions  $f_{p \rightarrow q}(x, Q)$  inside hadrons. At leading order, the gluon distribution function  $f_{p \rightarrow g}(x, Q)$  enters directly in hadron-hadron scattering processes with jet final states. Modern global parton distribution fits are carried out to next-to-leading (NLO) order and in some cases next-to-next-to-leading order (NNLO) which allows  $\alpha_s(Q^2)$ ,  $f_{p \rightarrow q}(x, Q)$  and  $f_{p \rightarrow g}(x, Q)$  to all mix and contribute in the theoretical formulae for all processes. Nevertheless, the broad picture described above still holds to some degree in global pdf analyses.

The data from DIS, DY and jet processes utilized in pdf fits cover a wide range in  $x$  and  $Q$ . HERA data (H1 [49]+ZEUS [50]) are predominantly at low  $x$ , while the fixed target DIS [52, 53, 54, 55, 56] and DY [57, 51] data are at higher  $x$ . Collider jet data [58, 59] cover a broad range in  $x$  and  $Q^2$  by themselves and are particularly important in the determination of the high  $x$  gluon distribution. There is considerable overlap, however, among the datasets with the degree of overlap increasing with time as the statistics of the HERA experiments increase. Parton distributions determined at a given  $x$  and  $Q^2$  'feed-down' or evolve to lower  $x$  values at higher  $Q^2$  values. DGLAP-based NLO (and NNLO) pQCD should provide an accurate description of the data (and of the evolution of the parton distributions) over the entire kinematic range present in current global fits. At very low  $x$  and  $Q$ , DGLAP evolution is believed to be no longer applicable and a BFKL description must be used. No clear evidence of BFKL physics is seen in the current range of data; thus all global analyses use conventional DGLAP evolution of pdf's.

Many processes have been calculated to NLO and there is the possibility of including

data from these processes in global fits. For example, the rapidity distributions for  $W^+$ ,  $W^-$  and Z production at the Tevatron and LHC should prove to be very useful in constraining  $u$  and  $d$  sea quarks.

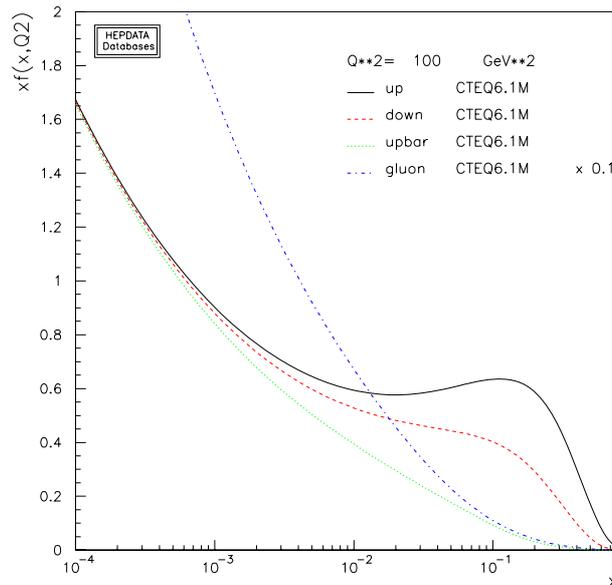
There is a remarkable consistency between the data in the pdf fits and the perturbative QCD theory fit to them. Both the CTEQ and MRST groups use over 2000 data points in their global pdf analyses and the  $\chi^2/\text{DOF}$  for the fit of theory to data is on the order of unity. For most of the data points, the statistical errors are smaller than the systematic errors, so a proper treatment of the systematic errors and their bin-to-bin correlations is important.

The accuracy of the extrapolation to higher  $Q^2$  depends on the accuracy of the original measurement, any uncertainty on  $\alpha_s(Q^2)$  and the accuracy of the evolution code. Most global pdf analyses are carried out at NLO; recently, the DGLAP evolution kernels have been calculated at NNLO [106], allowing a full NNLO evolution to be carried out, and NNLO pdf's calculated in this manner are available [101, 107]. However, not all processes in the global fits, and specifically inclusive jet production, are available at NNLO. Thus, any current NNLO global pdf analyses are still approximate for this reason, but in practice the approximation should work well. Current programs in use by CTEQ and MRST should be able to carry out the evolution using NLO DGLAP to an accuracy of a few percent over the hadron collider kinematic range, except perhaps at very large  $x$  and very small  $x$ . (See the discussion in Section 6 regarding the validity of NLO analysis at the LHC.) The kinematics appropriate for the production of a state of mass  $M$  and rapidity  $y$  at the LHC was shown in Figure 3 in Section 2. For example, to produce a state of mass 100 GeV and rapidity 2 requires partons of  $x$  values 0.05 and 0.001 at a  $Q^2$  value of  $1 \times 10^4 \text{ GeV}^2$ . Also shown in the figure is a view of the kinematic coverage of the fixed target and HERA experiments used in the global pdf fits.

#### 4.3. Parameterizations and Schemes

A global pdf analysis carried out at next-to-leading order needs to be performed in a specific renormalization and factorization scheme. The evolution kernels are in a specific scheme and to maintain consistency, any hard scattering cross section calculations used for the input processes or utilizing the resulting pdf's need to have been implemented in that same renormalization scheme. Almost universally, the  $\overline{MS}$  scheme is used; pdf's are also available in the DIS scheme, a fixed flavour scheme (a la GRV [108]) and several schemes that differ in their specific treatment of the charm quark mass.

It is also possible to use only leading-order matrix element calculations in the global fits which results in leading-order parton distribution functions, which have been made available by both the CTEQ and MRST groups. For many hard matrix elements for processes used in the global analysis, there exist K factors (NLO/LO) significantly different from unity. Thus, one expects there to be noticeable differences of the LO parton distributions from the NLO ones. All global analyses use a generic form for the



**Figure 24.** The CTEQ6.1 parton distribution functions evaluated at a  $Q$  of 10 GeV.

parameterization of both the quark and gluon distributions at some reference value  $Q_o$ :

$$F(x, Q_o) = A_o x^{A_1} (1-x)^{A_2} P(x; A_3, \dots) \quad (37)$$

The reference value  $Q_o$  is usually chosen in the range of 1-2 GeV. The parameter  $A_1$  is associated with small- $x$  Regge behavior while  $A_2$  is associated with large- $x$  valence counting rules..

The first two factors, in general, are not sufficient to describe either quark or gluon distributions. The term  $P(x; A_3, \dots)$  is a suitable chosen smooth function, depending on one or more parameters, that adds more flexibility to the pdf parameterization. In general, both the number of free parameters and the functional form can have an influence on the global fit. The pdf's made available to the world from the global analysis groups can either be in a form where the  $x$  and  $Q^2$  dependence is parameterized, or the pdf's for a given  $x$  and  $Q^2$  range can be interpolated from a grid that is provided, or the grid can be generated given the starting parameters for the pdf's (see the discussion on LHAPDF). All techniques should provide an accuracy on the output pdf distributions on the order of a few percent.

The parton distributions from the CTEQ6.1 pdf's release are plotted in Figure 24 at a  $Q$  value of 10 GeV. The gluon distribution is dominant at  $x$  values of less than 0.01 with the valence quark distributions dominant at higher  $x$ . One of the major influences of the HERA data has been to steepen the the gluon distribution at low  $x$ .

#### 4.4. Uncertainties on Pdf's

In addition to having the best estimates for the values of the pdf's in a given kinematic range, it is also important to understand the allowed range of variation of the pdf's,

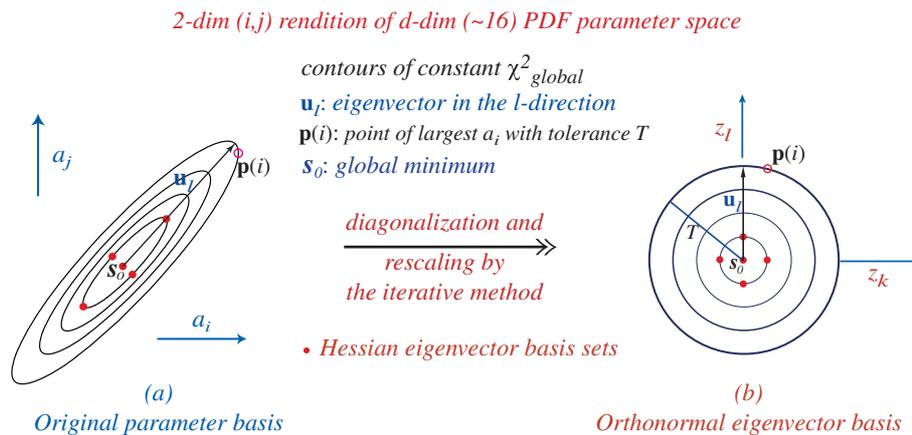
i.e. their uncertainties. A conventional method of estimating parton distribution uncertainties has been to compare different published parton distributions. This is unreliable since most published sets of parton distributions (for example from CTEQ and MRST) adopt similar assumptions and the differences between the sets do not fully explore the uncertainties that actually exist.

The sum of the quark distributions ( $\sum f_{p \rightarrow q}(x, Q) + f_{p \rightarrow \bar{q}}(x, Q)$ ) is, in general, well-determined over a wide range of  $x$  and  $Q^2$ . As stated above, the quark distributions are predominantly determined by the DIS and DY data sets which have large statistics, and systematic errors in the few percent range ( $\pm 3\%$  for  $10^{-4} < x < 0.75$ ). Thus the sum of the quark distributions is basically known to a similar accuracy. The individual quark flavours, though, may have a greater uncertainty than the sum. This can be important, for example, in predicting distributions that depend on specific quark flavours, like the W asymmetry distribution [109] and the W and Z rapidity distributions.

The largest uncertainty of any parton distribution, however, is that on the gluon distribution. The gluon distribution can be determined indirectly at low  $x$  by measuring the scaling violations in the quark distributions, but a direct measurement is necessary at moderate to high  $x$ . The best direct information on the gluon distribution at moderate to high  $x$  comes from jet production at the Tevatron.

There has been a great deal of recent activity on the subject of pdf uncertainties. Two techniques in particular, the Lagrange Multiplier and Hessian techniques, have been used by CTEQ and MRST to estimate pdf uncertainties [110, 111, 112]. The Lagrange Multiplier technique is useful for probing the pdf uncertainty of a given process, such as the W cross section, while the Hessian technique provides a more general framework for estimating the pdf uncertainty for any cross section. In addition, the Hessian technique results in tools more accessible to the general user.

In the Hessian method a large matrix ( $20 \times 20$  for CTEQ,  $15 \times 15$  for MRST), with dimension equal to the number of free parameters in the fit, has to be diagonalized. The result is 20 (15) orthonormal eigenvector directions for CTEQ (MRST) which provide the basis for the determination of the pdf error for any cross section. This process is shown schematically in Figure 25. The eigenvectors are now admixtures of the 20 pdf parameters left free in the global fit. There is a broad range for the eigenvalues, over a factor of one million. The eigenvalues are distributed roughly linearly as  $\ln \epsilon_i$ , where  $\epsilon_i$  is the eigenvalue for the  $i$ -th direction. The larger eigenvalues correspond to directions which are well-determined; for example, eigenvectors 1 and 2 are sensitive primarily to the valence quark distributions at moderate  $x$ , a region where they are well-constrained. The theoretical uncertainty on the determination of the W mass at both the Tevatron and the LHC depends primarily on these 2 eigenvector directions, as W production at the Tevatron proceeds primarily through collisions of valence quarks. The most significant eigenvector directions for determination of the W mass at the LHC correspond to larger eigenvector numbers, which are primarily determined by sea quark distributions. In most cases, the eigenvector can not be directly tied to the behavior of a particular pdf in a specific kinematic region. A longer discussion of the meaning of



**Figure 25.** A schematic representation of the transformation from the pdf parameter basis to the orthonormal eigenvector basis.

the eigenvectors for the CTEQ6.1 pdf analysis can be found at the benchmark website.

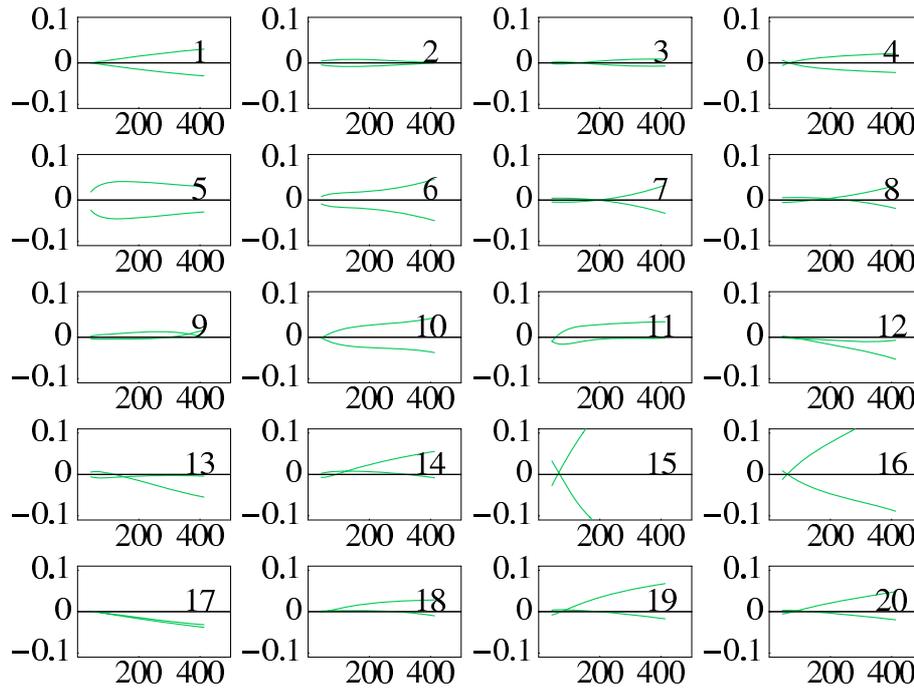
There are two things that can happen when new pdf's (eigenvector directions) are added: a new direction in parameter space can be opened to which some cross sections will be sensitive to (such as eigenvector 15 in the CTEQ6.1 error pdf set which is sensitive to the high  $x$  gluon behavior and thus influences the high  $p_T$  jet cross section at the Tevatron). In this case, a smaller parameter space is an underestimate of the true pdf error since it did not sample a direction important for some physics. In the second case, adding new eigenvectors does not appreciably open new parameter space and the new parameters should not contribute much pdf error to most physics processes (although the error may be redistributed somewhat among the new and old eigenvectors).

Each error pdf results from an excursion along the “+” and “−” directions for each eigenvector. The excursions are symmetric for the larger eigenvalues, but may be asymmetric for the more poorly determined directions. There are 40 pdf's for the CTEQ6.1 error set and 30 for the MRST error set. In Figure 26, the pdf errors are shown in the “+” and “−” directions for the 20 CTEQ eigenvector directions for predictions for inclusive jet production at the Tevatron. The excursions are symmetric for the first 10 eigenvectors but can be asymmetric for the last 10.

In order to calculate the pdf error for an observable, a *Master Equation* should be used:

$$\begin{aligned}
 \Delta X_{max}^+ &= \sqrt{\sum_{i=1}^N [\max(X_i^+ - X_0, X_i^- - X_0, 0)]^2} \\
 \Delta X_{max}^- &= \sqrt{\sum_{i=1}^N [\max(X_0 - X_i^+, X_0 - X_i^-, 0)]^2}
 \end{aligned} \tag{38}$$

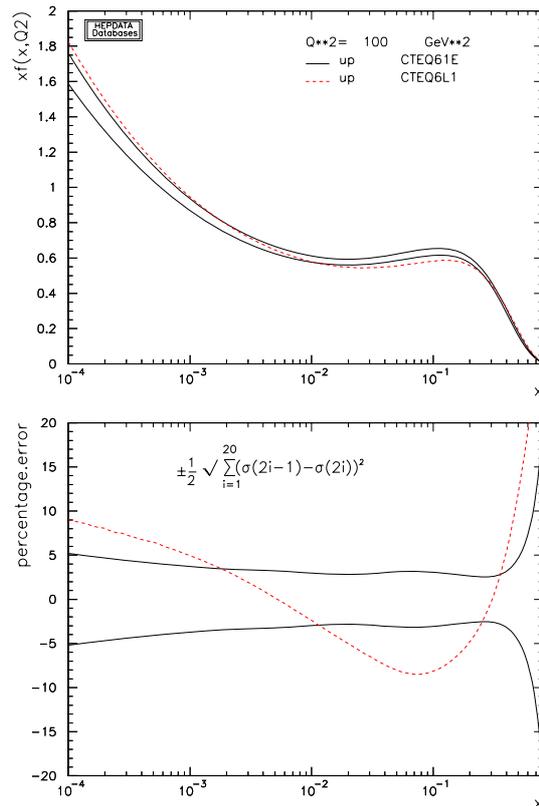
Note that variations in the same direction are always added together. Thus, there may be more or less than 20 terms for either the “+” or “−” directions. There are other versions of the *Master Equation* in current use but the version listed above is the



**Figure 26.** The pdf errors for the CDF inclusive jet cross section in Run 1 for the 20 different eigenvector directions.

“official” recommendation of the authors.

Perhaps the most controversial aspect of pdf uncertainties is the determination of the  $\Delta\chi^2$  excursion from the central fit that is representative of a reasonable error. CTEQ chooses a  $\Delta\chi^2$  value of 100 (corresponding to a 90% CL limit) while MRST uses a value of 50. Thus, in general, the pdf uncertainties for any cross section will be larger for the CTEQ set than for the MRST set. The difference in the criterion indicates the difficulty of treating the error analysis for such a large disparate sample of data in a statistically rigorous and unique manner. But in any case, both groups are in agreement that a  $\chi^2$  excursion of 1 (for a 1- $\sigma$  error) is too low of an estimate in a global pdf fit. The global fits use data sets arising from a number of different processes and different experiments; there is a non-negligible tension between some of the different data sets. Pdf fits performed within a single experiment, or with a single data type may be able to set tighter tolerances. The uncertainties for all predictions should be linearly dependent on the tolerance parameter used; thus, it should be reasonable to scale the uncertainty for an observable from the 90% CL limit provided by the CTEQ/MRST error pdf’s to a one-sigma error by dividing by a factor of 1.6. Such a scaling will be a better approximation for observables more dependent on the low number eigenvectors, where the  $\chi^2$  function is closer to a quadratic form. Note that such a scaling may result in an underestimate of the *true* pdf uncertainty, as the central results for CTEQ and MRST often differ by an amount similar to this one-sigma error. (See the discussion in Section 2 regarding the W cross section predictions at the Tevatron.)



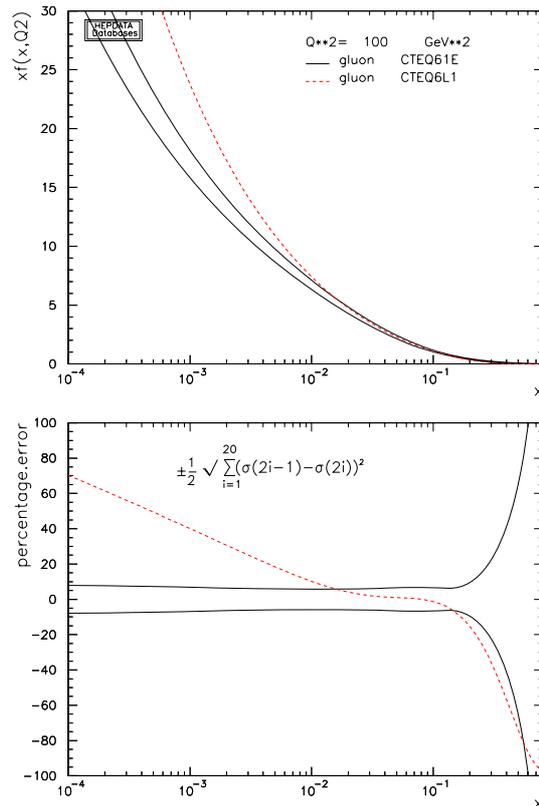
**Figure 27.** The pdf uncertainty for the up quark distribution from the CTEQ6.1 pdf set and a comparison to the CTEQ6L1 up quark distribution.

The CTEQ and MRST uncertainties for the u quark and gluon distributions are shown in Figures 27, 28, 29 and 30. The pdf luminosity uncertainties for various pdf combinations for some kinematic situations appropriate to the LHC will be discussed in Section 6.

Both the CTEQ and MRST groups typically use a fixed value of  $\alpha_s(M_Z)$  (equal to the PDG world average [113] in their global fits. An additional uncertainty in the determination of pdf's results from the uncertainty in the value of  $\alpha_s(M_Z)$  used in the global fits. Figure 31 and Figure 32 show the results of varying the the value of  $\alpha_s$  on the gluon and up quark distributions [114]. As expected, the gluon distribution has a greater sensitivity on  $\alpha_s$  due to the coupling mentioned earlier.

#### 4.5. NLO and LO pdf's

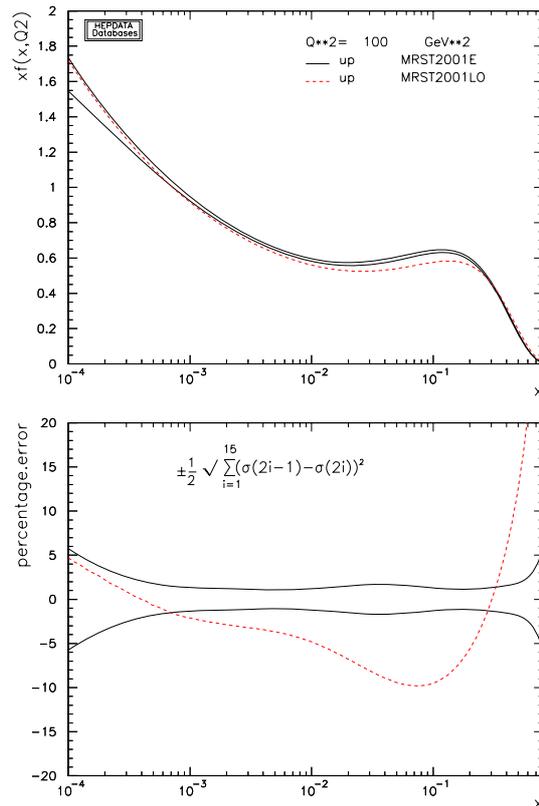
In the past, the global pdf fitting groups have produced sets of pdf's in which leading order rather than next-to-leading order matrix elements, along with the 1-loop  $\alpha_s$  rather than the 2-loop  $\alpha_s$ , have been used to fit the input datasets. The resultant leading order pdf's have most often been used in conjunction with leading order matrix element programs or parton shower Monte Carlos. However, the leading order pdf's of a given set will tend to differ from the central pdf's in the NLO fit, and in fact will



**Figure 28.** The pdf uncertainty for the gluon distribution from the CTEQ6.1 pdf set and a comparison to the CTEQ6L1 gluon distribution.

most often lie outside the pdf error band. Such is the case for the up quark distribution shown in Figures 27 and 29 and the gluon distribution shown in Figures 28 and 30, where the LO pdf's are plotted along with the NLO pdf error bands.

The global pdf fits are dominated by the high statistics, low systematic error deep inelastic scattering data and the differences between the LO and NLO pdf's are determined most often by the differences between the LO and NLO matrix elements for deep inelastic scattering. As the NLO corrections for most processes of interest at the LHC are reasonably small, the use of NLO pdf's in conjunction with LO matrix elements will most often give a closer approximation of the full NLO result (although the result remains formally LO). In addition, it is often useful to examine variations in acceptances in Monte Carlos using the families of NLO error pdf's; thus, it is important to also compare to the predictions using the central (NLO) pdf. It is our recommendation that NLO pdf's be used for predictions at the LHC, even with LO matrix element programs and parton shower Monte Carlos. There are two consequences: the pdf's must be positive-definite in the kinematic regions of interest as they will be used to develop the initial state showering history and (2) underlying event tunes must be available using the NLO pdf's. An underlying event model that uses multiple parton interactions depends strongly on the slope of the low  $x$  gluon distribution. The NLO gluon distribution tends



**Figure 29.** The pdf uncertainty for the up quark distribution from the MRST2001 pdf set and a comparison to the MRST2001LO up quark distribution.

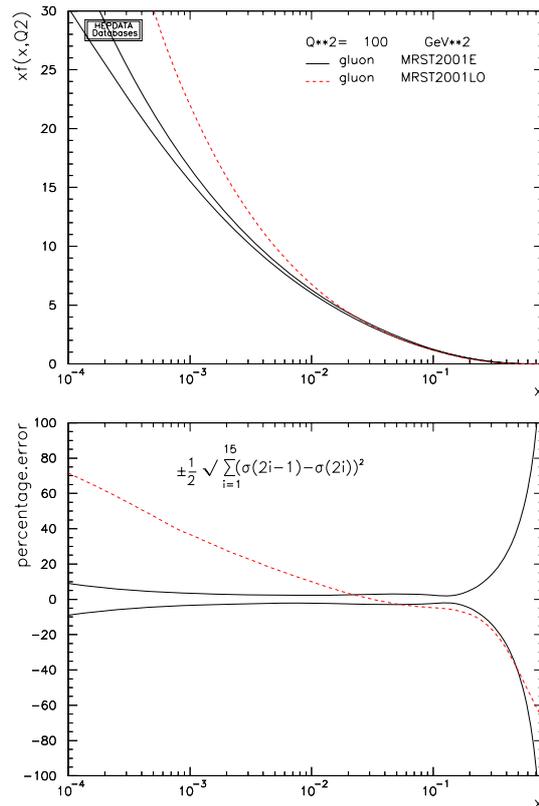
to have a much shallower slope than does the LO gluon and thus a different set of parameters will be needed for the tune. A NLO tune is currently available for Pythia using the (NLO) CTEQ6.1 pdf's that is equivalent to Tune A (see Section 5), which uses CTEQ5L. An equivalent tune for Herwig (with Jimmy) is also available [115]. At the end of the day, though, the accuracy of the LO prediction is still only LO<sup>+</sup>

The transition from NLO to NNLO results in much smaller changes to the pdf's as can be observed in Figure 33.

#### 4.6. Pdf uncertainties and Sudakov form factors

As discussed previously, it is often useful to use the error pdf sets with parton shower Monte Carlos. The caveat still remains that a true test of the acceptances would use a NLO MC. Similar to their use with matrix element calculations, events can be generated once using the central pdf and the pdf weights stored for the error pdf's. These pdf

<sup>+</sup> It has been pointed out in [116] that  $\overline{MS}$  pdf's are appropriate for calculations of inclusive cross sections but are not suitable for the exclusive predictions inherent with Monte Carlo event generators. It is necessary to define pdf's specific to the parton showering algorithm for each Monte Carlo. However, the error introduced by using  $\overline{MS}$  pdf's with leading order parton shower Monte Carlos should be of a higher logarithmic order than the accuracy we are seeking.

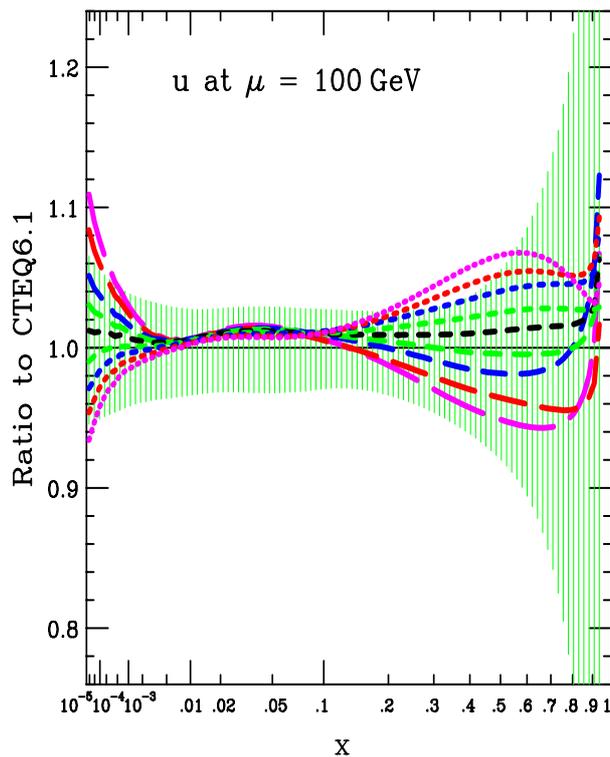


**Figure 30.** The pdf uncertainty for the gluon distribution from the MRST2001 pdf set and a comparison to the MRST2001LO gluon distribution.

weights then can be used to construct the pdf uncertainty for any observable. Some sample code for Pythia is given on the benchmark webpage. One additional complication with respect to their use in matrix element programs is that the parton distributions are used to construct the initial state parton showers through the backward evolution process. The space-like evolution of the initial state partons is guided by the ratio of parton distribution functions at different  $x$  and  $Q^2$  values. Thus the Sudakov form factors in parton shower Monte Carlos will be constructed using only the central pdf and not with any of the individual error pdf's and this may lead to some errors for the calculation of the pdf uncertainties of some observables. However, it was demonstrated in Reference [82] that the pdf uncertainty for Sudakov form factors in the kinematic region relevant for the LHC is minimal, and the weighting technique can be used just as well with parton shower Monte Carlos as with matrix element programs.

#### 4.7. LHAPDF

Libraries such as PDFLIB [117] have been established that maintain a large collection of available pdf's. However, PDFLIB is no longer supported, making it more difficult for easy access to the most up-to-date pdf's. In addition, the determination of the pdf uncertainty of any cross section typically involves the use of a large number of pdf's (on



**Figure 31.** Variation of the up quark distribution for different values of  $\alpha_s$ .  $\alpha_s$  varies from 0.110 (short-dash) to 0.124 (long-dash) in increments of 0.002.

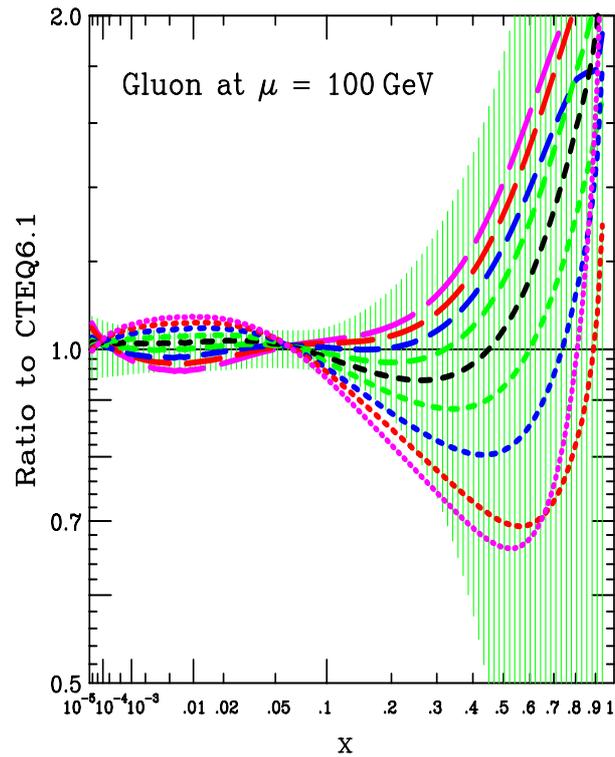
the order of 30-100) and PDFLIB is not set up for easy accessibility for a large number of pdf's.

At Les Houches in 2001, representatives from a number of pdf groups were present and an interface (Les Houches accord 2, or LHAPDF) [119] that allows the compact storage of the information needed to define a pdf was defined. Each pdf can be determined either from a grid in  $x$  and  $Q^2$  or by a few lines of information (basically the starting values of the parameters at  $Q = Q_o$ ) and the interface carries out the evolution to any  $x$  and  $Q$  value, at either LO or NLO as appropriate for each pdf.

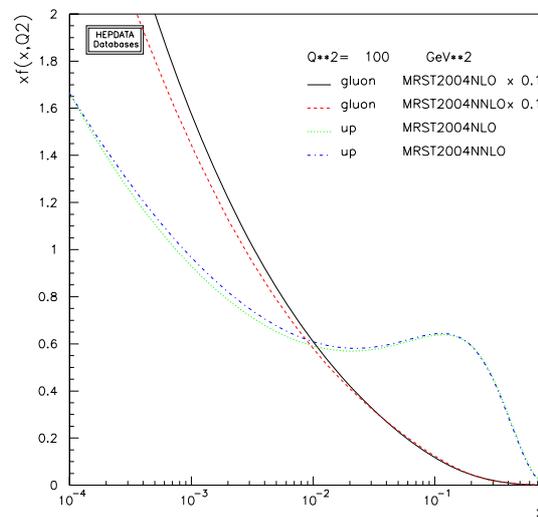
The interface is as easy to use as PDFLIB and consists essentially of 3 subroutine calls:

- call `InitPDFset(name)`: called once at the beginning of the code; *name* is the file name of the external pdf file that defines the pdf set (for example, CTEQ, GKK [118] or MRST)
- call `InitPDF(mem)`: *mem* specifies the individual member of the pdf set
- call `evolvePDF(x,Q,f)`: returns the pdf momentum densities for flavour *f* at a momentum fraction  $x$  and scale  $Q$

Responsibility for LHAPDF has been taken over by Durham [120] and regular updates/improvements have been produced. It is currently included in the matrix element program MCFM [121] and will be included in future versions of other matrix



**Figure 32.** Variation of the gluon distribution for different values of  $\alpha_s$ .  $\alpha_s$  varies from 0.110 (short-dash) to 0.124 (long-dash) in increments of 0.002.



**Figure 33.** A comparison of the NLO and NNLO gluon and up quark distributions for the MRST2004 sets of pdf's.

element and parton shower Monte Carlo programs. Recent modifications make it possible to include all error pdf's in memory at the same time. Such a possibility reduces the amount of time needed for pdf error calculations on any observable. The matrix element result can be calculated once using the central pdf and the relative (pdf) × (pdf)

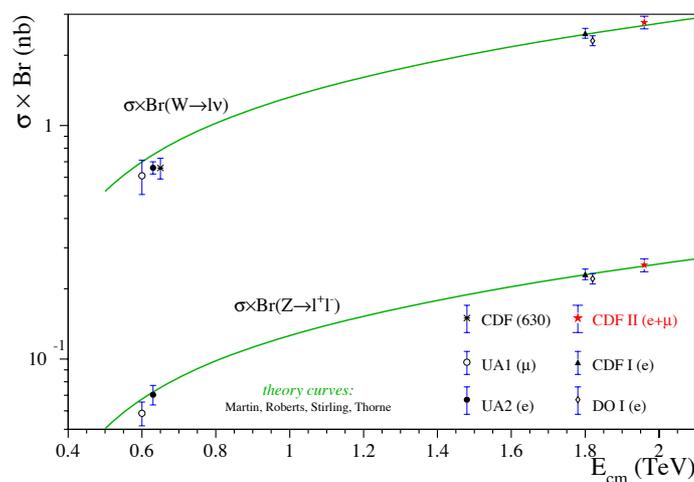
parton-parton luminosity can be calculated for each of the error pdf's (or the values of  $x_1, x_2$ , the flavour of partons 1 and 2 and the value of  $Q^2$  can be stored). Such a pdf re-weighting has been shown to work both for exact matrix element calculations as well as for matrix element+parton shower calculations. In addition, a new routine LHAGLUE [120] provides an interface from PDFLIB to LHAPDF making it possible to use the PDFLIB subroutine calls that may be present in older programs.

## 5. Comparisons to Tevatron data

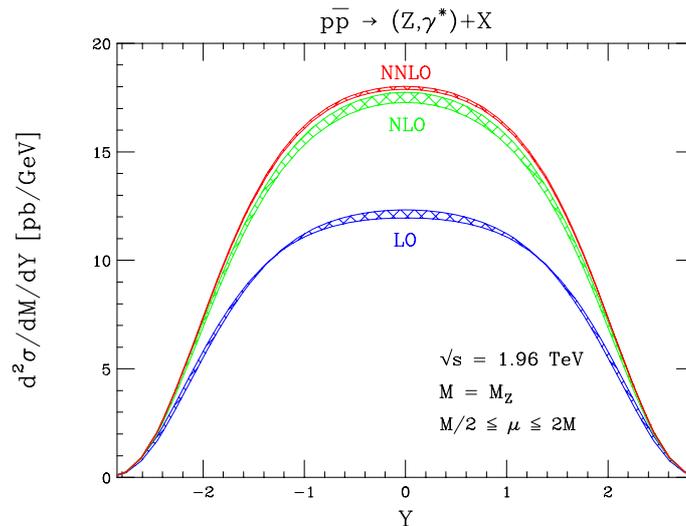
### 5.1. $W/Z$ production

As mentioned previously,  $W$  production at hadron-hadron colliders serves as a precision benchmark for Standard Model physics. Experimentally, the systematic errors are small. The decay leptons are easy to trigger on and the backgrounds are under good control. Theoretically, the cross section and rapidity distributions are known to NNLO. The  $W$  and  $Z$  cross sections measured at the Tevatron are shown in Figure 34 for both Run 1 and Run 2 [7]. The experimental cross sections agree well with the predictions and the Run 2 cross sections show the rise expected from the increase in center-of-mass energy over Run 1. In Figure 4 in Section 2, the Run 2 cross sections from CDF and D0 are compared to predictions at LO, NLO and NNLO. The NNLO predictions are a few percent larger than the NLO predictions. Good agreement is observed with both. It is noteworthy that the largest experimental uncertainty for the  $W$  and  $Z$  cross sections is the uncertainty in the luminosity, typically on the order of 5%. The theoretical systematic errors are primarily from the pdf uncertainty. If other Tevatron cross sections were normalized to the  $W$  cross section, then both the theoretical and experimental systematic errors for those cross sections could be reduced.

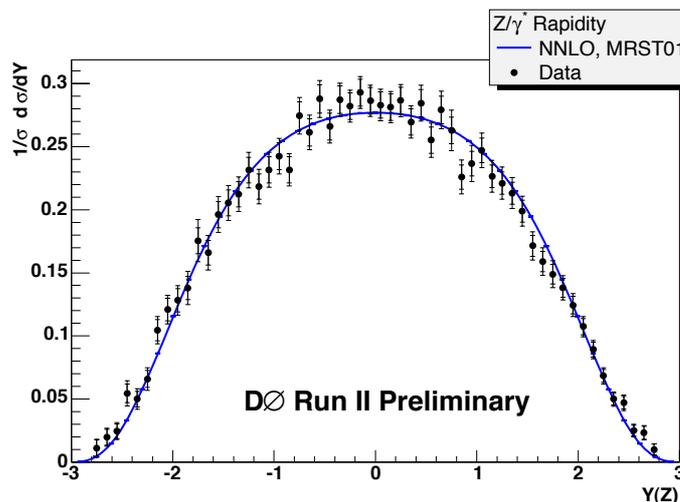
Predictions of the rapidity distribution for  $Z (\rightarrow e^+e^-)$  production at the Tevatron are shown in Figure 35 [65, 123]. There is a shape change in going from LO to NLO,



**Figure 34.**  $W$  and  $Z$  cross sections as a function of the center-of-mass energy.



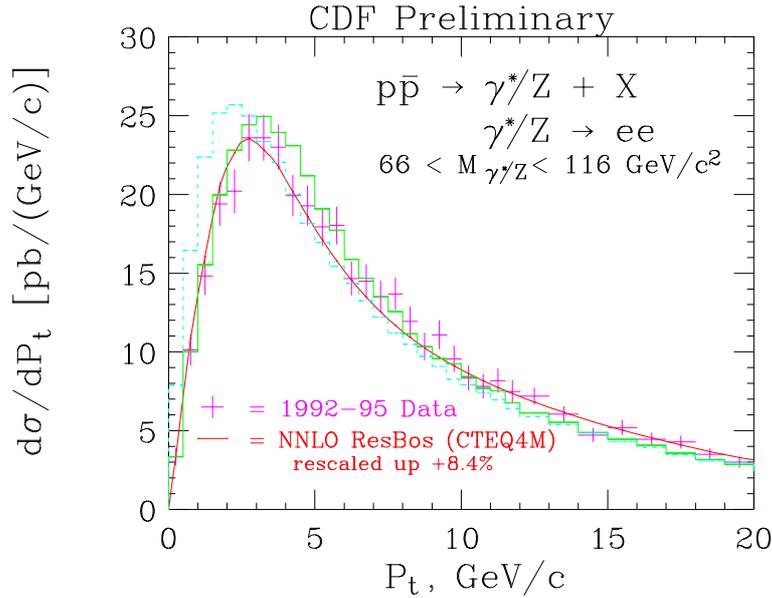
**Figure 35.** Predictions for the rapidity distribution of an on-shell  $Z$  boson in Run 2 at the Tevatron at LO, NLO and NNLO. The bands indicate the variation of the renormalization and factorization scales within the range  $M_Z/2$  to  $2M_Z$ .



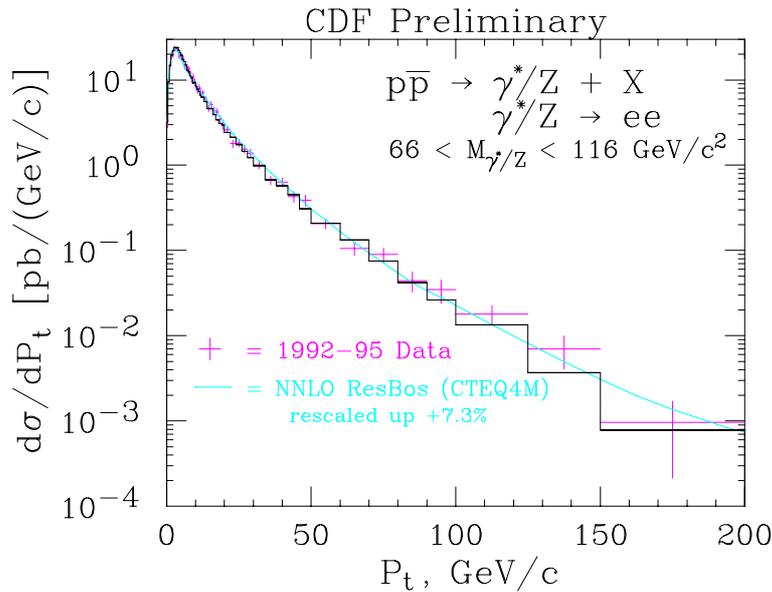
**Figure 36.**  $Z$  rapidity distribution from D0 in Run 2.

as well as an increase in normalization, while the transition from NLO to NNLO is essentially just a small  $K$ -factor with little change in shape. The  $Z$  rapidity distribution measured by D0 in Run 2 is shown in Figure 36; the measurement agrees with the NNLO prediction over the entire rapidity range. Such cross sections may be useful as input to global pdf fits in the near future.

The transverse momentum distribution for  $Z$  bosons at the Tevatron (CDF Run 1) is shown in Figures 37 (low  $p_T$ ) and 38 (all  $p_T$ ) along with comparisons to Pythia and ResBos. The ResBos predictions agree well with the data over the entire kinematic region, while the Pythia predictions require the addition of an intrinsic  $k_T$  of approximately 2 GeV/ $c$  for best agreement at lower transverse momentum. This intrinsic  $k_T$  is not totally due to the finite size of the proton; it mostly results from



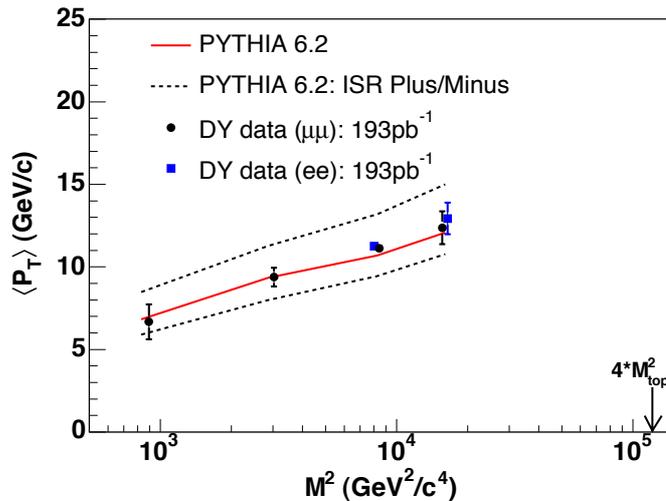
**Figure 37.** The transverse momentum distribution (low  $p_T$ ) for  $Z \rightarrow e^+e^-$  from CDF in Run 1, along with comparisons to predictions from Pythia and ResBos. The Pythia solid-green curve has had an additional 2 GeV/c of  $k_T$  added to the parton shower.



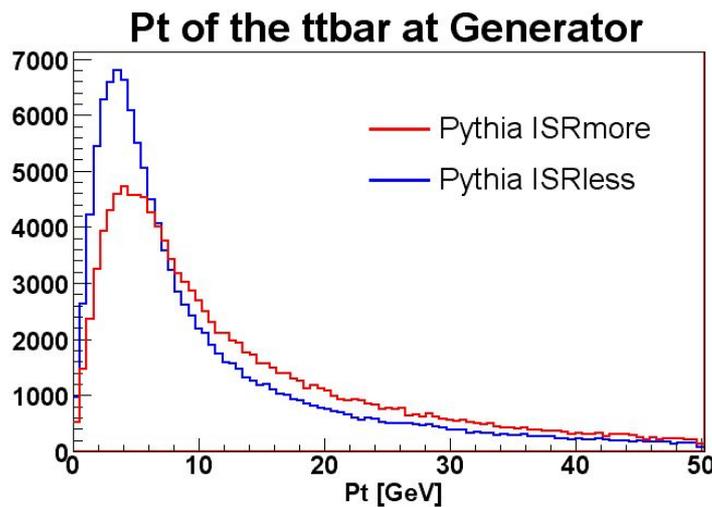
**Figure 38.** The transverse momentum distribution (full  $p_T$  range) for  $Z \rightarrow e^+e^-$  from CDF in Run 1, along with comparisons to predictions from Pythia and ResBos.

the finite cutoff in the parton shower evolution and the need to add the missing component back ‘by-hand’. ResBos correctly describes the non-perturbative as well as the perturbative region.

The average transverse momentum for Drell-Yan production has been measured as a function of the Drell-Yan pair mass in CDF in Run 2. This is shown in Figure 39.



**Figure 39.** The average transverse momentum for Drell-Yan pairs from CDF in Run 2, along with comparisons to predictions from Pythia.



**Figure 40.** The Pythia predictions for the  $t\bar{t}$  transverse momentum using the ‘Plus/Minus’ tunes.

The average transverse momentum (neglecting the high  $p_T$  tail) increases roughly logarithmically with the square of the Drell-Yan mass, as predicted by (18) in Section 3. The data agree well with the default Pythia 6.2 prediction using Tune A. Also shown are two predictions involving tunes of Pythia that give larger/smaller values for the average Drell-Yan transverse momentum as a function of the Drell-Yan mass. The *Plus/Minus* tunes were used to estimate the initial-state-radiation uncertainty for the determination of the top mass in CDF. Most of the  $t\bar{t}$  cross section at the Tevatron arises from  $q\bar{q}$  initial states, so the Drell-Yan measurements serve as a good model. The predictions for the  $t\bar{t}$  transverse momentum from the *Plus/Minus* tunes are shown in Figure 40. The information regarding these tunes is available on the benchmarks webpage.

## 5.2. Underlying event

What is meant by a minimum bias event is somewhat murky, and the exact definition will depend on the trigger of each experiment. The description of the underlying event energy and of minimum bias events requires a non-perturbative and/or semi-perturbative phenomenological model. There are currently a number of models available, primarily inside parton shower Monte Carlo programs, to predict both of these processes. We discuss several of the popular models below. An understanding of this soft physics is interesting in its own right but is also essential for precision measurements of hard interactions where the soft physics effects need to be subtracted.

Perhaps the simplest model for the underlying event is the uncorrelated soft scattering model present in Herwig. Basically, the model is a parametrization of the minimum bias data taken by the UA5 experiment [122] at the CERN  $p\bar{p}$  Collider. The model tends to predict underlying event distributions softer than measured at the Tevatron and has a questionable extrapolation to higher center-of-mass energies. A newer model for the underlying event in Herwig is termed “Jimmy” \* and describes the underlying event in terms of multiple parton interactions at a scale lower than the hard scale and with the number of such parton scatterings depending on the impact parameter overlap of the two colliding hadrons.

The Pythia model for the underlying event also utilizes a multiple parton interaction framework with the total rate for parton-parton interactions assumed to be given by perturbative QCD. A cutoff,  $p_{Tmin}$ , is introduced to regularize the divergence as the transverse momentum of the scattering goes to zero. The rate for multiple parton interactions depend strongly on the value of the gluon distribution at low  $x$ . The cutoff,  $p_{Tmin}$ , is the main free parameter of the model and basically corresponds to an inverse colour screening distance. A tuning of the Pythia underlying event parameters (Tune A) basically succeeds in describing all of the global event properties in events at the Tevatron. With the new version of Pythia (version 6.4) [81, 11], a new model for the underlying event is available, similar in spirit to the old multiple parton interaction model, but with a more sophisticated treatment of colour, flavour and momentum correlations in the remnants.

*5.2.1. Inclusive jet production* It is useful to consider the measurement of inclusive jet production at the Tevatron as (1) it probes the highest transverse momentum range accessible at the Tevatron, (2) it has a large impact on global pdf analyses, and (3) many of the subtleties regarding measurements with jets in the final state and the use of jet algorithms come into play.

As shown in Figure 41, a dijet event at a hadron-hadron collider consists of a hard collision of two incoming partons (with possible gluon radiation from both the incoming and outgoing legs) along with softer interactions from the remaining partons in the colliding hadrons (“the underlying event energy”).

\* <http://hepforge.cedar.ac.uk/jimmy/>

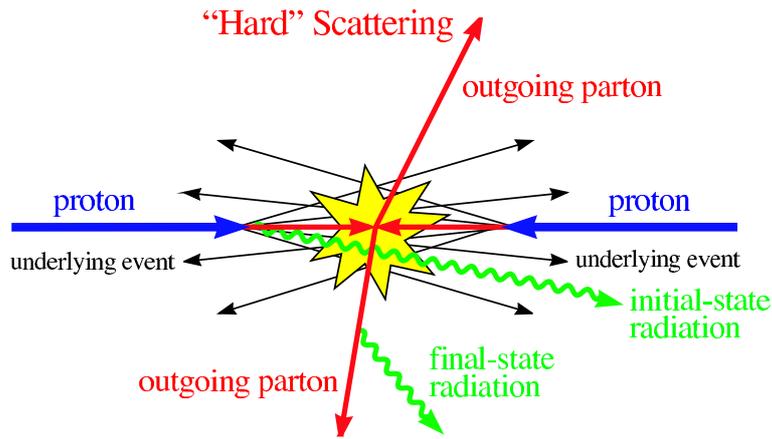


Figure 41. Schematic cartoon of a  $2 \rightarrow 2$  hard scattering event.

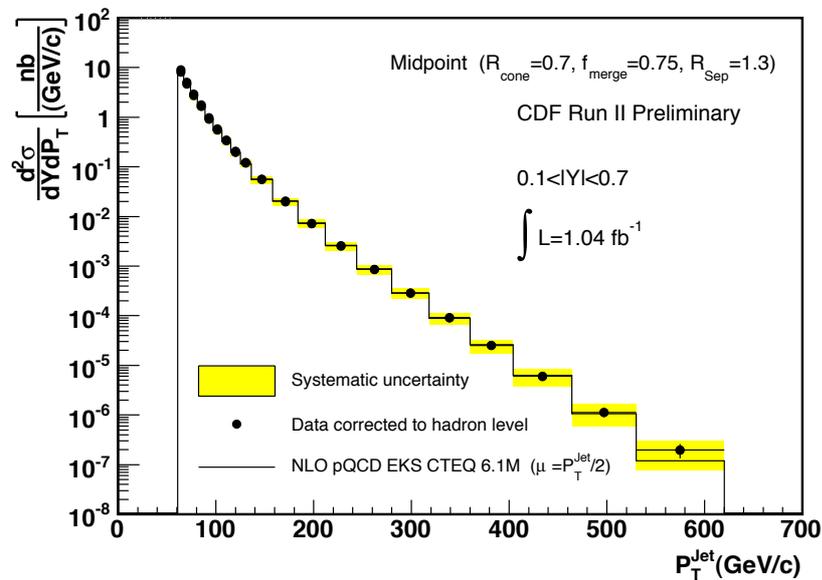


Figure 42. The inclusive jet cross section from CDF in Run 2.

The inclusive jet cross section measured by the CDF Collaboration in Run 2 is shown in Figure 42, as a function of the jet transverse momentum. Due to the higher statistics compared to Run 1, and the higher center-of-mass energy, the reach in transverse momentum has increased by approximately  $150 \text{ GeV}/c$ . The measurement uses the midpoint cone algorithm with a cone radius of 0.7. As discussed previously, the midpoint algorithm places additional seeds (directions for jet cones) between stable cones having a separation of less than twice the size of the clustering cones. The midpoint algorithm uses four-vector kinematics for clustering individual partons, particles or energies in calorimeter towers, and jets are described using rapidity ( $y$ ) and transverse

momentum ( $p_T$ ) ‡.

For comparison of data to theory, the calorimeter tower energies clustered into a jet must first be corrected for the detector response. The calorimeters in the CDF experiment respond differently to electromagnetic showers than to hadronic showers, and the difference varies as a function of the transverse momentum of the jet. The detector response corrections are determined using a detector simulation in which the parameters have been tuned to test-beam and in-situ calorimeter data. Pythia 6.216, with Tune A, is used for the production and fragmentation of jets. The same clustering procedure is applied to the final state particles in Pythia as is done for the data. The correction is determined by matching the calorimeter jet to the corresponding particle jet. An additional correction accounts for the smearing effects due to the finite energy resolution of the calorimeter. At this point, the jet is said to be determined at the “hadron level.”

For data to be compared to a parton level calculation, either the data must be corrected from the hadron level to the parton level or the theory must be corrected to the hadron level. Here we describe the former; the latter just involves the inverse corrections. It is our “official recommendation” that, where possible, the data be presented in both ways, with a clear indication of how the corrections have been performed. The hadronization corrections consist of two components: the subtraction from the jet of the underlying event energy not associated with the hard scattering and the correction for a loss of energy outside a jet due to the fragmentation process. The hadronization corrections can be calculated by comparing the results obtained from Pythia at the hadron level to the results from Pythia when the underlying event and the parton fragmentation into hadrons has been turned off. The underlying event energy is due to the interactions of the spectator partons in the colliding hadrons and the size of the correction depends on the size of the jet cone. It is approximately 0.5 GeV/c for a cone of radius 0.7 and is similar to the amount of energy observed in minimum bias events with a high track multiplicity. The rule-of-thumb has always been that the underlying event energy in a jet event looks very much like that observed in minimum bias events, i.e. that there is a rough factorization of the event into a hard scattering part and a soft physics part.

Studies have been carried out with inclusive jet production in CDF, examining the transverse momentum carried by charged particles inside and outside of jets [124, 125]. For example, the geometry for one study is shown in Figure 43, where the “towards” and “away” regions have been defined with respect to the direction of the leading jet. The transverse momenta in the two transverse regions are shown in Figure 44.

The region with the largest transverse momentum is designated as the max region and the one with the lowest, the min region. As the lead jet transverse momentum increases, the momentum in the max region increases; the momentum in the min region does not. The amount of transverse momentum in the min region is consistent with that observed in minimum bias events at the Tevatron. At the partonic level, the max region

‡ Self-contained versions of the CDF Midpoint cone algorithm, as well as the Run 1 cone algorithm JetClu, are available at the benchmark website, in both Fortran and C++ versions

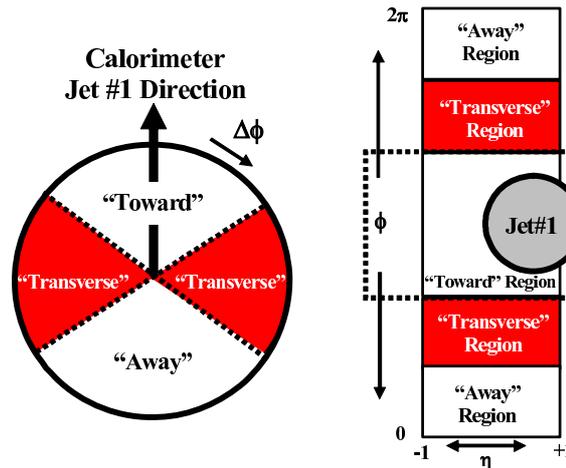


Figure 43. Definition of the “toward”, “away” and “transverse” regions.

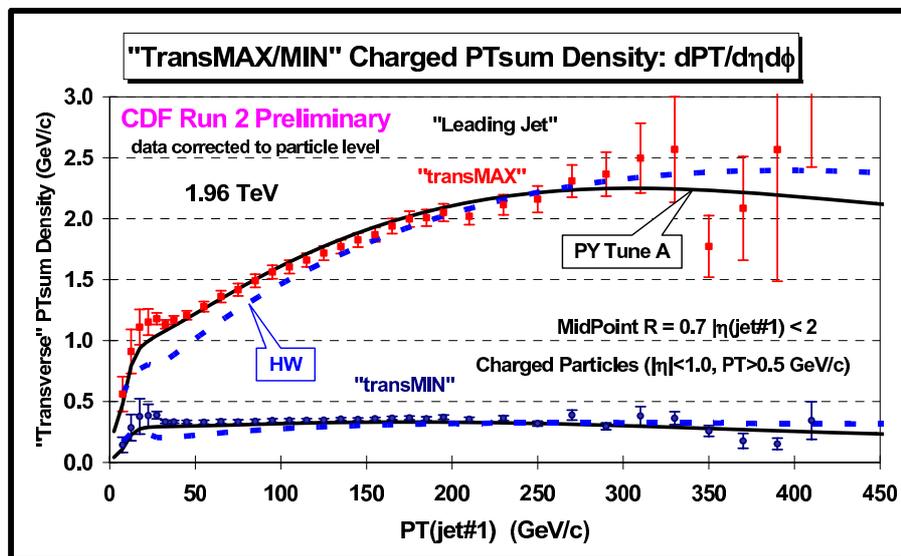
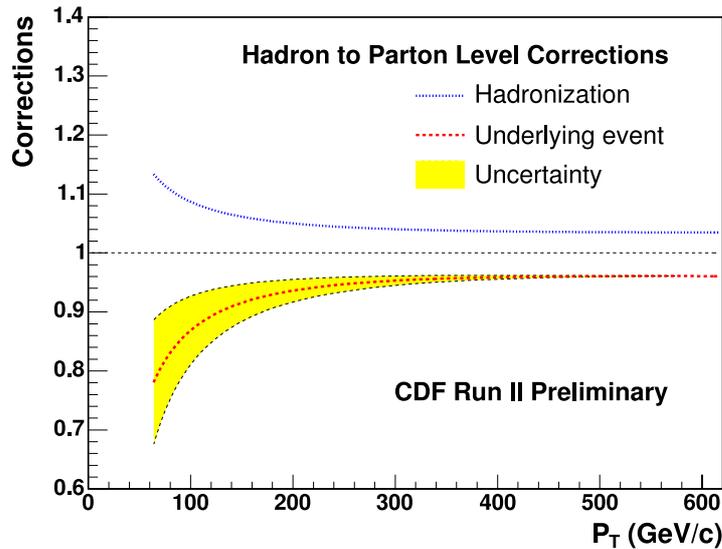


Figure 44. The sum of the transverse momenta of charged particles inside the TransMAX and TransMIN regions, as a function of the transverse momentum of the leading jet.

can receive contributions from the extra parton present in NLO inclusive jet calculations. The min region can not. There is good agreement between the Tevatron data and the Pythia tunes. Tune A was determined using the LO pdf CTEQ5L; an equivalently good agreement is also observed with a tune using the NLO pdf CTEQ6.1 ††. As discussed previously, such a tune is necessary for the use of a NLO pdf such as CTEQ6.1 with the Pythia Monte Carlo.

The fragmentation correction accounts for the daughter hadrons ending up outside the jet cone from mother partons whose trajectories lie inside the cone (also known

††The parameters for the tune are given on the benchmarks webpage.

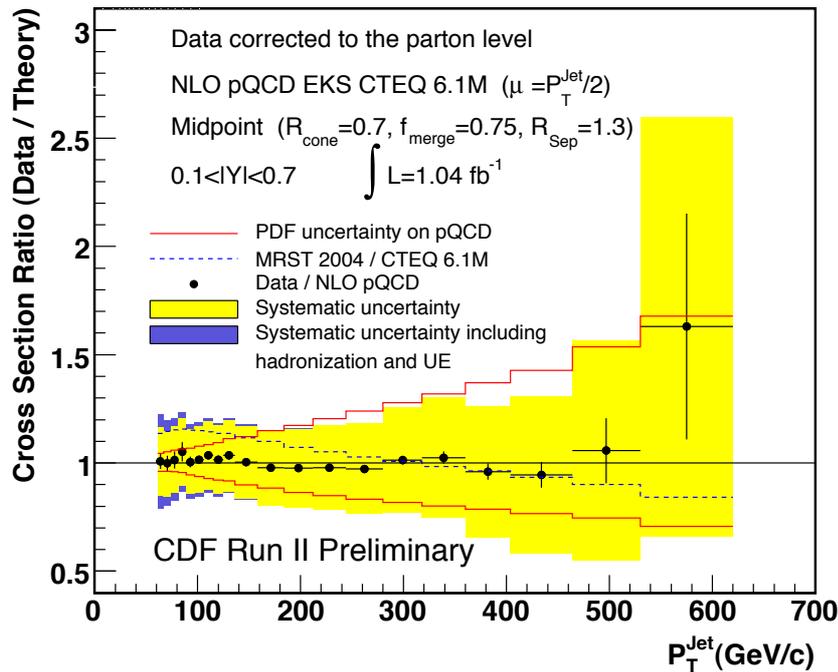


**Figure 45.** Fragmentation and underlying event corrections for the CDF inclusive jet result, for a cone size  $R = 0.7$ .

as *splash-out*); it does not correct for any out-of-cone energy arising from perturbative effects as these should be correctly accounted for in a NLO calculation. It is purely a power correction to the cross section. The numerical value of the splash-out energy is roughly constant at 1 GeV/c for a cone of radius 0.7, independent of the jet transverse momentum. This constancy may seem surprising. But, as the jet transverse momentum increases, the jet becomes more collimated; the result is that the energy in the outmost annulus (the origin of the splash-out energy) is roughly constant. The correction for splash-out derived using parton shower Monte Carlo can be applied to a NLO parton level calculation to the extent to which both the parton shower and the 2 partons in a NLO jet correctly describe the jet shape.

The two effects (underlying event and splash-out) go in opposite directions so there is a partial cancellation in the correction to parton level. For a jet cone of 0.7, the underlying event correction is larger, as seen in Figure 45, for the case of inclusive jet production at CDF. For a jet cone radius of 0.4, the fragmentation correction remains roughly the same size but the underlying event corrections scales by the ratio of the cone areas; as a result the two effects basically cancel each other out over the full transverse momentum range at the Tevatron.

A comparison of the inclusive jet cross section measured by CDF in Run 2 with the midpoint cone algorithm [127] to NLO pQCD predictions using the EKS [126] program with the CTEQ6.1 and MRST2004 pdf's is shown in Figure 46. A renormalization/factorization scale of  $(p_T^{jet}/2)$  has been used in the calculation. Typically, this leads to the highest predictions for inclusive jet cross sections at the Tevatron, as discussed in Section 2. There is good agreement with the CTEQ6.1

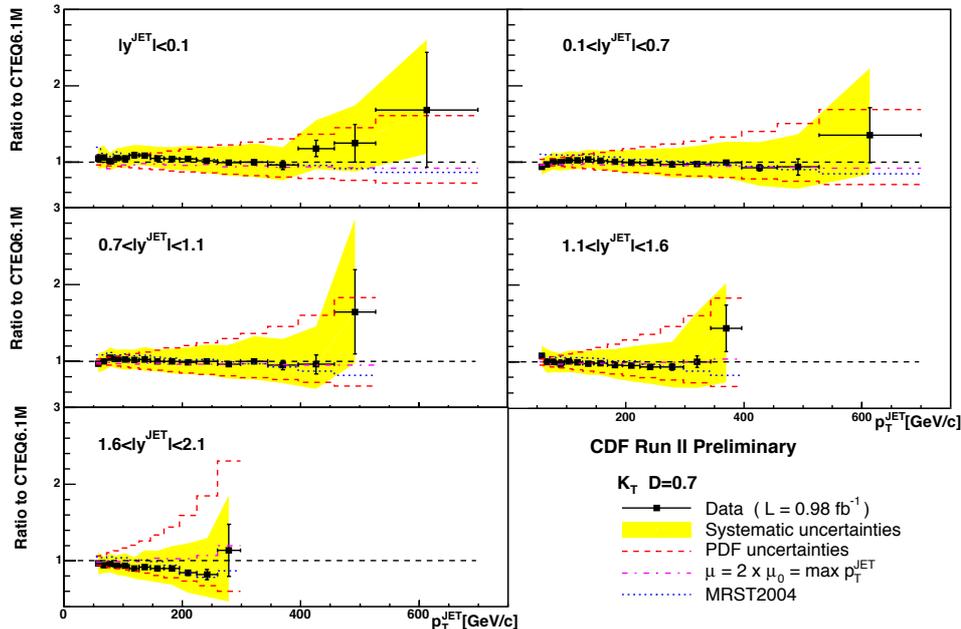


**Figure 46.** The inclusive jet cross section from CDF in Run 2 compared on a linear scale to NLO theoretical predictions using CTEQ6.1 and MRST2004 pdf's.

predictions over the transverse momentum range of the prediction. The MRST2004 predictions are slightly higher at lower  $p_T$  and slightly lower at higher  $p_T$ , but still in good overall agreement. As noted before, the CTEQ6.1 and MRST2004 pdf's have a higher gluon at large  $x$  as compared to previous pdf's, due to the influence of the Run 1 jet data from CDF and D0. This enhanced gluon provides a good agreement with the high  $p_T$  Run 2 measurement as well which, as stated before, extends approximately 150 GeV/c higher in transverse momentum. The red curves indicate the pdf uncertainty for the prediction using the CTEQ6.1 pdf error set. The yellow band indicates the experimental systematic uncertainty, which is dominated by the uncertainty in the jet energy scale (on the order of 3%). The purple band shows the effect of the uncertainty due to the hadronization and underlying event, which is visible only for transverse momenta below 100 GeV/c. In Figure 47, the jet cross sections measured with the midpoint cone algorithm are shown for the full rapidity coverage of the CDF experiment. Good agreement is observed in all rapidity regions with the CTEQ6.1 predictions. It is also important to note that for much of the kinematic range, the experimental systematic errors are less than pdf uncertainties; thus, the use of this data in future global pdf fits should serve to further constrain the gluon pdf.

For many events, the jet structure is clear and the jets to which the individual towers should be assigned are fairly un-ambiguous. However, in other events such as Figure 48, the complexity of the energy depositions means that different algorithms will





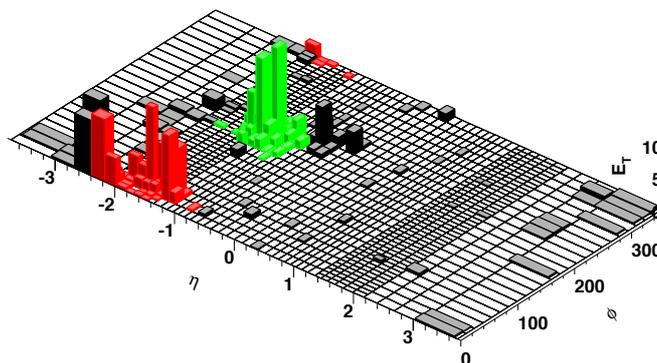
**Figure 49.** The inclusive jet cross section from CDF in Run 2, for several rapidity intervals using the  $k_T$  jet algorithm, compared on a linear scale to NLO theoretical predictions using CTEQ6.1 pdf's.

shower Monte Carlos, but, as discussed in Section 3, NLO calculations can place at most 2 partons in a jet.

In Figure 49, the experimental jet cross sections using the  $k_T$  algorithm from CDF Run 2 [128] are compared to NLO predictions using the JETRAD [83] program. Similar good agreement to that obtained for the midpoint cone algorithm is observed. This is an important observation. The two different jet algorithms have different strengths and weaknesses. It is our recommendation that, where possible, analyses at the Tevatron and LHC use both algorithms in order to obtain more robust results.

A particular complexity with the cone algorithm occurs when two jets overlap; a decision must be made whether to merge the two jets into one, or to separate them. This is an experimental decision; in CDF in Run 2, the two overlapping jets are merged when more than 75% of the smaller jet energy overlaps with the larger jet. When the overlap is less, the towers are assigned to the nearest jet. D0 uses a criterion of a 50% fraction. NLO theory is agnostic on the subject as there is no overlap between the two partons that can comprise a jet.

Another problem that can arise on the particle or calorimeter level, but not on the NLO parton level, occurs when particles or calorimeter towers remain unclustered in any jet, due to the strong attraction of a nearby larger jet peak that will attract away any trial jet cone placed at the location of the original particles/calorimeter towers.



**Figure 50.** An example of a Monte Carlo inclusive jet event where the midpoint algorithm has left substantial energy unclustered.

The result will be what [84] calls “dark towers”, i.e. clusters that have a transverse momentum large enough to be designated either a separate jet or to be included in an existing nearby jet, but which are not clustered into either. Such a Monte Carlo event is shown in Figure 50, where the towers unclustered into any jet are shaded black. A simple way of understanding these dark towers begins by defining a “Snowmass potential” in terms of the 2-dimensional vector  $\vec{r} = (y, \phi)$  via

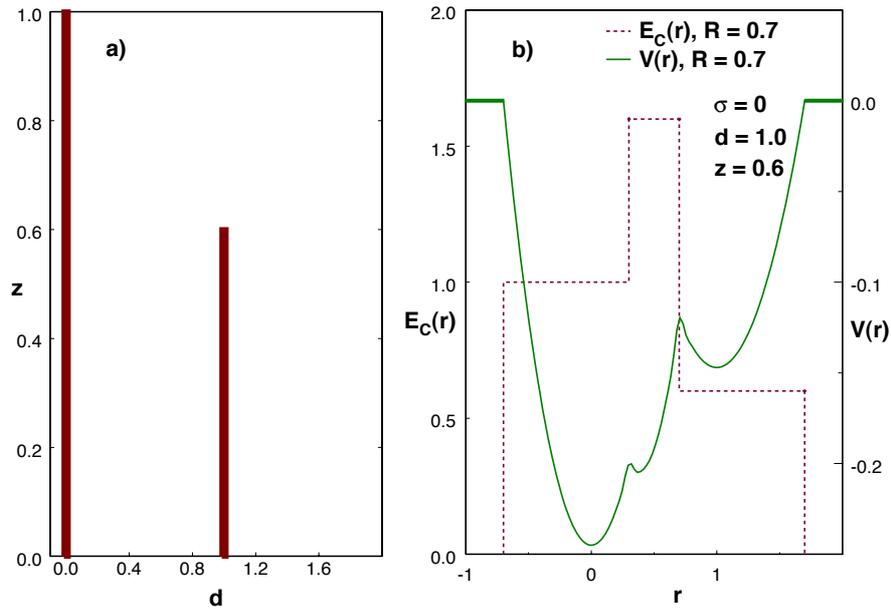
$$V(\vec{r}) = -\frac{1}{2} \sum_j p_{T,j} \left( R_{cone}^2 - (\vec{r}_j - \vec{r})^2 \right) \Theta \left( R_{cone}^2 - (\vec{r}_j - \vec{r})^2 \right). \quad (39)$$

The flow is then driven by the “force”  $\vec{F}(\vec{r}) = -\vec{\nabla} V(\vec{r})$  which is thus given by,

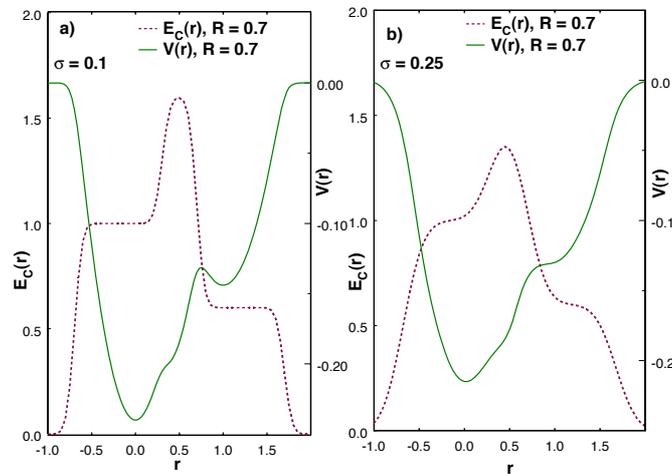
$$\begin{aligned} \vec{F}(\vec{r}) &= \sum_j p_{T,j} (\vec{r}_j - \vec{r}) \Theta \left( R_{cone}^2 - (\vec{r}_j - \vec{r})^2 \right) \\ &= \left( \vec{r}_{C(\vec{r})} - \vec{r} \right) \sum_{j \subset C(\vec{r})} p_{T,j}, \end{aligned} \quad (40)$$

where  $\vec{r}_{C(\vec{r})} = (\bar{y}_{C(\vec{r})}, \bar{\phi}_{C(\vec{r})})$  and the sum runs over  $j \subset C(\vec{r})$  such that  $\sqrt{(y_j - y)^2 + (\phi_j - \phi)^2} \leq R_{cone}$ . As desired, this force pushes the cone to the stable cone position.

In Figure 51 (left), two partons are placed a distance  $\Delta R = 0.9$  apart; the second parton has a fractional energy 0.6 that of the first parton. On the right, the potential  $V(r)$  and the energy contained inside a cone of radius 0.7 are plotted plotted for this parton configuration. At the parton level, there are three positions where minima are present in the potential: at the position of the left parton, the right parton and the midpoint between the two partons. The midpoint jet algorithm applied to a NLO parton level calculation, as discussed in Section 3, would find all three solutions. In Figure 52(left), the spatial distributions of the two partons have been smeared with a spatial resolution  $\sigma$  of 0.1, as would take place for example due to the effects of parton showering and hadronization. The central minimum is wiped out by the smearing.

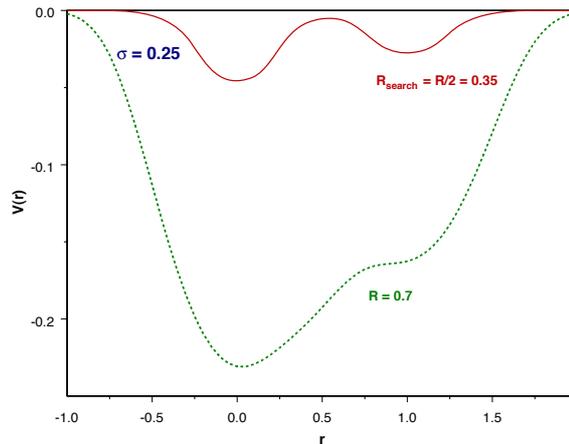


**Figure 51.** A schematic depiction of a specific parton configuration and the results of applying the midpoint cone jet clustering algorithm. The potential discussed in the text and the resulting energy in the jet are plotted.



**Figure 52.** A schematic depiction of the effects of smearing on the midpoint cone jet clustering algorithm.

On the right, the energy distributions of the two partons have been smeared with a resolution  $\sigma$  of 0.25 and both the midpoint and right minima have been wiped out. Any attempt to place the centroid of a jet cone at the position of the right parton will result in the centroid “sliding” to the position of the left parton and the energy corresponding to the right parton remaining unclustered in any jet. This is the origin of the dark towers. The effective smearing in the data lies between a  $\sigma$  of 0.1 and 0.25.



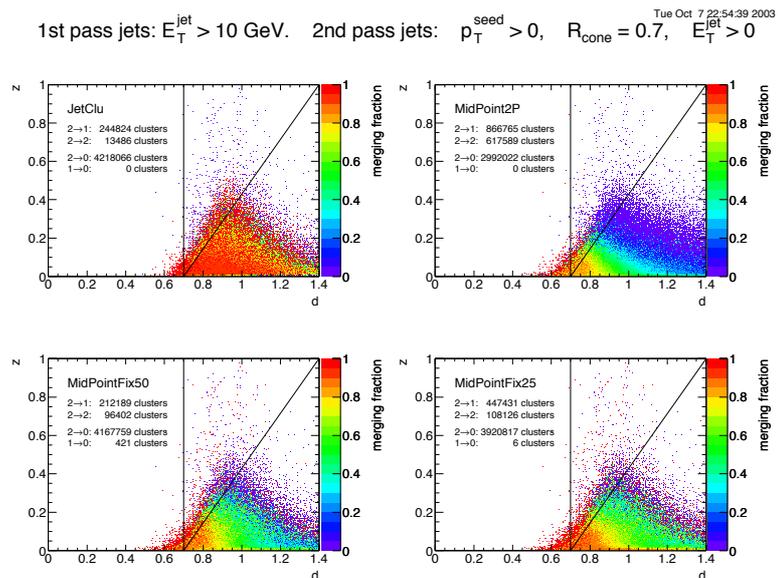
**Figure 53.** A schematic depiction of the effects of smearing on the midpoint cone jet clustering algorithm and the result of using a smaller initial search cone.

A solution recommended in [84] is to start the jet clustering procedure with a cone of radius half that of the final radius, and only as a final step to expand the cone to the full radius. In this way, there is less influence from nearby jet peaks in the formation of stable jet cones, and clusters arising from a hard scatter are more likely to be included in a jet, either combined in a single jet or to be reconstructed as separate jets, depending on the clusters' relative energy and separation. In the example discussed earlier, the potential, in the presence of smearing, recovers the second minimum, but not the central minimum (that includes both partons in a single jet), as shown in Figure 53. Thus, the experimental algorithm will not combine the two jets arising from the two partons, even though the corresponding midpoint algorithm applied to the partons (with  $R_{sep} = 2.0$ ) would. It was originally thought that with the addition of a midpoint seed, the value of  $R_{sep}$  used with the NLO theory could be returned to its *natural* value of 2.0. Now it is realized that the effects of parton showering/hadronization require it to remain at approximately 1.3 for the NLO jet algorithm to best model the experimental one. The theory cross section with  $R_{sep} = 1.3$  is approximately 5% smaller than with  $R_{sep} = 2.0$ , independent of the jet transverse momentum.

The net impact of adding the step with the smaller initial search cone is a 5% increase in the experimental inclusive jet cross section, due to the merging, in some events, of the dark towers with the larger nearby jet. There is also a cost to the use of this algorithm of a larger infrared sensitivity of the algorithm, at the level of 1% †. We reiterate that the current level of theoretical sophistication (NLO) is insensitive to this subtlety in the algorithm and note that this solution is still controversial. For example, it is currently in use by CDF but not by D0 ‡. More work remains to be done to prepare

† Note that (1) all cone algorithms are IR-sensitive at some level and (2) this is an issue of sensitivity, not of safeness.

‡ See the benchmark webpages for more discussion of the controversy

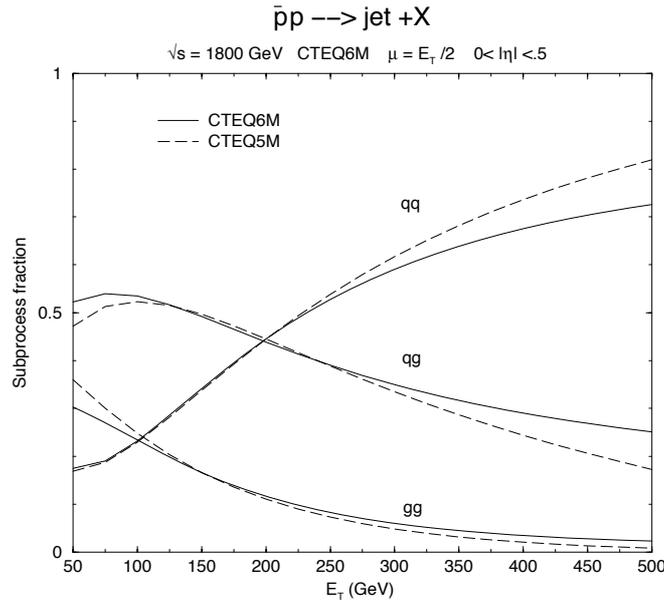


**Figure 54.** The probability for two nearby jets to be merged for several different cone jet algorithms. MidPointFix50 corresponds to the smaller initial search cone algorithm currently used by CDF; there is no initial search cone with the MidPoint2P version of the algorithm.

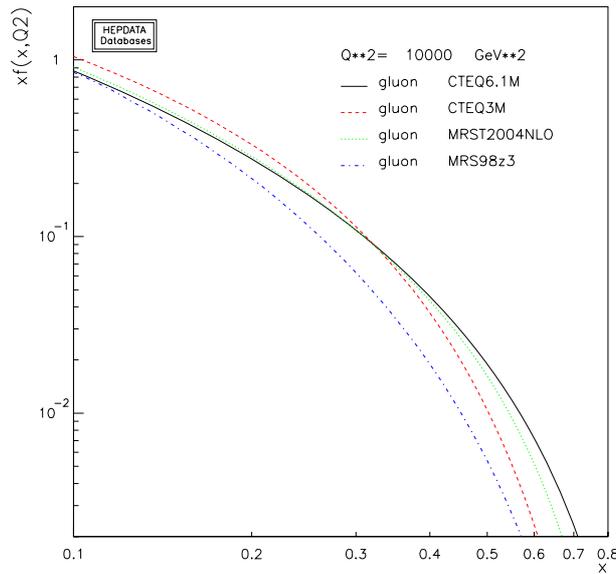
cone algorithms for precision measurements at the LHC. Possible solutions are discussed on the benchmarks webpage.

Figure 54 shows the probability for two nearby jets to be merged as a function of the separation  $d$  and the energy fraction  $z$  (given by  $p_T^{\text{jet}2}/p_T^{\text{jet}1}$ ) for the midpoint algorithm with and without the smaller initial search cone. The density of points in each kinematic region indicates the relative population of events. The distribution is peaked at  $d$  near the value of the cone radius (0.7) due to the effects of the collinear singularity and at low  $z$  due to the soft singularity. The colour of each point indicates the percentage of the time that the two jets are merged by the jet algorithm. This depends on the details of the way in which the energy of the two jets is distributed, i.e. the stochastic process of parton showering and hadronization starting from the same initial partons. Ideally, for the detector algorithm at the detector level to match the jet algorithm at the NLO parton level, all jets above the diagonal line and to the left of the vertical line would be merged. All others would be reconstructed as separate jets. The use of the smaller initial search cone creates more merging in the region where the two partons in a NLO calculation would be merged; it may also create more merging at low ( $d, z$ ) where the two partons would not be merged. This emphasizes the difficulty of matching the behavior of jet algorithms at the parton and hadron levels, especially with the cone algorithm.

Inclusive jet production receives contributions from  $gg$ ,  $gq$  and  $qq(q\bar{q})$  initial states as shown in Figure 55. The experimental precision of the measurement, along with the remaining theoretical uncertainties, means that the cross sections do not serve as a

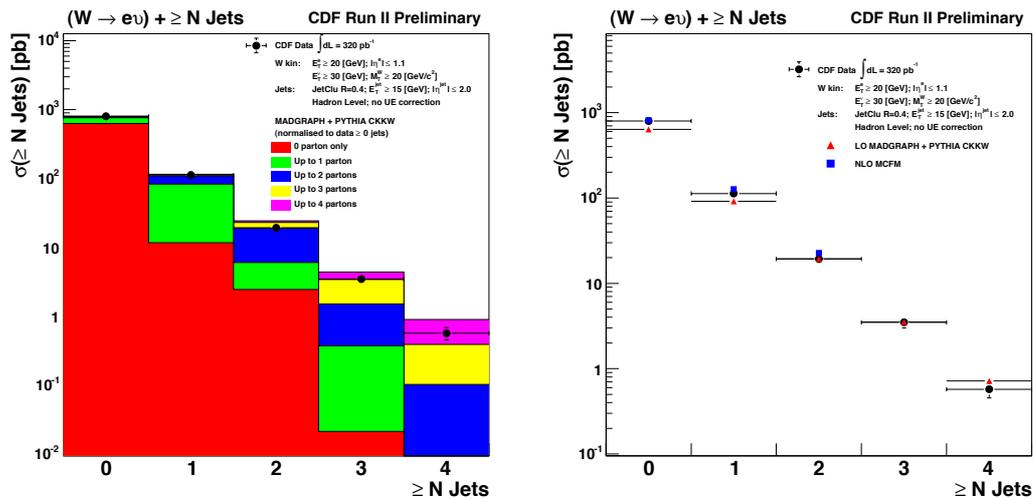


**Figure 55.** The subprocess contributions to inclusive jet production at the Tevatron for the CTEQ5M and CTEQ6M pdf's. The impact of the larger larger gluon at high  $x$  for CTEQ6 is evident.



**Figure 56.** Gluon pdf's before and after the inclusion of Tevatron inclusive jet data.

meaningful constraint on the quark or anti-quark distributions. However, they do serve as an important constraint on the gluon distribution, especially at high  $x$ . Figure 56 shows the gluon distributions for the CTEQ and MRST groups prior to the inclusion of any inclusive jet data and the latest sets which include the Run 1 data from both CDF and D0. The influence of the high  $E_T$  Run 1 jet cross section on the high  $x$  gluon is evident. There is always the danger of sweeping new physics under the rug of



**Figure 57.** The rate of production of  $W + n$  jets at CDF, compared to CKKW (left) and also NLO QCD using MCFM (right). The measurements and predictions were performed with a cone jet of radius 0.4 and with a requirement of 15 GeV/c or greater. The CKKW predictions are normalized to the first bin and use a scale of 10 GeV/c for the matching; the MCFM predictions are absolutely normalized.

pdf uncertainties. Thus, it is important to measure the inclusive jet cross section over as wide a kinematic range as possible. New physics tends to be central while a pdf explanation should be universal, i.e. fit the data in all regions.

**5.2.2.  $W/Z + jets$**  The production of a  $W$  or  $Z$  boson in conjunction with jets is an interesting process in its own right as well as a background to many SM and non-SM physics signals. Jet multiplicities of up to 7 have been measured at the Tevatron. The NLO cross sections have been calculated only for  $W/Z + 2$  jets; predictions for the higher jet multiplicity final states are accessible through matrix element (+ parton shower) predictions and in fact can be considered as a prime testing ground for the accuracy of such predictions as well as for measurements of  $\alpha_s$ . Some comparisons to a recent measurement of  $W \rightarrow e\nu + \geq n$  jets from CDF are shown below. In this analysis, the data have been reconstructed using the JetClu cone algorithm (with a Midpoint cone analysis to come) with a cone radius of 0.4. A smaller jet cone size is preferred for final states that may be “complicated” by the presence of a large number of jets. The data have been compared, at the hadron level, to predictions using matrix element information from ALPGEN and parton shower and hadronization information from Pythia.

The jet multiplicity distribution for  $W + n$  jets measured at the Tevatron is shown in Figure 57. In the left-hand plot, the data have been compared to the CKKW predictions from [88] (normalized to the first bin of the data). The CKKW predictions use matrix element information for up to 4 partons in the final state. Thus, most of the cross section

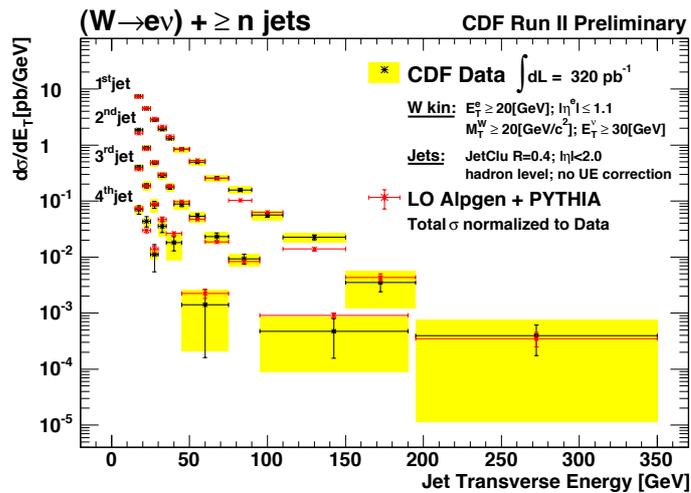
in the  $\geq 4$ -jet bin has been constructed from the 4 parton matrix element information, but a non-negligible number of events have been generated from a 2 or 3 parton final state, with the extra jets coming from the parton shower. In the right-hand plot, the jet multiplicity distribution is shown again, this time compared as well to the NLO (LO) prediction from MCFM for the 1, 2 (3) jet final states. The CKKW prescription agrees well with the NLO calculation for the jet multiplicities where it is available and agrees reasonably well with the Tevatron data for the range shown. Note that the production of each additional jet in this inclusive distribution is suppressed by a factor of the order of 0.2, or approximately  $\alpha_s$ .

A comparison of the measured cross sections for  $W + \geq n$  jets in CDF Run 2 as a function of the jet transverse momentum, to predictions from ALPGEN+PYTHIA is shown in Figure 58. The agreement is good. Note that this data is in a form (at the hadron level, corrected for detector effects) that makes it convenient for comparison to any hadron level Monte Carlo prediction §. Such a form should be the norm for measurements at both the Tevatron and LHC.

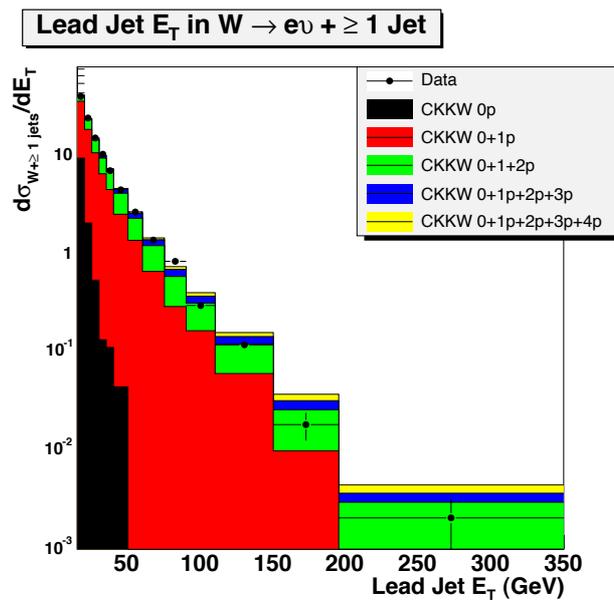
Comparisons with the NLO predictions of MCFM will be available in the near future. There is little change in normalization in going from LO to NLO predictions; as we saw in Section 3, the  $K$ -factor for these processes was close to unity. The major impact of the NLO corrections for the two highest  $p_T$  jets is to soften the distributions. The NLO calculation allows some of the momentum of the hard partons to be carried off by gluon radiation. A similar effect also occurs with the CKKW calculation where again there is the possibility for the parton momentum to be decreased by additional branchings. This is an instance of where parton showering contains some of the physics present in NLO calculations.

The transverse momentum distribution for the highest  $p_T$  jet in  $W +$  jets events at CDF is shown in Figure 59, along with the CKKW predictions. Both the total CKKW prediction and its decomposition into  $W + 1$  jet,  $W + 2$  jet,  $\dots$  contributions are indicated. As the transverse momentum of the lead jet increases, more contributions are seen to come from the higher jet multiplicity CKKW components. The exact subdivision of contributions will depend on where the cutoff was applied in the formation of the CKKW sample. Note that when the lead jet has a transverse momentum of the order of 200 GeV/c, the events are dominated by the higher multiplicity matrix elements, i.e. there is a large probability for the event to contain a second (or more) jet given the large  $p_T$  of the first. If the same transverse momentum cut were applied to all jets in the event, then the expectation would be that adding an additional jet would lead to a penalty of  $\alpha_s$  (as we have seen for the inclusive jet multiplicity distribution). In this case, since the additional jet is soft compared to the leading jet,  $\alpha_s$  is multiplied by a log relating the energies of the leading and additional jet. The log more than makes up for the factor of  $\alpha_s$ . Another way of looking at the jet multiplicity is to say that there

§ As mentioned before, the corrections for underlying event and for fragmentation basically cancel each other out for a cone of radius 0.4, so that the hadron level predictions are essentially parton level predictions as well



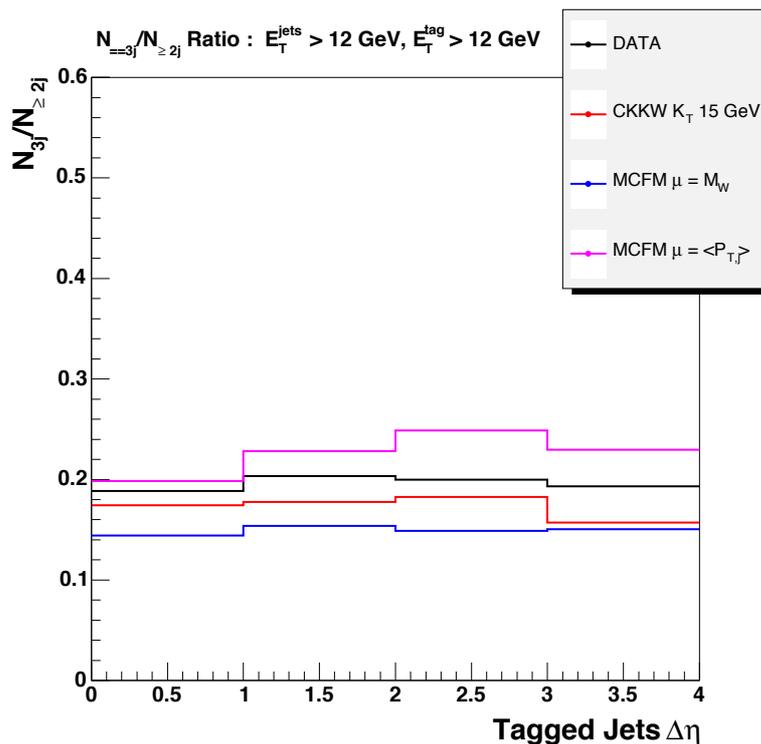
**Figure 58.** A comparison of the measured cross sections for  $W + \geq n$  jets in CDF Run 2 to predictions from ALPGEN+PYTHIA. The experimental cross sections have been corrected to the hadron level.



**Figure 59.** The  $E_T$  distribution of the lead jet in  $W + \geq 1$  jet events, along with the CKKW decomposition.

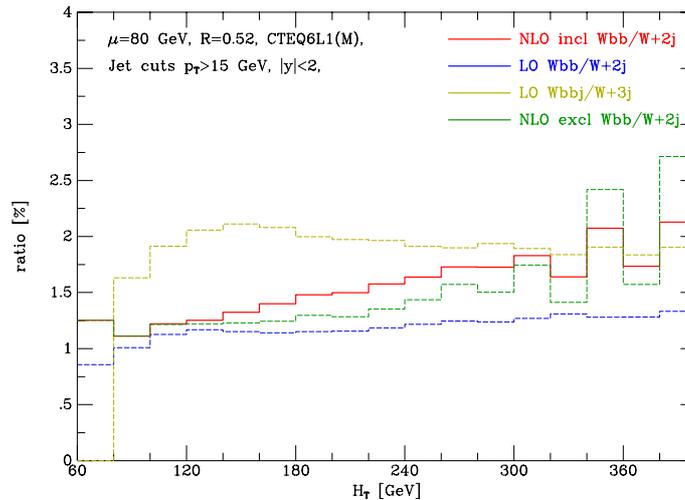
would be a Sudakov suppression on each of the incoming and outgoing parton legs if a requirement was imposed that no extra jet be produced. An estimate the approximate size of the Sudakov suppression using the figures given in Section 3.

A category of  $W/Z +$  jets event that has drawn particular interest lately is where two jets are produced separated by a large rapidity gap. Such a configuration (involving a  $Z$ ) serves as a background to vector boson fusion (VBF) production of a Higgs, where the two forward-backward jets serve to tag the event. As the Higgs is produced through



**Figure 60.** Predictions and a measurement from CDF Run 2 for the rate for the production of a third jet in  $W + \geq 2$  jet events, as a function of the rapidity separation of the two lead jets.

a colour-less exchange, the rate for an additional jet to be produced in the central region between the two tagging jets should be suppressed with respect to the QCD production of  $Z + \geq 2$  jets. The probability for an additional jet to be emitted in QCD  $W$  (rather than  $Z$  in order to obtain a higher rate)  $+2$  jet events, plus the ability of various theoretical predictions to describe this rate, is a measurement that can be carried out at the Tevatron prior to the turn-on of the LHC. Such a measurement is shown in Figures 60, where the rate for a 3rd jet to be emitted is shown versus the rapidity separation of the two tagging jets. It is evident that (1) the rate for a 3rd jet to be produced is large and (2) that the observed rate is in agreement with the CKKW predictions, and is bracketed by the predictions of MCFM for two choices of scale. Since the prediction is for  $W + 3$  jets, the MCFM calculation is at LO and retains a large scale dependence. The  $W/Z + 3$  jets process is one to which a high priority has been given for calculation to NLO, as will be discussed in Section 6. The rate for an additional jet to be emitted is roughly independent of the rapidity separation of the two tagging jets. The agreement of the data with the CKKW predictions is heartening for two reasons: (1) it indicates that CKKW predictions will most likely provide accurate predictions for similar topologies at the LHC and (2) the rate for additional jet production in  $W/Z + 2$  widely separated jet events is high, leading to an effective veto in VBF Higgs searches at the LHC.



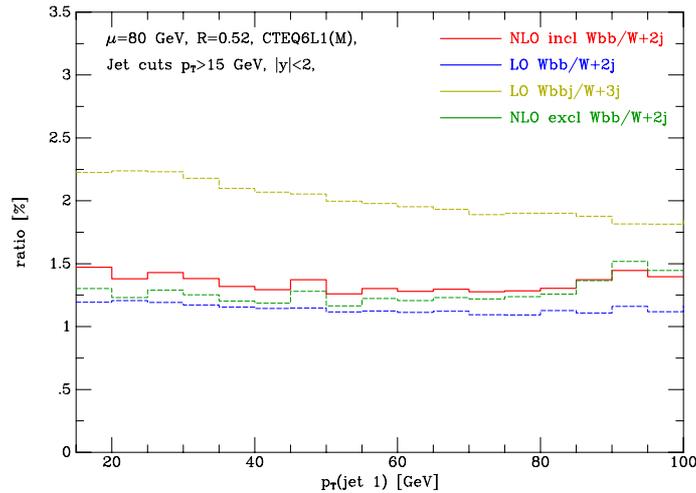
**Figure 61.** The ratio of the cross sections for  $Wb\bar{b}$  and  $Wjj$  is plotted at LO and NLO as a function of the variable  $H_T$ . The ratio is observed to have a different slope at NLO than at LO.

For many of the analyses at the Tevatron, it is useful to calculate the rate of leading order parton shower Monte Carlo predictions. For example, the Method 2 technique in CDF’s top analysis uses the calculated ratio of  $[Wb\bar{b} + (n - 2) \text{ jets}]/[W + n \text{ jets}]$  (for  $n = 1, 2$ ) and the measured rate for  $W + n$  jets to calculate the  $Wb\bar{b} + (n - 2)$  jet background to top production. The ratio of the two processes should be more reliable than the absolute leading order prediction for  $Wb\bar{b} + (n - 2)$  jets. The above statement should be true as long as the NLO corrections (the  $K$ -factors) do not have an appreciable difference in shape between the two processes. As discussed previously, NLO corrections are available for  $Wb\bar{b}$  and  $Wjj$  but not for higher jet multiplicities. A recent study has shown that the relative shape of the two processes is indeed different at NLO if highly exclusive variables (such as  $H_T$ , the sum of the transverse energy of all final state objects in the event) are used for the measurement [129]. This is shown in Figure 61. Inclusive variables, such as the transverse momentum of the leading jet seem to be safe in this regard, as can be seen in Figure 62. There is some indication that the change in shape observed in going from LO to NLO can also be reproduced by including higher jet multiplicities using the CKKW procedure.

## 6. Benchmarks for the LHC

### 6.1. Introduction

Scattering at the LHC is not simply *rescaled* scattering at the Tevatron. There are small typical momentum fractions  $x$  for many of the key processes; thus, there is a dominance of sea quark and gluon scattering as compared to valence quark scattering at the Tevatron. There is a large phase space for gluon emission and thus intensive QCD backgrounds for many of the signatures of new physics. Many of the scales relating to



**Figure 62.** The ratio of the cross sections for  $Wb\bar{b}$  and  $Wjj$  is plotted at LO and NLO as a function of the  $p_T$  of the leading jet. The ratio is observed to have a similar slope at NLO as at LO.

interesting processes are large compared to the  $W$  mass; thus, electroweak corrections can become important even for nominally QCD processes. In this section, we will try to provide some useful benchmarks for LHC predictions.

## 6.2. Parton-parton luminosities at the LHC

||

It is useful to introduce the idea of differential parton-parton luminosities. Such luminosities, when multiplied by the dimensionless cross section  $\hat{\sigma}$  for a given process, provide a useful estimate of the size of an event cross section at the LHC. Below we define the differential parton-parton luminosity  $dL_{ij}/d\hat{s} dy$  and its integral  $dL_{ij}/d\hat{s}$ :

$$\frac{dL_{ij}}{d\hat{s} dy} = \frac{1}{s} \frac{1}{1 + \delta_{ij}} [f_i(x_1, \mu) f_j(x_2, \mu) + (1 \leftrightarrow 2)] . \quad (41)$$

The prefactor with the Kronecker delta avoids double-counting in case the partons are identical. The generic parton-model formula

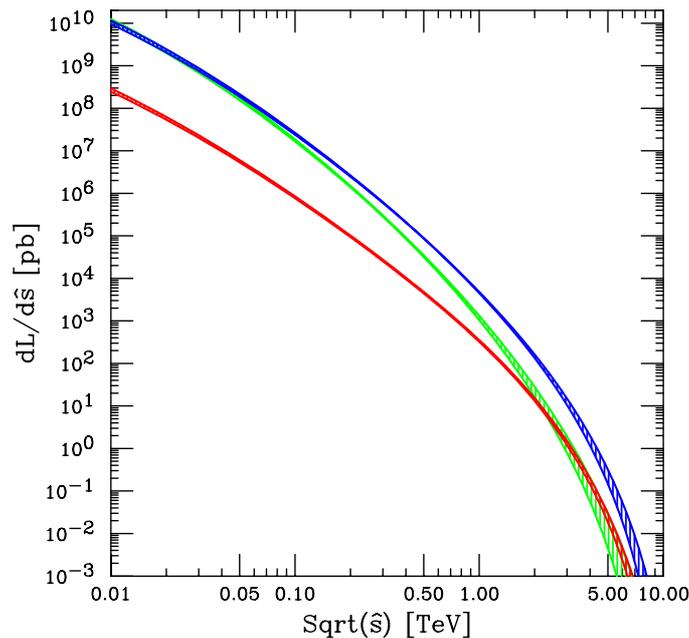
$$\sigma = \sum_{i,j} \int_0^1 dx_1 dx_2 f_i(x_1, \mu) f_j(x_2, \mu) \hat{\sigma}_{ij} \quad (42)$$

can then be written as

$$\sigma = \sum_{i,j} \int \left( \frac{d\hat{s}}{\hat{s}} dy \right) \left( \frac{dL_{ij}}{d\hat{s} dy} \right) (\hat{\sigma}_{ij}) . \quad (43)$$

(Note that this result is easily derived by defining  $\tau = x_1 x_2 = \hat{s}/s$  and observing that the Jacobian  $\partial(\tau, y)/\partial(x_1, x_2) = 1$ .)

|| Parts of this discussion also appeared in a contribution to the Les Houches 2005 proceedings [130] by A. Belyaev, J. Huston and J. Pumplin



**Figure 63.** The parton-parton luminosity  $\left[\frac{1}{\hat{s}} \frac{dL_{ij}}{d\hat{r}}\right]$  in picobarns, integrated over  $y$ . Green= $gg$ , Blue= $g(d + u + s + c + b) + g(\bar{d} + \bar{u} + \bar{s} + \bar{c} + \bar{b}) + (d + u + s + c + b)g + (\bar{d} + \bar{u} + \bar{s} + \bar{c} + \bar{b})g$ , Red= $d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$ .

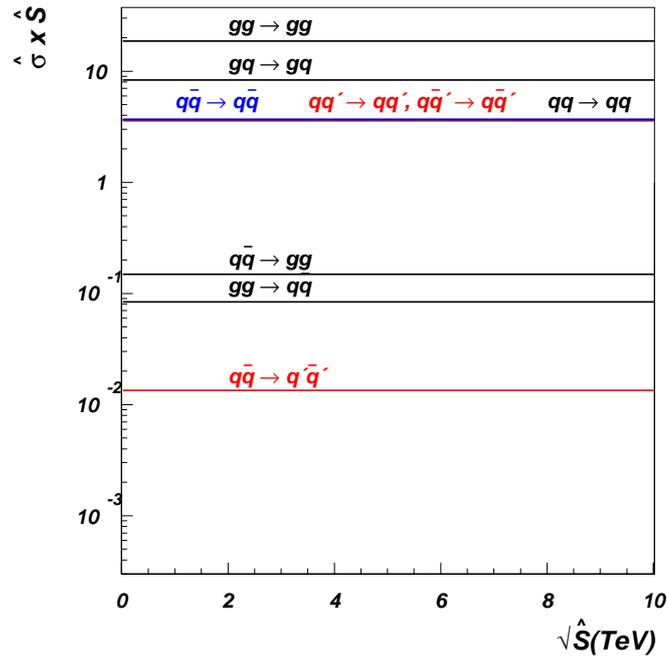
(43) can be used to estimate the production rate for a hard scattering process at the LHC as follows. Figure 63 shows a plot of the luminosity function integrated over rapidity,  $dL_{ij}/d\hat{s} = \int (dL_{ij}/d\hat{s} dy) dy$ , at the LHC  $\sqrt{s} = 14$  TeV for various parton flavour combinations, calculated using the CTEQ6.1 parton distribution functions [6]. The widths of the curves indicate an estimate for the PDF uncertainties. We assume  $\mu = \sqrt{\hat{s}}$  for the scale  $\mathfrak{M}$ . As expected, the  $gg$  luminosity is large at low  $\sqrt{\hat{s}}$  but falls rapidly with respect to the other parton luminosities. The  $gq$  luminosity is large over the entire kinematic region plotted.

Figures 64 and 65 present the second product,  $[\hat{s}\hat{\sigma}_{ij}]$ , for various  $2 \rightarrow 2$  partonic processes with massless and massive partons in the final state respectively. The parton level cross sections have been calculated for a parton  $p_T > 0.1 \times \sqrt{\hat{s}}$  cut and for fixed  $\alpha_s = 0.118$  using the CalcHEP package [131]. For the case of massive partons in the final state, there is a threshold behavior not present with massless partons. Note also that the threshold behavior is different for  $q\bar{q}$  and  $gg$  initial states. The  $gg$  processes can proceed through the  $t$ -channel as well as the  $s$ -channel and this is responsible for the extra structure.

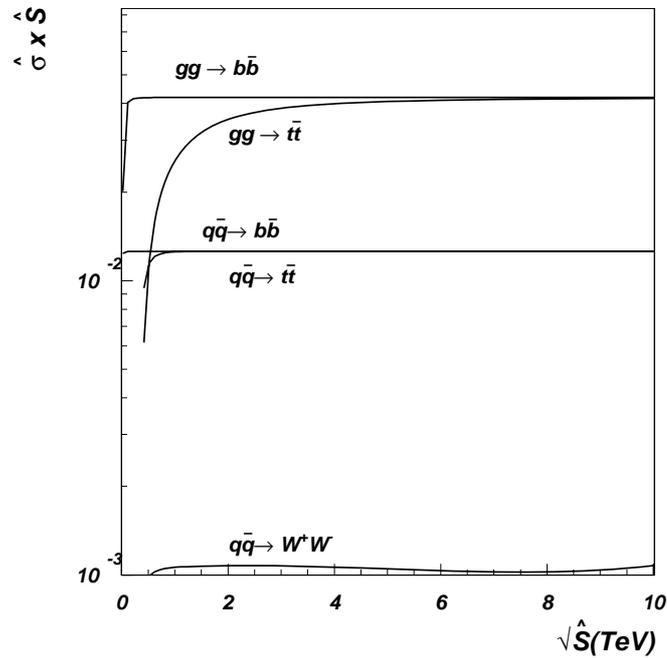
The products  $[\hat{s}\hat{\sigma}_{ij}]$  are plotted for massless and massive final state partons as a function of the ratio  $p_T/\sqrt{\hat{s}}$  in Figures 66 and 67. One can use (43) in the form

$$\sigma = \frac{\Delta\hat{s}}{\hat{s}} \left( \frac{dL_{ij}}{d\hat{s}} \right) (\hat{s}\hat{\sigma}_{ij}). \quad (44)$$

$\mathfrak{M}$  Similar plots made with earlier PDFs are shown in Ellis, Stirling, Webber [3]

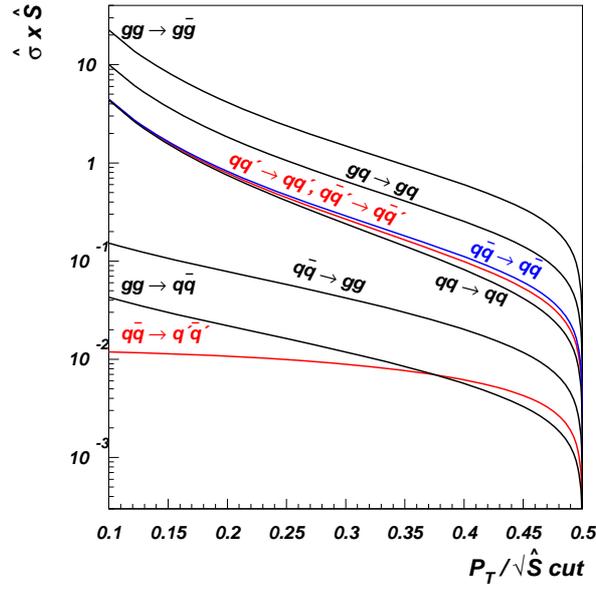


**Figure 64.** Parton level cross sections ( $\hat{\sigma}_{ij}$ ) for various processes involving massless partons in the final state.

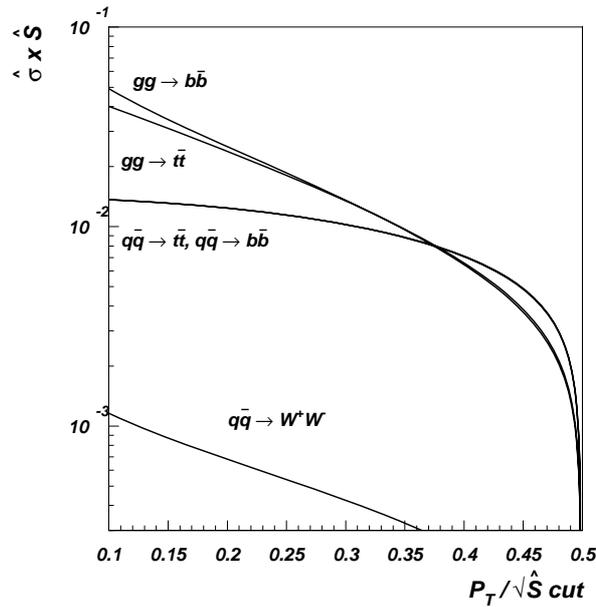


**Figure 65.** Parton level cross sections ( $\hat{\sigma}_{ij}$ ) for various processes involving massive partons in the final state.

and Figures 64, 66, 65, 67 to estimate the QCD production cross sections for a given  $\Delta\hat{s}$

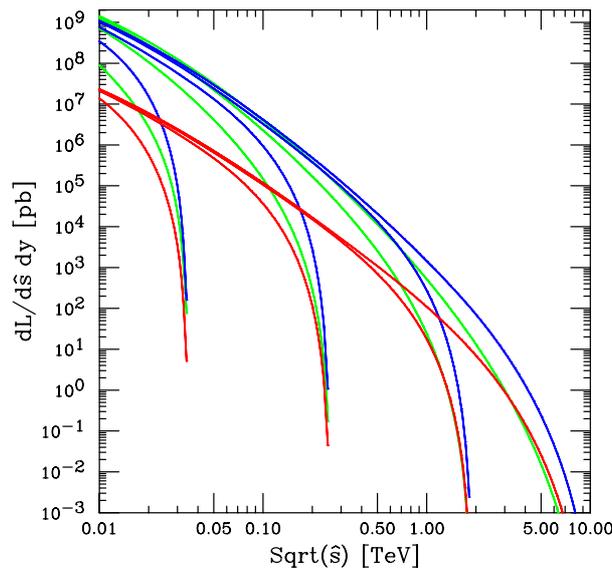


**Figure 66.** Parton level cross sections ( $\hat{\sigma}_{ij}$ ) for various processes involving massless partons in the final state as a function of the variable  $p_T/\sqrt{\hat{s}}$ .



**Figure 67.** Parton level cross sections ( $\hat{\sigma}_{ij}$ ) for various processes involving massive partons in the final state as a function of the variable  $p_T/\sqrt{\hat{s}}$ . The calculations were performed for  $\sqrt{\hat{s}} = 2$  TeV, in the region where  $\hat{\sigma}x\hat{s}$  has a relatively flat behavior with  $\sqrt{\hat{s}}$ .

interval and a particular value of  $p_T/\sqrt{\hat{s}}$ . For example, for the  $gg \rightarrow gg$  rate for  $\hat{s}=1$  TeV and  $\Delta\hat{s} = 0.01\hat{s}$ , we have  $dL_{gg}/d\hat{s} \simeq 10^3$  pb and  $\hat{s}\hat{\sigma}_{gg} \simeq 20$  leading to  $\sigma \simeq 200$  pb (for



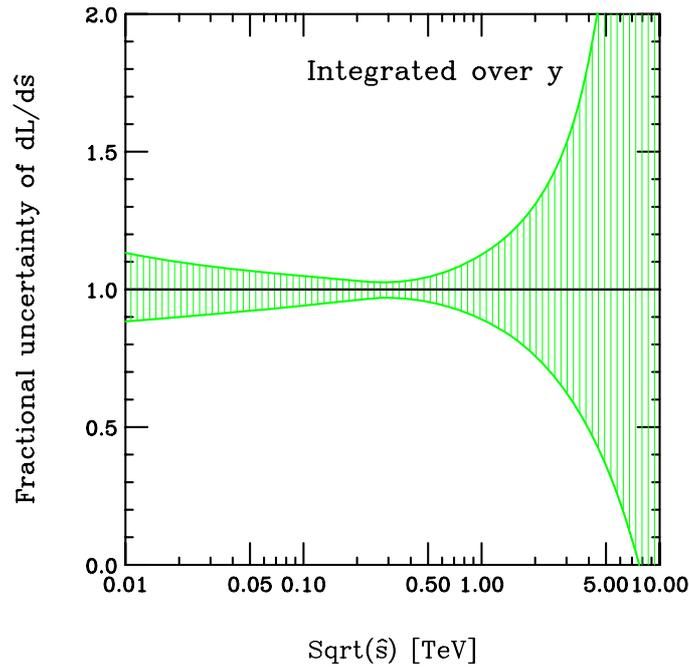
**Figure 68.**  $dL_{\text{Luminosity}}/dy$  at rapidities  $y = 0, 2, 4, 6$ . Green= $gg$ , Blue= $g(d + u + s + c + b) + g(\bar{d} + \bar{u} + \bar{s} + \bar{c} + \bar{b}) + (d + u + s + c + b)g + (\bar{d} + \bar{u} + \bar{s} + \bar{c} + \bar{b})g$ , Red= $d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$ .

the  $p_T^q > 0.1 \times \sqrt{\hat{s}}$  cut we have used above). Note that for a given small  $\Delta\hat{s}/\hat{s}$  interval, the corresponding invariant mass  $\Delta\sqrt{\hat{s}}/\sqrt{\hat{s}}$  interval, is  $\Delta\sqrt{\hat{s}}/\sqrt{\hat{s}} \simeq \frac{1}{2}\Delta\hat{s}/\hat{s}$ . One should also mention that all hard cross sections presented in Figure 64 are proportional to  $\alpha_s^2$  and have been calculated for  $\alpha_s = 0.118$ , so production rates can be easily rescaled for a certain  $\alpha_s$  at a given scale.

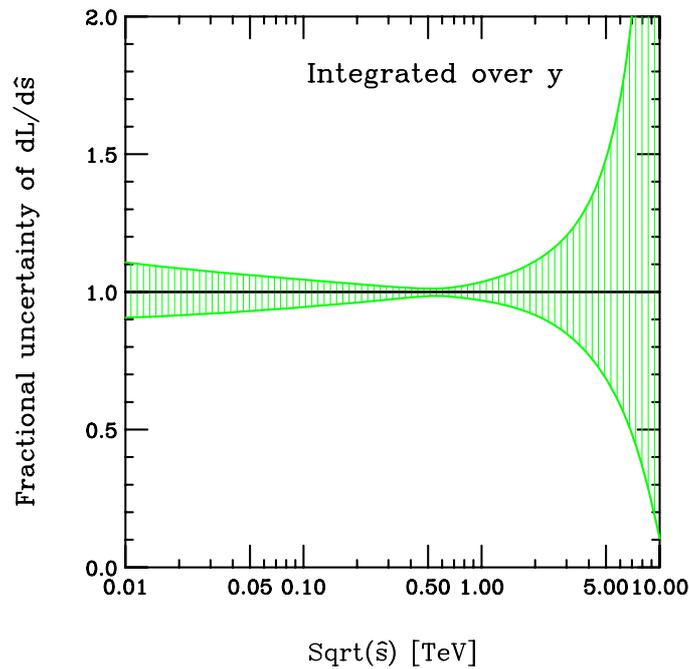
One can further specify the parton-parton luminosity for a specific rapidity  $y$  and  $\hat{s}$ ,  $dL_{ij}/d\hat{s} dy$ . If one is interested in a specific partonic initial state, then the resulting differential luminosity can be displayed in families of curves as shown in Figure 68, where the differential parton-parton luminosity at the LHC is shown as a function of the subprocess center-of-mass energy  $\sqrt{\hat{s}}$  at various values of rapidity for the produced system for several different combinations of initial state partons. One can read from the curves the parton-parton luminosity for a specific value of mass fraction and rapidity. (It is also easy to use the Durham PDF plotter to generate the pdf curve for any desired flavour and kinematic configuration <sup>+</sup>.)

It is also of great interest to understand the uncertainty for the parton-parton luminosity for specific kinematic configurations. Some representative parton-parton luminosity uncertainties, integrated over rapidity, are shown in Figures 69, 70 and 71. The PDF uncertainties were generated from the CTEQ6.1 Hessian error analysis using the standard  $\Delta\chi^2 = 100$  criterion. Except for kinematic regions where one or both partons is a gluon at high  $x$ , the pdf uncertainties are of the order of 5–10%. Luminosity uncertainties for specific rapidity values are available at the SM benchmark website. Even tighter constraints will be possible once the LHC Standard Model data is included

<sup>+</sup> <http://durpdg.dur.ac.uk/hepdata/pdf3.html>

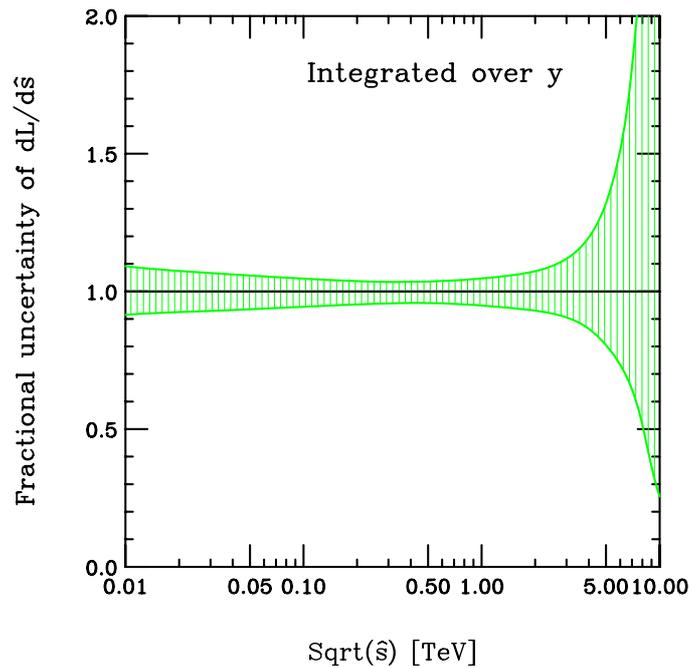


**Figure 69.** Fractional uncertainty of the  $gg$  luminosity integrated over  $y$ .



**Figure 70.** Fractional uncertainty for the parton-parton luminosity integrated over  $y$  for  $g(d+u+s+c+b) + g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b}) + (d+u+s+c+b)g + (\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$ .

in the global pdf fits. Again, the uncertainties for individual PDF's can also be calculated online using the Durham pdf plotter. Often it is not the pdf uncertainty for a cross section that is required, but rather the pdf uncertainty for an acceptance for a given

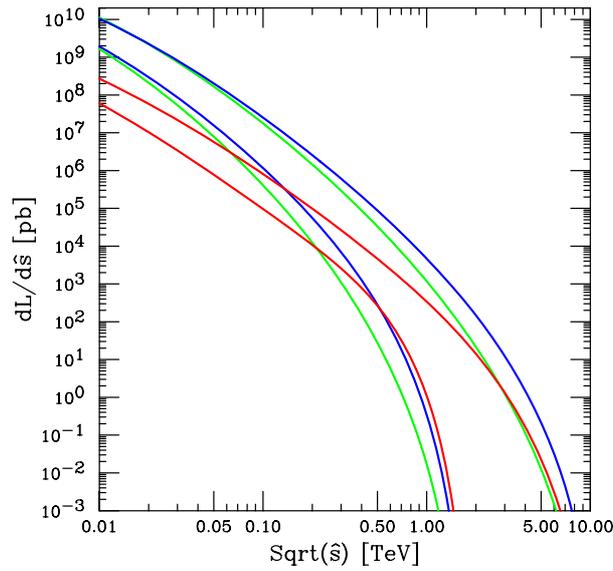


**Figure 71.** Fractional uncertainty for the luminosity integrated over  $y$  for  $d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$ .

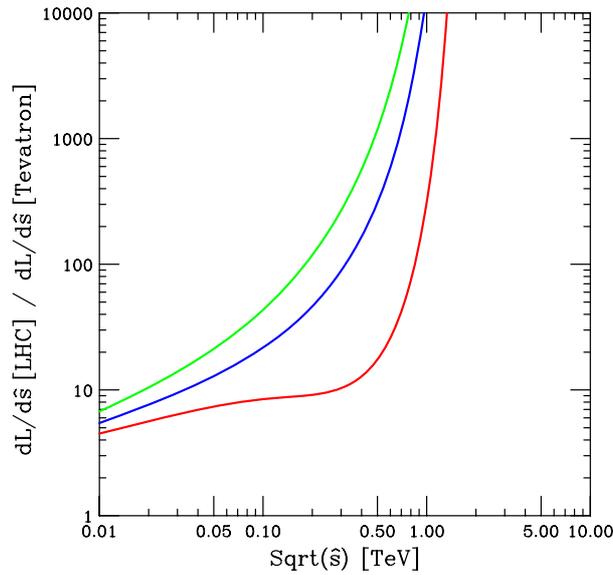
final state. The acceptance for a particular process may depend on the input pdf's due to the rapidity cuts placed on the jets, leptons, photons, etc and the impacts of the varying longitudinal boosts of the final state caused by the different pdf pairs. An approximate “rule of thumb” is that the pdf uncertainty for the acceptance is an order of magnitude smaller than the uncertainty for the cross section itself.

In Figure 72, the pdf luminosity curves shown in Figure 63 are overlaid with equivalent luminosity curves from the Tevatron. In Figure 73, the ratios of the pdf luminosities at the LHC to those at the Tevatron are plotted. The most dramatic increase in pdf luminosity at the LHC comes from  $gg$  initial states, followed by  $gq$  initial states and then  $q\bar{q}$  initial states. The latter ratio is smallest because of the availability of valence anti-quarks at the Tevatron at moderate to large  $x$ . Consider chargino pair production at a  $\sqrt{\hat{s}}$  of 0.4 TeV. This process proceeds through  $q\bar{q}$  annihilation; thus, there is only a factor of 10 enhancement at the LHC compared to the Tevatron. Top pair production, on the other hand, proceeds both through  $q\bar{q}$  annihilation and  $gg$  scattering. If we consider a specific value of  $\sqrt{\hat{s}}$  of 0.4 TeV (near  $t\bar{t}$  threshold), the  $q\bar{q}$  annihilation component again is only a factor of 10 larger at the LHC than at the Tevatron. The  $gg$  component, on the other hand, is over a factor of 500 larger, leading to (1) the large dominance of  $gg$  scattering for top pair production at the LHC, in contrast to the situation at the Tevatron and (2) a total  $t\bar{t}$  cross section a factor of 100 larger than at the Tevatron.

Backgrounds to interesting physics at the LHC proceed mostly through  $gg$  and  $gq$  initial states. Thus, there will be a commensurate increase in the rate for background

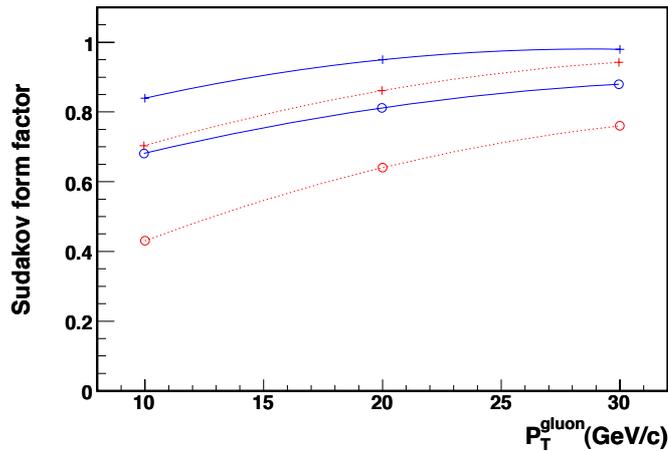


**Figure 72.** The parton-parton luminosity  $\left[\frac{1}{\hat{s}} \frac{dL_{ij}}{d\tau}\right]$  in pb integrated over  $y$ . Green= $gg$ , Blue= $g(d+u+s+c+b) + g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b}) + (d+u+s+c+b)g + (\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$ , Red= $d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$ . The top family of curves are for the LHC and the bottom for the Tevatron.



**Figure 73.** The ratio of parton-parton luminosity  $\left[\frac{1}{\hat{s}} \frac{dL_{ij}}{d\tau}\right]$  in pb integrated over  $y$  at the LHC and Tevatron. Green= $gg$  (top), Blue= $g(d+u+s+c+b) + g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b}) + (d+u+s+c+b)g + (\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$  (middle), Red= $d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$  (bottom).

processes at the LHC. For this reason, the jet cuts applied to  $t\bar{t}$  analyses at the LHC are set a factor of 2 larger than at the Tevatron, in order to reduce the backgrounds from  $W + 4$  jet production, relative to the lepton + 4 jets final state from  $t\bar{t}$  decay. The

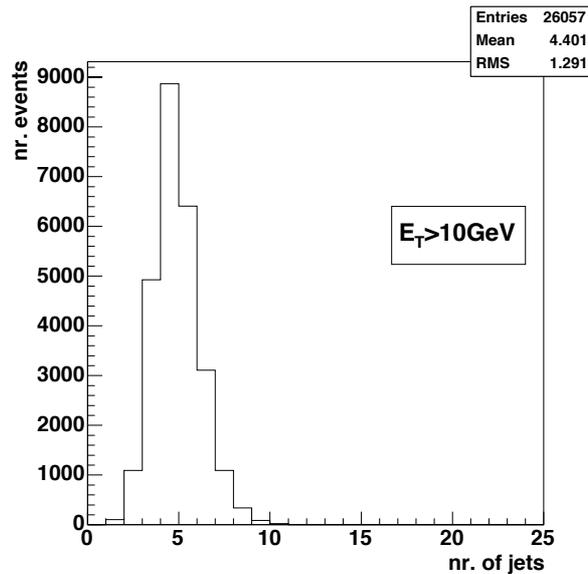


**Figure 74.** The Sudakov form factors for initial state quarks and gluons at a hard scale of 200 GeV/c as a function of the transverse momentum of the emitted gluon. The form factors are for quarks (blue-solid) and gluons (red-dashed) at parton  $x$  values of 0.3 (crosses) and 0.03 (open circles).

larger  $t\bar{t}$  cross section and the higher luminosity at the LHC make up for the lower event efficiency for the  $t\bar{t}$  analysis.

### 6.3. Sudakovs for the LHC

As discussed, many of the interesting processes at the LHC will proceed through  $gg$  scattering at low  $x$ . This will have an impact on the number of extra jets produced through initial state radiation. It is informative to compare the Sudakov form factors for  $t\bar{t}$  production at the Tevatron and LHC. At the Tevatron,  $t\bar{t}$  production proceeds primarily (85%) through  $q\bar{q}$  with  $gg$  being responsible for 15%, with the partons near an average  $x$  value of 0.3. At the LHC, the percentages are roughly reversed (or more precisely 90% for  $gg$ ) and the scattering takes place at an average  $x$  value of a factor of 7 lower (which we approximate here as  $x = 0.03$ ). The relevant Sudakov form factors are shown in Figure 74, as a function of the transverse momentum of the emitted gluon, at a hard scale of 200 GeV/c (roughly appropriate for  $t\bar{t}$  production). It is evident that because of the higher percentage of  $gg$  production and the lower average  $x$  of the incident partons, that the jet multiplicity will be significantly higher for  $t\bar{t}$  production at the LHC than at the Tevatron. We can make some rough estimates from these plots. The probability for no gluon of 10 GeV/c or greater to be radiated from an initial quark leg with  $x=0.3$  is 0.85. The probability for no such gluon to be radiated from either quark leg at the Tevatron is  $0.85 \times 0.85 = 0.72$ , i.e. a 0.28 chance of radiating such a gluon. A similar exercise for two incident gluons of  $x=0.3$  gives a chance of radiating a 10 GeV/c gluon of 0.51. As the  $q\bar{q}$  initial state makes up 85% of the Tevatron cross section, with  $gg$  15%, the total probability of emitting at least one 10 GeV/c gluon is 0.3. Using 90% for  $gg$  at the LHC and 10% for  $q\bar{q}$ , gives a 0.8 probability of radiating



**Figure 75.** The jet multiplicity in  $t\bar{t}$  events with a lepton + jets final state at the LHC. A cut of 10 GeV has been applied to the jets.

such a hard gluon. This is borne out in Figure 75 where the jet multiplicity in  $t\bar{t}$  events at the LHC from a CMS Monte Carlo study is shown, with a jet transverse energy cut of 10 GeV/c<sup>\*</sup>. The probability for additional jets is large ‡ For this reason, and because of the likelihood of jets being produced from the underlying event, the jet thresholds will need to be increased for analyses at the LHC, compared to those at the Tevatron.

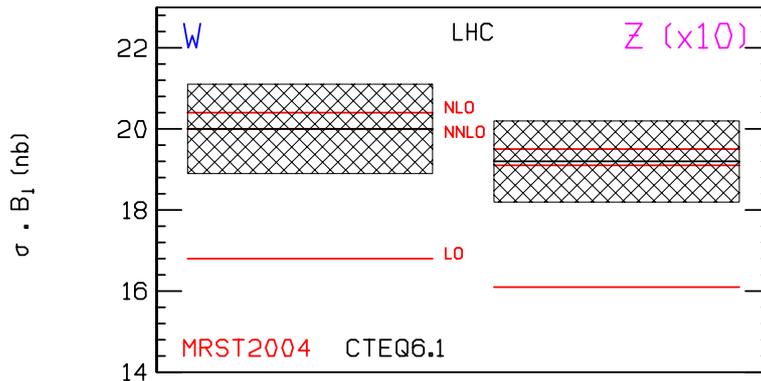
#### 6.4. Stability of NLO global analyses

The W and Z cross sections at the LHC are shown in Figure 76. The MRST2004 predictions are shown at LO, NLO and NNLO; also shown are the CTEQ6.1 predictions at NLO, along with the CTEQ6.1 pdf error band. There is a significant increase in cross section when going from LO to NLO, and then a small decrease when going from NLO to NNLO. The NLO MRST2004 and CTEQ6.1 predicted cross sections agree with each other, well within the CTEQ6.1 pdf uncertainty band.

Most of the absolute predictions for LHC observables have been carried out at NLO, i.e. with NLO pdf's and with NLO matrix elements. Such predictions have worked well at the Tevatron, but the LHC explores a new region of  $x$  for a hadron-hadron collider. Thus, it is important to understand whether the NLO formalism carries over to the LHC with the same degree of accuracy. In recent years, some preliminary next-to-next-leading-order (NNLO) global pdf analyses have been carried out either for DIS alone [132], or in a global analysis context [134]. The differences with respect to the

\* We would like to thank Jan Heyninck for providing this plot

‡ Note that some of the radiated gluons will be outside the rapidity cuts of the top analysis; some of the jets from the  $t\bar{t}$  decay may also either be outside the rapidity cuts or below the transverse energy threshold.

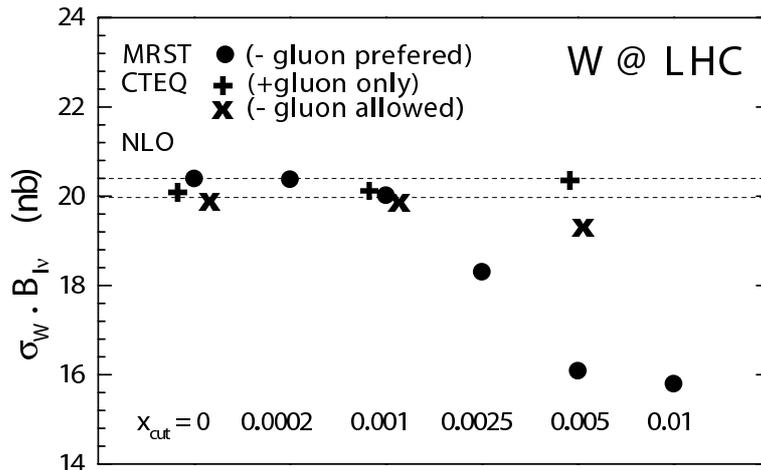


**Figure 76.** Predicted cross sections for  $W$  and  $Z$  production at the LHC using MRST2004 and CTEQ6.1 PDF's. The overall PDF uncertainty of the NLO CTEQ6.1 prediction is approximately 5%, consistent with Figure 71.

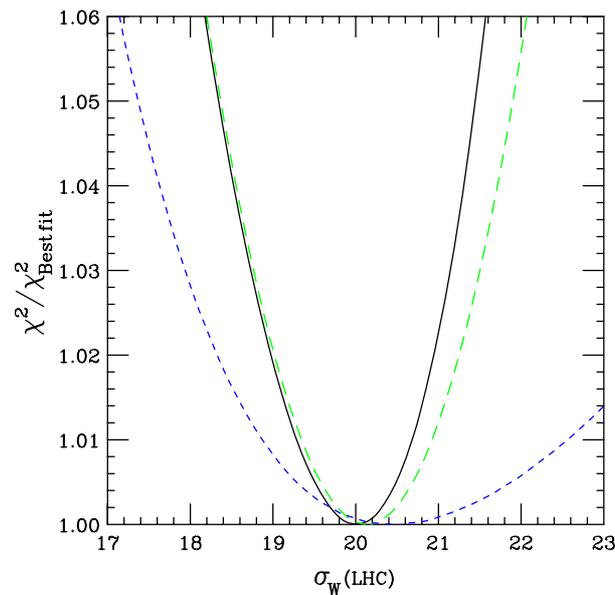
corresponding NLO analyses are reasonably small. However, most observables of interest (including inclusive jet production) have not yet been calculated at NNLO.

All other considerations being equal, a global analysis at NNLO must be expected to have a higher accuracy. However, NLO analyses can be adequate as long as their accuracy is sufficient for the task, and as long as their predictions are stable with respect to certain choices inherent in the analysis. Examples of those choices are the functional forms used to parametrize the initial nonperturbative parton distribution functions, and the selection of experimental data sets included in the fit — along with the kinematic cuts that are imposed on that data. In a recent MRST analysis [133], a 20% variation in the cross section predicted for  $W$  production at the LHC — a very important “standard candle” process for hadron colliders — was observed when data at low  $x$  and  $Q^2$  were removed from the global fit. This is a fundamental consideration, since the presence of this instability would significantly impact the phenomenology of a wide range of physical processes for the LHC. The instability was removed in the same study when the analysis was carried out at NNLO.

A similar study was also carried out by CTEQ in which this instability was not observed, and the predictions for the  $W$  cross section at the LHC remained stable (but with increased uncertainties) when more severe low  $(x, Q^2)$  kinematic cuts were performed [135]. The predicted cross sections, as a function of the  $x$  cut on the data removed, are shown in Figure 77 for the two studies. In Figure 78, the uncertainty on the predictions for the  $W$  cross section at the LHC is shown from a study that uses the Lagrange Multiplier Method. As more data are removed at low  $x$  and  $Q^2$ , the resulting uncertainty on the predicted  $W$  cross section increases greatly. If the gluon is parameterized in a form that allows it to become negative at low  $x$ , such as the parameterization that MRST used, then the uncertainty becomes even larger, especially for the lower range of the cross section.



**Figure 77.** Predicted total cross section of  $W^+ + W^-$  production at the LHC for the fits obtained in the CTEQ stability study, compared to the MRST results. The overall PDF uncertainty of the prediction is  $\sim 5\%$ , as observed in Figure 71.



**Figure 78.** Lagrange multiplier results for the  $W$  cross section (in nb) at the LHC using a positive-definite gluon. The three curves, in order of decreasing steepness, correspond to three sets of kinematic cuts, standard/intermediate/strong.

### 6.5. The future for NLO calculations

Unfortunately, our ability to take advantage of the benefits of NLO predictions is severely limited by the calculations available. Currently, no complete NLO QCD calculation exists for a process involving more than 5 particles. This means that we are limited to the consideration of scenarios such as 4 jet production at LEP (effectively a  $1 \rightarrow 4$  process) and a number of  $2 \rightarrow 3$  processes of interest at hadron colliders.

The bottleneck for these calculations is the evaluation of the loop diagrams. The real radiation component consists of the evaluation of lowest order matrix elements (known) combined with a method for systematically removing their singularities, several of which are well understood. In contrast, the evaluation of the virtual contribution must at present be performed on a case-by-case basis. Even the most basic integral that can appear depends non-trivially on the number of external legs in the loop, as well as the masses of all the external and internal particles. To complicate matters further, the presence of powers of loop momenta in the integrals (from fermion propagators or multi-gluon vertices) leads to “tensor integrals” which typically are expressed as the sum of many terms. Taken all together, this means that traditional methods based on an analytical Feynman diagram approach are already limited by the computing power currently available. It seems unlikely that further progress, such as the NLO calculation of a  $2 \rightarrow 4$  process, will be made in this fashion.

An alternative to this brute-force Feynman diagram approach is provided by “unitarity-based” methods, which are based on sewing together tree level amplitudes (for a review, see [136]). These methods rely on the Cutkosky rules to generate the 1-loop amplitude by summing over all possible cuts of sewn-together tree amplitudes. Although this knowledge in itself only suffices for sufficiently supersymmetric theories, results for 1-loop amplitudes in non-supersymmetric theories (such as QCD) can be obtained by consideration of collinear limits of the amplitudes. Such techniques were used to obtain relatively compact expressions for the 1-loop amplitudes required to evaluate the process  $e^+e^- \rightarrow 4$  partons [47].

More recently the “twistor-inspired” methods, which have proven so efficient at generating compact expressions for tree-level amplitudes, have been considered at the 1-loop level [137]. The MHV rules were shown to reproduce known results at the 1-loop level [138] and their application has passed from calculations in  $N = 4$  supersymmetric Yang-Mills theory to the evaluation of amplitudes in QCD ([139] and references therein). Similarly, the on-shell recursion relations have proven equally useful at one-loop [140], even allowing for the calculation of MHV amplitudes involving an arbitrary number of gluons [141]. These methods will doubtless be refined and extended to more complicated calculations over time.

An alternative path to 1-loop amplitudes is provided by numerical approaches. Such techniques have long been applied at lowest order, with the main attraction deriving from the promise of a generic solution that can be applied to any process (given sufficient computing power). The techniques are generally either completely numerical evaluations of loop integrals [142, 143], or they are semi-numerical approaches that perform only the tensor integral reductions numerically [144, 145, 146, 147, 148, 149]. In most cases these algorithms have only been applied to individual integrals or diagrams, although one recent application evaluates the 1-loop matrix elements for the Higgs +4 parton process (which is otherwise uncalculated) [150] ††. Once more, the study of these approaches is

†† All partons are treated as outgoing so this process corresponds to a Higgs +2 parton final state

in its infancy and the coming years promise significant further development.

### 6.6. A realistic NLO wishlist for multi-parton final states at the LHC

A somewhat whimsical experimenter's *wishlist* was first presented at the Run 2 Monte Carlo workshop at Fermilab in 2001 †. Since then the list has gathered a great deal of notoriety and has appeared in numerous LHC-related theory talks. It is unlikely that  $WWW + b\bar{b} + 3\text{jets}$  will be calculated at NLO soon, no matter the level of physics motivation, but there are a number of high priority calculations, primarily of backgrounds to new physics, that are accessible with the present technology. However, the manpower available before the LHC turns on is limited. Thus, it is necessary to prioritize the calculations, both in terms of the importance of the calculation and the effort expected to bring it to completion.

A prioritized list, determined at the Les Houches 2005 Workshop, is shown in Table 2, along with a brief discussion of the physics motivation. Note that the list contains only  $2 \rightarrow 3$  and  $2 \rightarrow 4$  processes, as these are feasible to be completed by the turn on of LHC.

| process<br>( $V \in \{Z, W, \gamma\}$ )     | relevant for   |
|---|--|
| 1. $pp \rightarrow V V + \text{jet}$        | $t\bar{t}H$ , new physics                                      |
| 2. $pp \rightarrow H + 2\text{jets}$        | $H$ production by vector boson fusion (VBF)                    |
| 3. $pp \rightarrow t\bar{t} b\bar{b}$       | $t\bar{t}H$  |
| 4. $pp \rightarrow t\bar{t} + 2\text{jets}$ | $t\bar{t}H$  |
| 5. $pp \rightarrow V V b\bar{b}$            | VBF $\rightarrow H \rightarrow VV$ , $t\bar{t}H$ , new physics |
| 6. $pp \rightarrow V V + 2\text{jets}$      | VBF $\rightarrow H \rightarrow VV$                             |
| 7. $pp \rightarrow V + 3\text{jets}$        | various new physics signatures                                 |
| 8. $pp \rightarrow V V V$                   | SUSY trilepton   |

**Table 2.** The wishlist of processes for which a NLO calculation is both desired and feasible in the near future.

First, a few general statements: in general, signatures for new physics will involve high  $p_T$  leptons and jets (especially  $b$  jets) and missing transverse momentum. Thus, backgrounds to new physics will tend to involve (multiple) vector boson production (with jets) and  $t\bar{t}$  pair production (with jets). The best manner in which to understand the normalization of a cross section is to measure it; however the rates for some of the complex final states listed here may be limited and (at least in the early days) must be

† <http://vmsstreamer1.fnal.gov/Lectures/MonteCarlo2001/Index.htm>

known from NLO theory. NLO is the first order at which both the normalization and shape can be calculated with any degree of confidence.

We now turn to a detailed look at each of these processes:

- $pp \rightarrow VV + \text{jet}$ : One of the most promising channels for Higgs production in the low mass range is through the  $H \rightarrow WW^*$  channel, with the W's decaying semi-leptonically. It is useful to look both in the  $H \rightarrow WW$  exclusive channel, along with the  $H \rightarrow WW + \text{jet}$  channel. The calculation of  $pp \rightarrow WW + \text{jet}$  will be especially important in understanding the background to the latter.
- $pp \rightarrow H + 2 \text{ jets}$ : A measurement of vector boson fusion (VBF) production of the Higgs boson will allow the determination of the Higgs coupling to vector bosons. One of the key signatures for this process is the presence of forward-backward tagging jets. Thus, QCD production of  $H + 2 \text{ jets}$  must be understood, especially as the rates for the two are comparable in the kinematic regions of interest.
- $pp \rightarrow t\bar{t}b\bar{b}$  and  $pp \rightarrow t\bar{t} + 2 \text{ jets}$ : Both of these processes serve as background to  $t\bar{t}H$ , where the Higgs decays into a  $b\bar{b}$  pair. The rate for  $t\bar{t}jj$  is much greater than that for  $t\bar{t}b\bar{b}$  and thus, even if 3  $b$ -tags are required, there may be a significant chance for the heavy flavour mistag of a  $t\bar{t}jj$  event to contribute to the background.
- $pp \rightarrow VVb\bar{b}$ : Such a signature serves as non-resonant background to  $t\bar{t}$  production as well as to possible new physics.
- $pp \rightarrow VV + 2 \text{ jets}$ : The process serves as a background to VBF production of Higgs.
- $pp \rightarrow V + 3 \text{ jets}$ : The process serves as background for  $t\bar{t}$  production where one of the jets may not be reconstructed, as well as for various new physics signatures involving leptons, jets and missing transverse momentum.
- $pp \rightarrow VVV$ : The process serves as a background for various new physics subprocesses such as SUSY tri-lepton production.

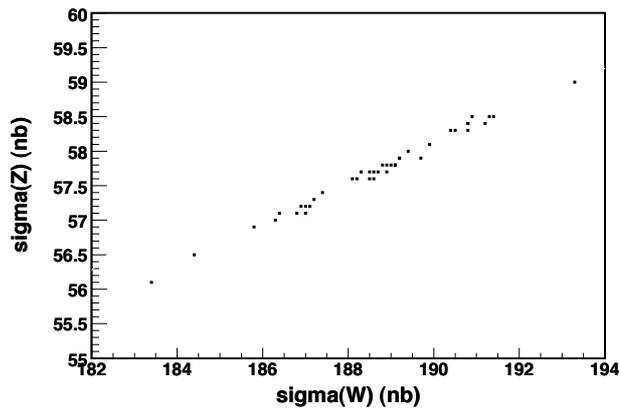
Work on (at least) the processes 1. to 4. of Table 2 is already in progress by several groups, and clearly all of them aim at a setup which allows for a straightforward application to other processes.

From an experimentalist's point-of-view, the NLO calculations discussed thus far may be used to understand changes in normalization and/or shape that occur for a given process when going from LO to NLO [129]. As discussed previously, direct comparisons to the data require either a determination of parton-to-hadron corrections for the theory or hadron-to-parton corrections for the data [151]. And as also discussed previously, for multi-parton final states, it is also necessary to model the effects of jet algorithms, when two or more partons may be combined into one jet.

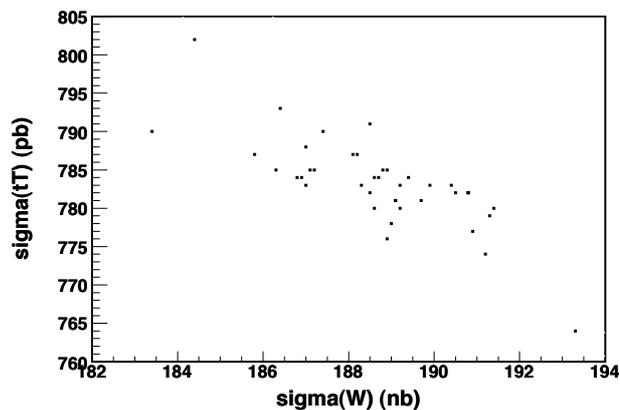
## 6.7. Some Standard Model Cross Sections for the LHC

### 6.7.1. $W/Z/Higgs$ production

Here we discuss some of the Standard Model benchmark cross sections at the LHC. A lack of space will keep the discussion short but a more

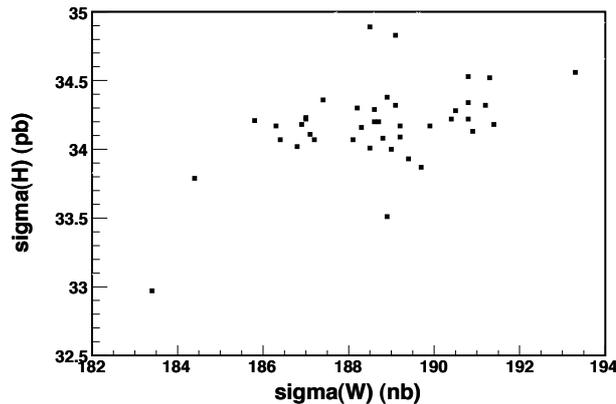


**Figure 79.** The cross section predictions for  $Z$  production versus the cross section predictions for  $W$  production at the LHC plotted using the 41 CTEQ6.1 pdf's.



**Figure 80.** The cross section predictions for  $t\bar{t}$  production versus the cross section predictions for  $W$  production at the LHC plotted using the 41 CTEQ6.1 pdf's.

complete treatment can be found at the benchmark website. The stability and PDF uncertainties for NLO  $W$  production at the LHC have been previously discussed. It is interesting to examine the pdf uncertainties of other processes at the LHC in relation to the pdf uncertainty for  $W$  production. The understanding gained may help to reduce the theoretical uncertainties for these processes. In the two figures below, the cross section predictions using the 41 CTEQ6.1 pdf's are presented. As shown in Figure 79, the cross section for  $Z$  production at the LHC is highly correlated with the cross section for  $W$  production. Both are sensitive to the low  $x$  quark pdf's, at a similar  $x$  value, which are driven by the gluon distribution at a slightly higher  $x$  value. Interestingly enough, as shown in Figure 80, the cross section for  $t\bar{t}$  production is anti-correlated with the  $W$  cross section. An increase in the  $W$  cross section is correlated with a decrease in the  $t\bar{t}$  cross section and vice versa. This is due to the fact that  $t\bar{t}$  production at the



**Figure 81.** The cross section predictions for Higgs production versus the cross section predictions for W production at the LHC plotted using the 41 CTEQ6.1 pdf's.

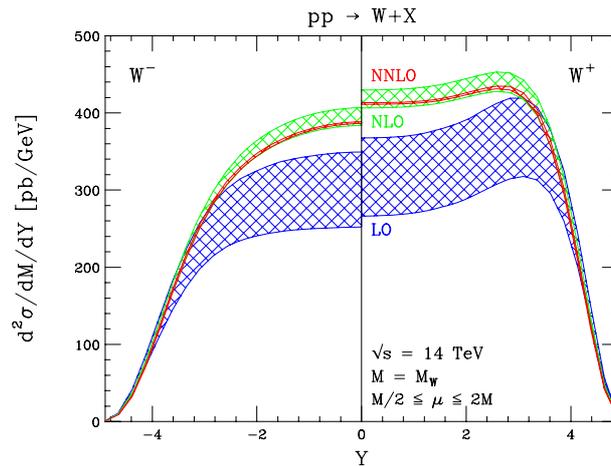
LHC proceeds primarily through  $gg$  fusion while  $W$  production is still  $q\bar{q}$ . An increase in the gluon distribution in the  $x$  range relevant for  $t\bar{t}$  production leads to a decrease in the quark distributions in the (lower)  $x$  range relevant for  $W$  production. In fact, the extremes for both cross sections are produced by eigenvector 5 (pdf's 9 and 10) which is most sensitive to the low  $x$  behavior of the gluon distribution.

The predictions from the 41 CTEQ6.1 PDF's for the production of a  $125 \text{ GeV}/c^2$  Higgs at the LHC are plotted versus similar predictions for the W cross section in Figure 81. Here, there is little correlation between the two cross sections. The PDF uncertainty for the Higgs cross section is fairly small as could be surmised from Figure 69. The low cross section point comes from error PDF 30, which is responsible for the high jet cross section predictions at high  $p_T$ ; the flow of momentum to high  $x$  removes it from the kinematic region responsible for producing a Higgs of this mass.

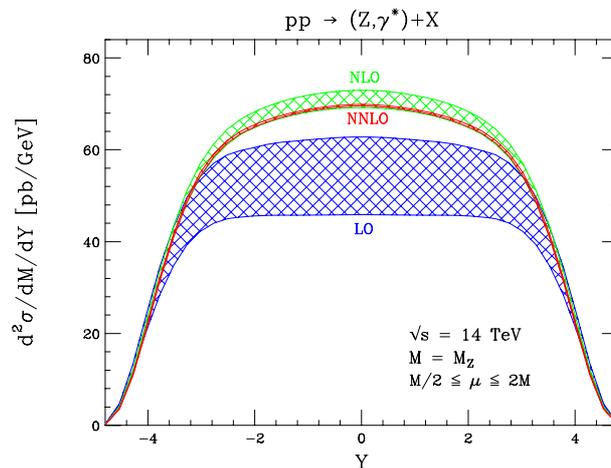
The rapidity distributions for  $W^+$  and  $W^-$  production at LO, NLO and NNLO are shown in Figure 82, while similar distributions for the  $Z$  are shown in Figure 83. The widths of the curves indicate the scale uncertainty for the cross section predictions. As expected, the scale dependence greatly decreases from LO to NLO to NNLO. There is a sizeable increase in the cross sections from LO to NLO, and a slight decrease (and basically no change in shape) in the cross sections from NLO to NNLO. The change from NLO to NNLO is within the NLO scale uncertainty band. As discussed in Section 4, the NNLO pdf's are still somewhat incomplete due to the lack of inclusion of inclusive jet production in the global fits. Although this is formally mixing orders, the result of using the NNLO matrix element with NLO pdf's for calculation of the  $Z$  rapidity distribution is shown in Figure 84. The predictions are similar as expected, but the prediction using the NLO pdf's is outside the (very precise) NNLO error band ‡.

The transverse momentum distributions for  $Z$  and Higgs production at the LHC

‡ We would like to thank Lance Dixon for providing these curves

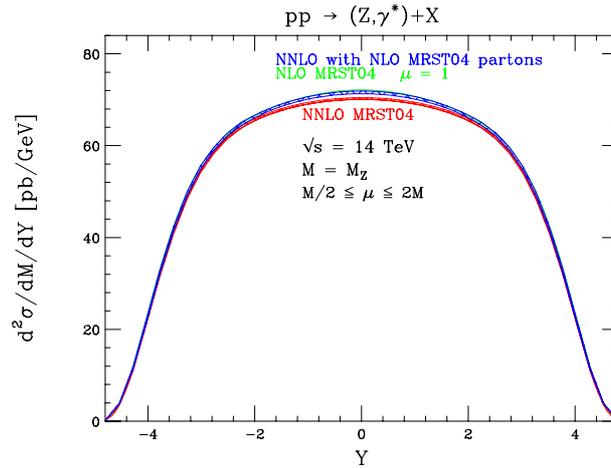


**Figure 82.** The rapidity distributions for  $W^+$  and  $W^-$  production at the LHC at LO, NLO and NNLO.

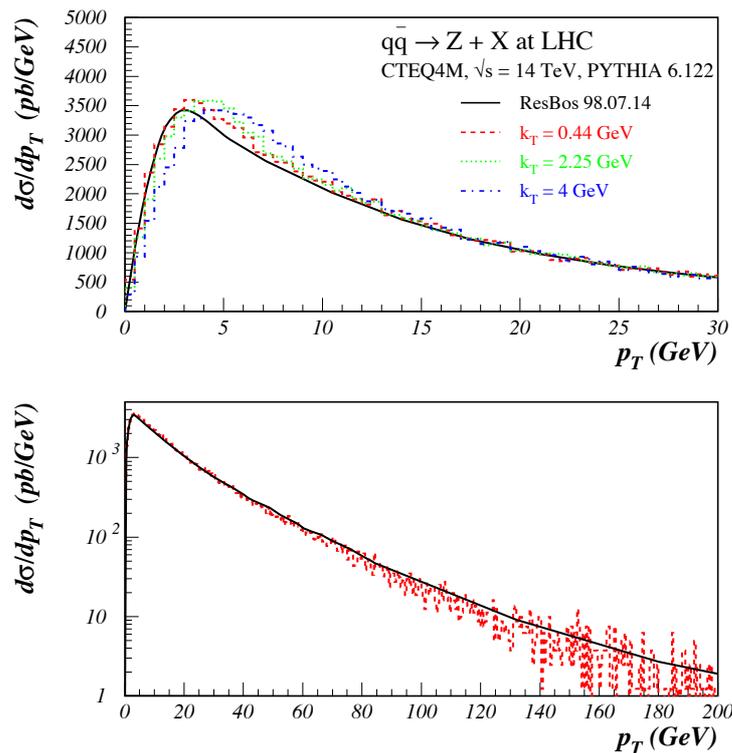


**Figure 83.** The rapidity distributions for  $Z$  production at the LHC at LO, NLO and NNLO.

are also important to understand. The  $Z$  will be one the Standard Model benchmark processes during the early running of the LHC and a knowledge of the production kinematics for the Higgs is crucial to its discovery. At low transverse momenta, the distributions are dominated by the effects of multiple soft gluon emission, while at higher  $p_T$ , hard gluon emission is the major contribution. In Figure 85, the  $Z$   $p_T$  distributions are shown using predictions from ResBos and from Pythia [152]. It is noteworthy that the  $p_T$  distribution looks similar to that observed at the Tevatron. There is not that much larger of a phase space for gluon emission of incident quarks at  $x=0.007$  ( $Z$  production at the LHC) than for incident quarks at  $x=0.05$  ( $Z$  production at the Tevatron). There is still the need for the input of a non-perturbative  $k_T$  to the parton shower predictions and it can be observed that the resultant transverse momentum distribution has a strong dependence on its value.

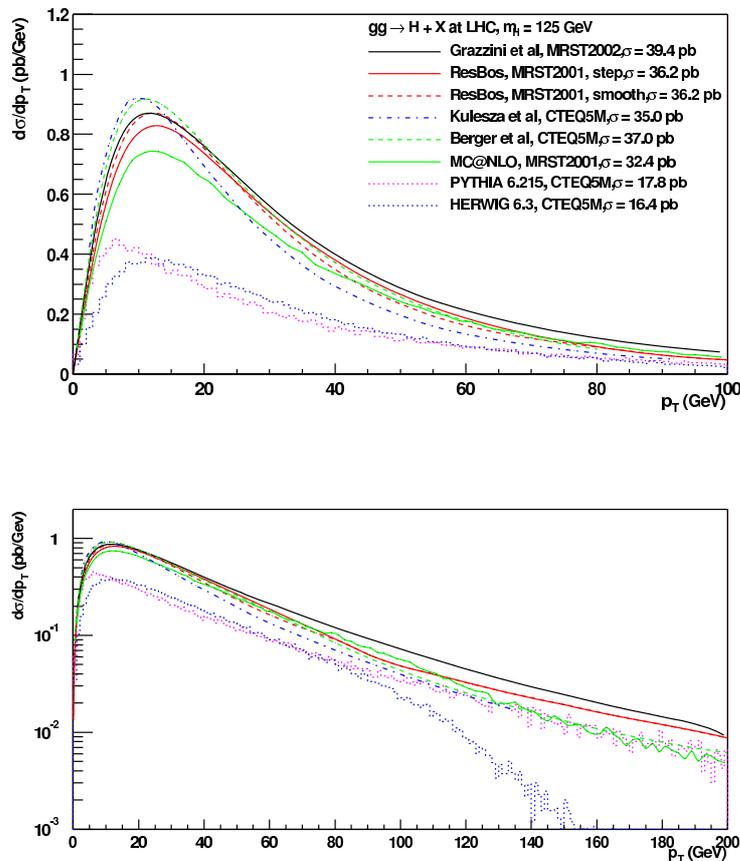


**Figure 84.** The rapidity distributions for  $Z$  production at the LHC at NNLO calculated with NNLO and with NLO pdf's.



**Figure 85.** Predictions for the transverse momentum distribution for  $Z$  production at the LHC.

In Figure 86 (un-normalized) and Figure 87 (normalized) are shown a large number of predictions for the transverse momentum distributions for the production of a  $125 \text{ GeV}/c^2$  mass Higgs at the LHC, through the  $gg$  fusion process. The impact of higher

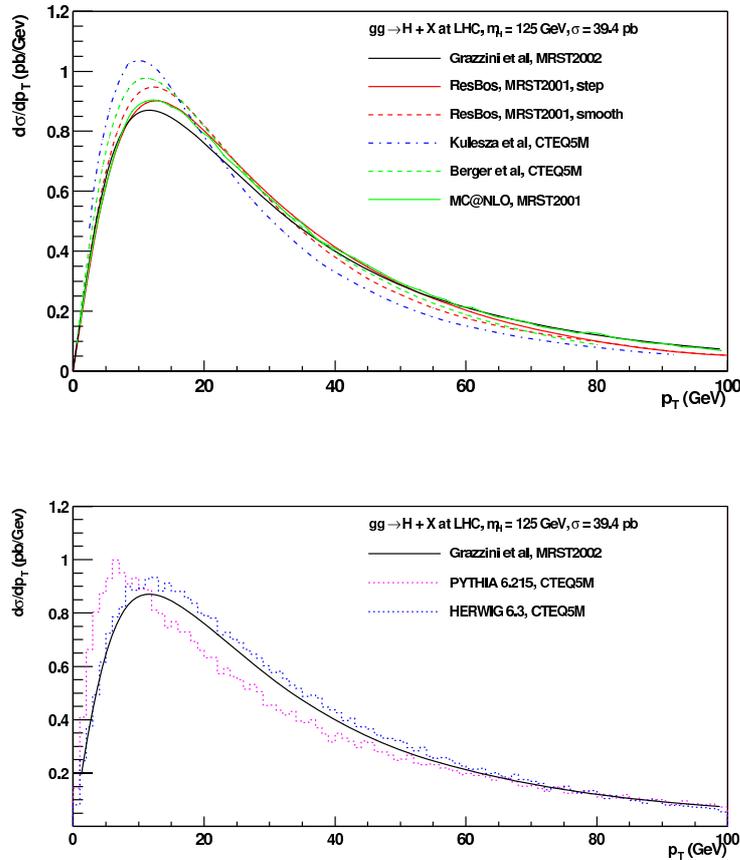


**Figure 86.** The predictions for the transverse momentum distribution for a 125 GeV mass Higgs at the LHC from a number of theoretical predictions. The predictions are shown with an absolute normalization.

order corrections (both NLO and NNLO) on the normalization of the Higgs cross section is evident in Figure 86. But in Figure 87, it can be observed that the shapes for the  $p_T$  distributions are basically the same. In particular, parton shower Monte Carlos provide a reasonable approximation to the more rigorous predictions shown. Two Pythia predictions are indicated: Pythia 6.215 is a virtuality-ordered parton shower while Pythia 6.3 is a  $p_T$ -ordered shower (and as different from Pythia 6.215 as Herwig is from Pythia). The peak of the Pythia 6.3 prediction is more in line with the other theoretical predictions.

Note that the  $p_T$  distribution peaks at approximately 1200 GeV/c while the  $p_T$  distribution for the  $Z$  peaks at 3 GeV/c. The peak is slightly higher because of the higher mass (125 GeV compared to 90 GeV) but the major shift is due to the larger colour factor for the gluon compared to the quark and the much larger phase space available for gluon radiation for  $x = 0.01$  gluons due to the  $z \rightarrow 0$  pole in the splitting function.

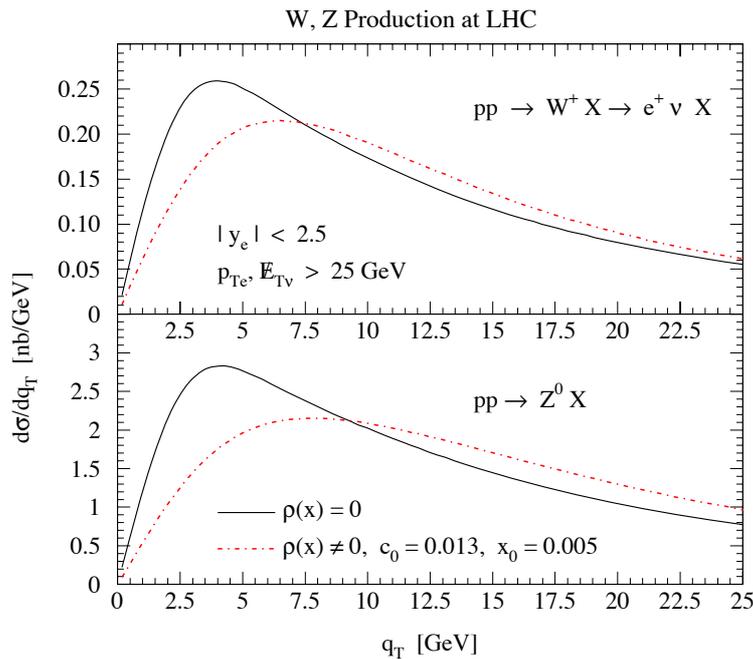
An analysis of semi-inclusive deep-inelastic scattering hadroproduction suggests a



**Figure 87.** The predictions for the transverse momentum distribution for a 125 GeV mass Higgs at the LHC from a number of theoretical predictions. The predictions have all been normalized to the same cross section for shape comparisons.

broadening of transverse momentum distributions for  $x$  values below  $10^{-3}$  to  $10^{-2}$  [153]. The  $q_T$  broadening at small  $x$  may be due to  $x$ -dependent higher order contributions (like BFKL) not included in current resummation formalisms. This would have a small impact at the Tevatron (except perhaps for  $W/Z$  production in the forward region) but may affect the predictions for  $W/Z/Higgs$   $p_T$  distributions for all rapidity regions at the LHC. The  $q_T$  broadening can be modeled in the Collins-Soper-Sterman formalism by a modification of the impact parameter-dependent parton densities. The  $q_T$  shifts for the  $W$  and  $Z$  transverse momentum distributions at the LHC are shown in Figure 88 [154]. The observed shifts would have important implications for the measurement of the  $W$  boson mass and a measurement of the  $W/Z$   $p_T$  distributions will be one of the important benchmarks to be established at the LHC.

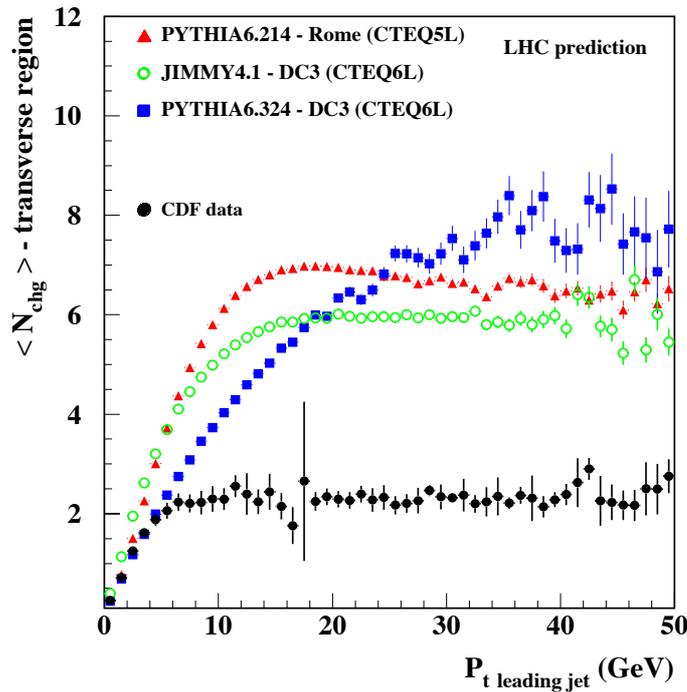
*6.7.2. Underlying Event at the LHC* As discussed previously, hard interactions at hadron-hadron colliders consist of a hard collision of two incoming partons along with softer interactions from the remaining partons in the colliding hadrons (“the underlying



**Figure 88.** The predictions for the transverse momentum distributions for  $W$  and  $Z$  production with and without the  $q_T$ -broadening effects.

event energy”). Some predictions for the underlying event charged multiplicities and charged momentum sum (defined in Section 5) at the LHC are shown in Figures 89 and 90. It is clear that (1) all predictions lead to a substantially larger charged particle multiplicity and charged particle momentum sum at the LHC than at the Tevatron and (2) there are large differences among the predictions from the various models. Investigations are continuing trying to reduce the energy extrapolation uncertainty of these models. This measurement will be one of the first to be performed at the LHC during the commissioning run in 2007 and will be used for subsequent Monte Carlo tunings for the LHC.

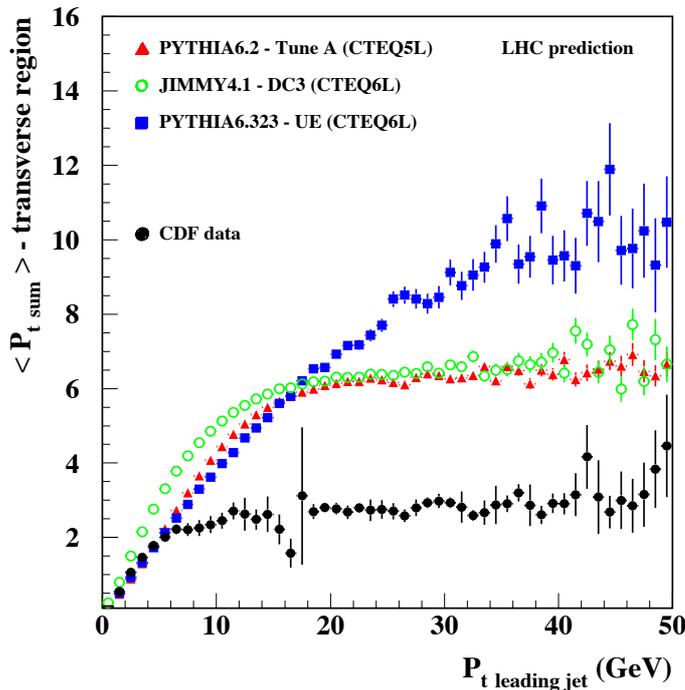
*6.7.3. Inclusive jet production* The increase of the center-of-mass energy to 14 TeV at the LHC will result in a dramatically larger accessible kinematic range. Inclusive jet cross sections can be measured out to the range of transverse momentum of 4 TeV/c in the central region and 1.5 TeV/c in the forward region. The predictions with the CTEQ6.1 central pdf’s and the 40 error pdf’s are shown in Figure 91 and 92 for three different rapidity regions [6]. The cross sections were generated with a renormalization and factorization scale equal to  $p_T^{jet}/2$ . The cross section predictions have a similar sensitivity to the error pdf’s as do the jet cross sections at the Tevatron for similar  $x_T$  values and the uncertainty on the predicted cross sections remain up to a factor of 2 at the highest  $p_T$  values. Measurements of the jet cross section over the full rapidity range at the LHC will serve to further constrain the high  $x$  gluon pdf and distinguish between possible new physics and uncertainties in pdf’s.



**Figure 89.** Pythia6.2 - Tune A, Jimmy4.1 - UE and Pythia6.323 - UE predictions for the average charged multiplicity in the transverse region in the underlying event for LHC  $pp$  collisions.

It is useful to plot the K-factors (the ratio of the NLO to LO cross sections) for the three different rapidity intervals shown above. As discussed previously, the value of the K-factor is a scale-dependent quantity; the K-factors shown in Figure 93 are with the nominal scale of  $p_T^{jet}/2$ . The K-factors have a somewhat complicated shape due to the interplay between the different subprocesses comprising inclusive jet production and the behaviors of the relevant pdf's in the different regions of parton momentum fraction  $x$ . In the central region, the K-factor is within 10% of unity for the observable range. There are no new parton-parton subprocesses that contribute at NLO but not at LO. Thus a LO prediction, using the NLO CTEQ6.1 pdf's, will reproduce fairly closely the NLO calculation. For rapidities between 1 and 2, the K-factor is within 20% of unity, dropping below one at higher transverse momentum. For forward rapidities, the K-factor drops almost immediately below one, due to the behavior of the high- $x$  pdf's that contribute to the cross section in this region. There is nothing wrong with the NLO prediction in this region; its relationship to the LO cross section has just changed due to the kinematics. LO predictions in this region will provide an overestimate to the NLO cross section.

Jet events at the LHC will be much more active than events in a similar  $p_T$  range at the Tevatron. The majority of the dijet production for the transverse momentum range less than 1 TeV/c will be with a  $gg$  initial state. As discussed previously, the

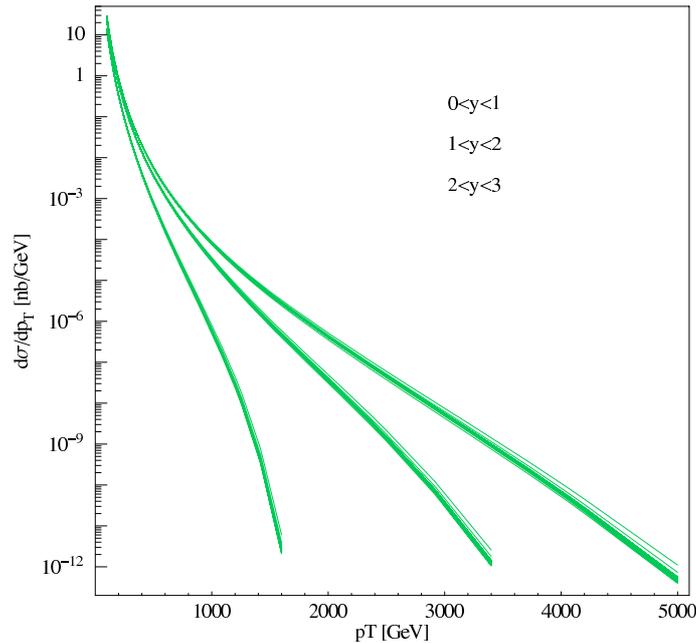


**Figure 90.** Pythia6.2 - Tune A, Jimmy4.1 - UE and Pythia6.323 - UE predictions for the average sum of the transverse momenta of charged particles in the transverse region in the underlying event for LHC  $pp$  collisions.

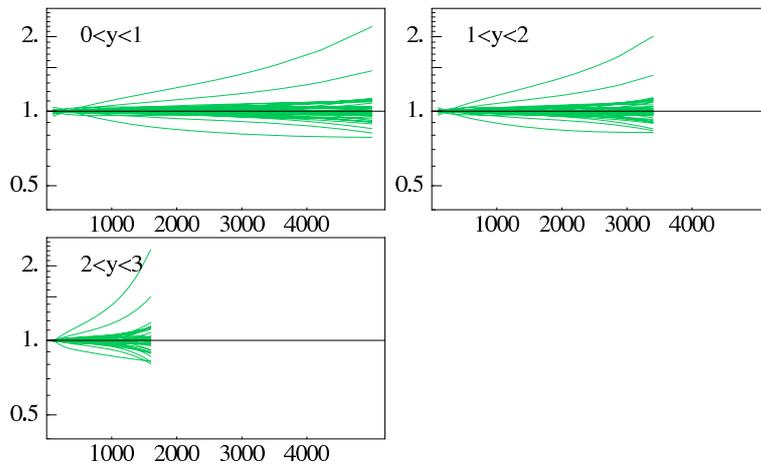
larger colour factor associated with the gluon and the greater phase space available at the LHC for gluon emission will result in an increased production of additional soft jets. In addition, there is an increased probability for the production of “mini-jets” from multiple-parton scattering among the spectator partons.

As discussed above, the LHC will produce jets at transverse momenta far larger than the  $W$  mass. Thus, virtual electroweak corrections of the form  $\alpha_W \log^2(p_T^2/M_W^2)$ , where  $\alpha_W = \alpha_{EM} \sin^2 \theta_W$ , may become important. For example, at  $p_T = 3$  TeV/c,  $\log^2(p_T^2/m_W^2) = 10$ . These double logs are a result of the lack of cancellation between real and virtual  $W$  emission in higher order calculations and may become competitive in size with NLO QCD corrections. The impact of the EW corrections reaches of the order of 40% at 4 TeV/c, as shown in Figure 94 below [155]. This is an indication that all Standard Model corrections, and not just those from higher order QCD, need to be understood in order to correctly interpret any new physics that may be present at the LHC.

**6.7.4.  $W + jets$**  Finally, we end with a brief discussion of the production of  $W + jets$  at the LHC. In Figure 95, the rate for production of  $W + \geq 1, 2, 3$  jets at the LHC is shown as a function of the transverse energy of the lead jet. The rate for 3 jet production is actually larger than the rate for 1 or 2 jet production for large lead jet transverse energy

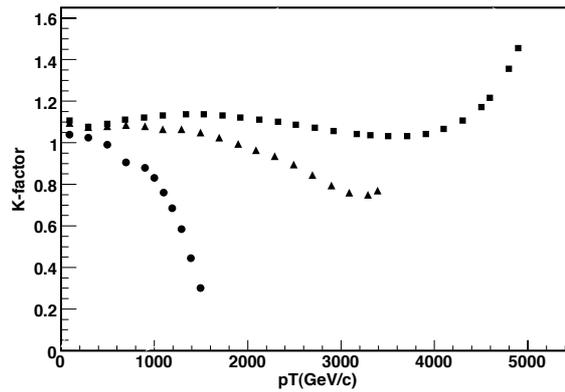


**Figure 91.** Inclusive jet cross section predictions for the LHC using the CTEQ6.1 central pdf and the 40 error pdf's.

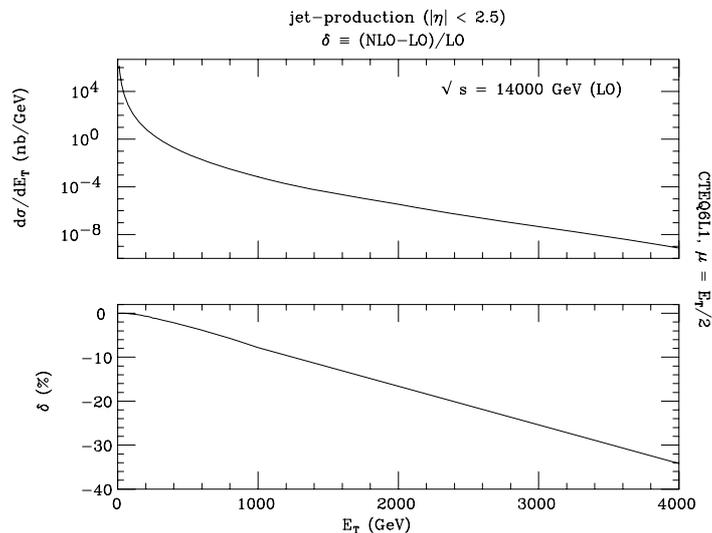


**Figure 92.** The ratios of the jet cross section predictions for the LHC using the CTEQ6.1 error pdf's to the prediction using the central pdf. The extremes are produced by eigenvector 15.

due to the Sudakov effects discussed throughout this paper. We also show the the rate for  $W + 3$  jet production (a LO calculation in MCFM) using both CTEQ6L1 (a leading order pdf) and CTEQ6.1 (a NLO pdf). The rate using the LO pdf is larger due to (1) the importance of the gluon distribution for production of  $W + 3$  jet final states and (2) the larger small  $x$  gluon present in the LO pdf compared to the NLO pdf.



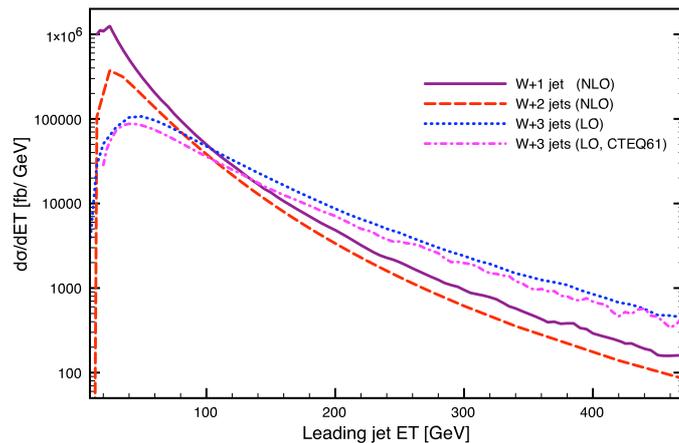
**Figure 93.** The ratios of the NLO to LO jet cross section predictions for the LHC using the CTEQ6.1 pdf's for the three different rapidity regions (0–1 (squares), 1–2 (triangles), 2–3 (circles)).



**Figure 94.** The effect of electroweak logarithms on jet cross sections at the LHC.

## 7. Summary

The LHC will present an unprecedented potential for the discovery of new physics. Thus, it is even more important to understand both the possible production mechanisms for the new physics and the Standard Model backgrounds for their signals. The experiences gained at the Tevatron, and the phenomenological tools developed in the last few years, will serve as a basis towards the establishment of the relevant Standard Model benchmarks at the LHC.



**Figure 95.** Predictions for the production of  $W + \geq 1, 2, 3$  jets at the LHC shown as a function of the transverse energy of the lead jet. A cut of 20 GeV/c has been placed on the other jets in the prediction.

## 8. Acknowledgements

We would like to thank A. Belyaev, L. Dixon, S. Ellis, S. Frixione, S. Gieseke, J. Heyninck, M. Mangano, S. Mrenna, S. Moretti, J. Pumplin, P. Skands, M. Toennesmann and HRH Elizabeth II for useful conversations.

## 9. References

- [1] S.D. Drell and T.M. Yan, *Ann. Phys.* **66** (1971) 578.
- [2] J.C. Collins and D.E. Soper, *Ann. Rev. Nucl. Part. Sci.* **37** (1987) 33.
- [3] R. K. Ellis, W. J. Stirling and B. R. Webber, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8**, 1 (1996).
- [4] L.N. Lipatov, *Sov. J. Nucl. Phys.* **20** (1975) 95;  
V.N. Gribov and L.N. Lipatov, *Sov. J. Nucl. Phys.* **15** (1972) 438;  
G. Altarelli and G. Parisi, *Nucl. Phys.* **B126** (1977) 298;  
Yu.L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 641.
- [5] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, *Phys. Lett. B* **604**, 61 (2004) [arXiv:hep-ph/0410230].
- [6] D. Stump, J. Huston, J. Pumplin, W. K. Tung, H. L. Lai, S. Kuhlmann and J. F. Owens, *JHEP* **0310**, 046 (2003) [arXiv:hep-ph/0303013].
- [7] D. Acosta *et al.* [CDF II Collaboration], *Phys. Rev. Lett.* **94**, 091803 (2005) [arXiv:hep-ex/0406078].
- [8] <http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/EW/>
- [9] S. D. Ellis, Z. Kunszt and D. E. Soper, *Phys. Rev. Lett.* **69**, 3615 (1992) [arXiv:hep-ph/9208249].
- [10] G. Corcella *et al.*, *JHEP* **0101**, 010 (2001) [arXiv:hep-ph/0011363].
- [11] T. Sjostrand, S. Mrenna and P. Skands, *JHEP* **0605**, 026 (2006) [arXiv:hep-ph/0603175].
- [12] M. L. Mangano, M. Moretti and R. Pittau, *Nucl. Phys. B* **632**, 343 (2002) [arXiv:hep-ph/0108069].
- [13] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *JHEP* **0307**, 001 (2003) [arXiv:hep-ph/0206293].
- [14] A. Pukhov *et al.*, arXiv:hep-ph/9908288.

- [15] E. Boos *et al.* [CompHEP Collaboration], Nucl. Instrum. Meth. A **534**, 250 (2004) [arXiv:hep-ph/0403113].
- [16] T. Stelzer and W. F. Long, Comput. Phys. Commun. **81**, 357 (1994) [arXiv:hep-ph/9401258].
- [17] F. Maltoni and T. Stelzer, JHEP **0302**, 027 (2003) [arXiv:hep-ph/0208156].
- [18] F. Caravaglios, M. L. Mangano, M. Moretti and R. Pittau, Nucl. Phys. B **539**, 215 (1999) [arXiv:hep-ph/9807570].
- [19] H. Murayama, I. Watanabe and K. Hagiwara, KEK-91-11
- [20] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, Phys. Rev. D **67**, 014026 (2003) [arXiv:hep-ph/0209271].
- [21] E. Witten, Commun. Math. Phys. **252**, 189 (2004) [arXiv:hep-th/0312171].
- [22] F. Cachazo, P. Svrcek and E. Witten, JHEP **0409**, 006 (2004) [arXiv:hep-th/0403047].
- [23] G. Georgiou and V. V. Khoze, JHEP **0405**, 070 (2004) [arXiv:hep-th/0404072].
- [24] J. B. Wu and C. J. Zhu, JHEP **0409**, 063 (2004) [arXiv:hep-th/0406146].
- [25] G. Georgiou, E. W. N. Glover and V. V. Khoze, JHEP **0407**, 048 (2004) [arXiv:hep-th/0407027].
- [26] L. J. Dixon, E. W. N. Glover and V. V. Khoze, JHEP **0412**, 015 (2004) [arXiv:hep-th/0411092].
- [27] S. D. Badger, E. W. N. Glover and V. V. Khoze, JHEP **0503**, 023 (2005) [arXiv:hep-th/0412275].
- [28] Z. Bern, D. Forde, D. A. Kosower and P. Mastrolia, Phys. Rev. D **72**, 025006 (2005) [arXiv:hep-ph/0412167].
- [29] R. Britto, F. Cachazo and B. Feng, Nucl. Phys. B **715**, 499 (2005) [arXiv:hep-th/0412308].
- [30] R. Britto, F. Cachazo, B. Feng and E. Witten, Phys. Rev. Lett. **94**, 181602 (2005) [arXiv:hep-th/0501052].
- [31] F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).
- [32] T. Kinoshita, J. Math. Phys. **3**, 650 (1962).
- [33] T. D. Lee and M. Nauenberg, Phys. Rev. **133**, B1549 (1964).
- [34] K. Kajantie, J. Lindfors and R. Raitio, Nucl. Phys. B **144**, 422 (1978).
- [35] Z. Kunszt and D. E. Soper, Phys. Rev. D **46**, 192 (1992).
- [36] R. K. Ellis, D. A. Ross and A. E. Terrano, Nucl. Phys. B **178**, 421 (1981).
- [37] S. Frixione, Z. Kunszt and A. Signer, Nucl. Phys. B **467**, 399 (1996) [arXiv:hep-ph/9512328].
- [38] S. Catani and M. H. Seymour, Nucl. Phys. B **485**, 291 (1997) [Erratum-ibid. B **510**, 503 (1997)] [arXiv:hep-ph/9605323].
- [39] Z. Nagy and Z. Trocsanyi, Nucl. Phys. B **486**, 189 (1997) [arXiv:hep-ph/9610498].
- [40] K. Fabricius, I. Schmitt, G. Kramer and G. Schierholz, Z. Phys. C **11**, 315 (1981).
- [41] W. T. Giele and E. W. N. Glover, Phys. Rev. D **46**, 1980 (1992).
- [42] A. Gehrmann-De Ridder, T. Gehrmann and E. W. N. Glover, JHEP **0509**, 056 (2005) [arXiv:hep-ph/0505111].
- [43] E. W. N. Glover and D. J. Miller, Phys. Lett. B **396**, 257 (1997) [arXiv:hep-ph/9609474].
- [44] Z. Bern, L. J. Dixon, D. A. Kosower and S. Weinzierl, Nucl. Phys. B **489**, 3 (1997) [arXiv:hep-ph/9610370].
- [45] J. M. Campbell, E. W. N. Glover and D. J. Miller, Phys. Lett. B **409**, 503 (1997) [arXiv:hep-ph/9706297].
- [46] E. W. N. Glover, Nucl. Phys. Proc. Suppl. **116**, 3 (2003) [arXiv:hep-ph/0211412].
- [47] Z. Bern, L. J. Dixon and D. A. Kosower, Nucl. Phys. B **513**, 3 (1998) [arXiv:hep-ph/9708239].
- [48] S. D. Badger and E. W. N. Glover, JHEP **0407**, 040 (2004) [arXiv:hep-ph/0405236].
- [49] H1 Collaboration: C. Adloff *et al.*, Eur. Phys. J. **C 13** (2000) 609 [hep-ex/9908059]; Eur. Phys. J. **C 19** (2001) 269 [hep-ex/0012052]; Eur. Phys. J. **C 21** (2001) 33 [hep-ex/0012053].
- [50] ZEUS Collaboration: S. Chekanov *et al.*, Eur. Phys. J. **C 21** (2001) 443 [hep-ex/0105090]; A.M. Cooper-Sarkar, Proceedings of International Europhysics Conference on HEP 2001, Budapest [hep-ph/0110386].
- [51] E866 Collaboration: R. S. Towell *et al.*, Phys. Rev. **D 64** (2001) 052002 [hep-ex/0103030].
- [52] CCFR Collaboration: U. K. Yang *et al.*, Phys. Rev. Lett. **86** (2001) 2742 [hep-ex/0009041].
- [53] BCDMS Collaboration: A. C. Benvenuti *et al.*, Phys. Lett. **B 223** (1989) 485.

- [54] BCDMS Collaboration: A. C. Benvenuti *et al.*, Phys. Lett. **B 236** (1989) 592.
- [55] New Muon Collaboration: M. Arneodo *et al.*, Nucl. Phys. **B 483** (1997) 3 [hep-ph/9610231]; and M. Arneodo *et al.*, Nucl. Phys. **B 487** (1997) 3 [hep-ex/9611022].
- [56] CCFR Collaboration: W. G. Seligman *et al.*, Phys. Rev. Lett. **79** (1997) 1213 [hep-ex/970107].
- [57] E605 Collaboration: G. Moreno *et al.*, Phys. Rev. **D 43** (1991) 2815.
- [58] CDF Collaboration: T. Affolder *et al.*, Phys. Rev. **D 64** (2001) 032001 [hep-ph/0102074].
- [59] DØ Collaboration: B. Abbott *et al.*, Phys. Rev. Lett. **86** (2001) 1707 [hep-ex/0011036]; and Phys. Rev. **D 64** (2001) 032003 [hep-ex/0012046].
- [60] S. Moch, J. A. M. Vermaseren and A. Vogt, Nucl. Phys. **B 688**, 101 (2004) [arXiv:hep-ph/0403192].
- [61] A. Vogt, S. Moch and J. A. M. Vermaseren, Nucl. Phys. **B 691**, 129 (2004) [arXiv:hep-ph/0404111].
- [62] R. Hamberg, W. L. van Neerven and T. Matsuura, Nucl. Phys. **B 359**, 343 (1991) [Erratum-ibid. **B 644**, 403 (2002)].
- [63] R. V. Harlander and W. B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002) [arXiv:hep-ph/0201206].
- [64] C. Anastasiou and K. Melnikov, Nucl. Phys. **B 646**, 220 (2002) [arXiv:hep-ph/0207004].
- [65] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, Phys. Rev. Lett. **91**, 182002 (2003) [arXiv:hep-ph/0306192].
- [66] C. Anastasiou, K. Melnikov and F. Petriello, Phys. Rev. Lett. **93**, 262002 (2004) [arXiv:hep-ph/0409088].
- [67] G. Sterman, Nucl. Phys. **B 281**, 310 (1987).
- [68] S. Catani and L. Trentadue, Nucl. Phys. **B 327**, 323 (1989).
- [69] S. Catani and L. Trentadue, Nucl. Phys. **B 353**, 183 (1991).
- [70] G. Parisi and R. Petronzio, Nucl. Phys. **B 154**, 427 (1979).
- [71] G. Altarelli, R. K. Ellis, M. Greco and G. Martinelli, Nucl. Phys. **B 246**, 12 (1984).
- [72] J. C. Collins and D. E. Soper, Nucl. Phys. **B 193**, 381 (1981) [Erratum-ibid. **B 213**, 545 (1983)].
- [73] J. C. Collins and D. E. Soper, Nucl. Phys. **B 197**, 446 (1982).
- [74] J. C. Collins, D. E. Soper and G. Sterman, Nucl. Phys. **B 250**, 199 (1985).
- [75] G. Sterman and W. Vogelsang, Phys. Rev. **D 71**, 014013 (2005) [arXiv:hep-ph/0409234].
- [76] A. Kulesza, G. Sterman and W. Vogelsang, Phys. Rev. **D 69**, 014012 (2004) [arXiv:hep-ph/0309264].
- [77] A. Kulesza, G. Sterman and W. Vogelsang, arXiv:hep-ph/0207148.
- [78] C. Balazs and C. P. Yuan, Phys. Rev. **D 56**, 5558 (1997) [arXiv:hep-ph/9704258].
- [79] A. Banfi, G. Corcella, M. Dasgupta, Y. Delenda, G. P. Salam and G. Zanderighi, arXiv:hep-ph/0508096.
- [80] T. Gleisberg, S. Hoche, F. Krauss, A. Schaliche, S. Schumann and J. C. Winter, JHEP **0402**, 056 (2004) [arXiv:hep-ph/0311263].
- [81] T. Sjostrand and P. Z. Skands, Eur. Phys. J. **C 39**, 129 (2005) [arXiv:hep-ph/0408302].
- [82] S. Gieseke, JHEP **0501**, 058 (2005) [arXiv:hep-ph/0412342].
- [83] W. T. Giele, E. W. N. Glover and D. A. Kosower, Phys. Rev. Lett. **73**, 2019 (1994) [arXiv:hep-ph/9403347].
- [84] S. D. Ellis, J. Huston and M. Tonnesmann, in *Proc. of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001)* ed. N. Graf, eConf **C010630**, P513 (2001) [arXiv:hep-ph/0111434].
- [85] E. Boos *et al.*, arXiv:hep-ph/0109068.
- [86] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, JHEP **0111**, 063 (2001) [arXiv:hep-ph/0109231].
- [87] T. Gleisberg, S. Hoeche, F. Krauss, A. Schaelicke, S. Schumann, G. Soff and J. Winter, Czech. J. Phys. **55**, B529 (2005) [arXiv:hep-ph/0409122].
- [88] S. Mrenna and P. Richardson, JHEP **0405**, 040 (2004) [arXiv:hep-ph/0312274].
- [89] Y. j. Chen, J. Collins and X. m. Zu, JHEP **0204**, 041 (2002) [arXiv:hep-ph/0110257].
- [90] J. C. Collins and X. Zu, JHEP **0503**, 059 (2005) [arXiv:hep-ph/0411332].
- [91] S. Frixione and B. R. Webber, JHEP **0206**, 029 (2002) [arXiv:hep-ph/0204244].
- [92] S. Frixione, P. Nason and B. R. Webber, JHEP **0308**, 007 (2003) [arXiv:hep-ph/0305252].

- [93] S. Frixione and B. R. Webber, arXiv:hep-ph/0506182.
- [94] Y. Kurihara, J. Fujimoto, T. Ishikawa, K. Kato, S. Kawabata, T. Munehisa and H. Tanaka, Nucl. Phys. B **654**, 301 (2003) [arXiv:hep-ph/0212216].
- [95] M. Kramer and D. E. Soper, Phys. Rev. D **69**, 054019 (2004) [arXiv:hep-ph/0306222].
- [96] D. E. Soper, Phys. Rev. D **69**, 054020 (2004) [arXiv:hep-ph/0306268].
- [97] M. Kramer, S. Mrenna and D. E. Soper, Phys. Rev. D **73**, 014022 (2006) [arXiv:hep-ph/0509127].
- [98] Z. Nagy and D. E. Soper, JHEP **0510**, 024 (2005) [arXiv:hep-ph/0503053].
- [99] Z. Nagy and D. E. Soper, arXiv:hep-ph/0601021.  
[100]
- [100] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, arXiv:hep-ph/0512250.
- [101] S. Alekhin, JETP Lett. **82**, 628 (2005) [Pisma Zh. Eksp. Teor. Fiz. **82**, 710 (2005)] [arXiv:hep-ph/0508248].
- [102] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **21**, 33 (2001) [arXiv:hep-ex/0012053].
- [103] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **30**, 1 (2003) [arXiv:hep-ex/0304003].
- [104] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Rev. D **67**, 012007 (2003) [arXiv:hep-ex/0208023].
- [105] S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **42**, 1 (2005) [arXiv:hep-ph/0503274].
- [106] S. Moch, J. A. M. Vermaseren and A. Vogt, Nucl. Phys. Proc. Suppl. **135**, 137 (2004) [arXiv:hep-ph/0408075].
- [107] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Phys. Lett. B **531**, 216 (2002) [arXiv:hep-ph/0201127].
- [108] M. Gluck, E. Reya and A. Vogt, Eur. Phys. J. C **5**, 461 (1998) [arXiv:hep-ph/9806404].
- [109] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **81**, 5754 (1998) [arXiv:hep-ex/9809001].
- [110] D. Stump *et al.*, Phys. Rev. D **65**, 014012 (2002) [arXiv:hep-ph/0101051].
- [111] J. Pumplin *et al.*, Phys. Rev. D **65**, 014013 (2002) [arXiv:hep-ph/0101032].
- [112] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. C **28**, 455 (2003) [arXiv:hep-ph/0211080].
- [113] S. Eidelman *et al.* [Particle Data Group], Phys. Lett. B **592**, 1 (2004).
- [114] J. Pumplin, A. Belyaev, J. Huston, D. Stump and W. K. Tung, JHEP **0602**, 032 (2006) [arXiv:hep-ph/0512167].
- [115] Rick Field's writeup in TeV4LHC
- [116] J. C. Collins and X. m. Zu, JHEP **0206**, 018 (2002) [arXiv:hep-ph/0204127].
- [117] H. Plothow-Besch, Int. J. Mod. Phys. A **10**, 2901 (1995).
- [118] W. T. Giele, S. A. Keller and D. A. Kosower, arXiv:hep-ph/0104052.
- [119] W. Giele *et al.*, arXiv:hep-ph/0204316.
- [120] M. R. Whalley, D. Bourilkov and R. C. Group, arXiv:hep-ph/0508110.
- [121] J. M. Campbell and R. K. Ellis, Phys. Rev. D **62**, 114012 (2000) [arXiv:hep-ph/0006304].
- [122] G. J. Alner *et al.* [UA5 Collaboration], Phys. Rept. **154**, 247 (1987).
- [123] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, Phys. Rev. D **69**, 094008 (2004) [arXiv:hep-ph/0312266].
- [124] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **70**, 072002 (2004) [arXiv:hep-ex/0404004].  
[125]
- [125] A. A. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **65**, 092002 (2002).
- [126] S. D. Ellis, Z. Kunszt and D. E. Soper, Phys. Rev. Lett. **64**, 2121 (1990).
- [127] A. Abulencia *et al.* [CDF Run II Collaboration], arXiv:hep-ex/0512020.
- [128] A. Abulencia *et al.* [CDF II Collaboration], Phys. Rev. Lett. **96**, 122001 (2006) [arXiv:hep-ex/0512062].
- [129] J. Campbell and J. Huston, Phys. Rev. D **70**, 094021 (2004).
- [130] C. Buttar *et al.*, arXiv:hep-ph/0604120.
- [131] A. Pukhov, arXiv:hep-ph/0412191.
- [132] S. Alekhin, [arXiv:hep-ph/0311184]; Phys. Rev. D **68**, 014002 (2003) [arXiv:hep-ph/0211096].
- [133] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. C **35**, 325 (2004)

- [arXiv:hep-ph/0308087].
- [134] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, *Phys. Lett. B* **531**, 216 (2002) [arXiv:hep-ph/0201127].
- [135] J. Huston, J. Pumplin, D. Stump and W. K. Tung, *JHEP* **0506**, 080 (2005) [arXiv:hep-ph/0502080].
- [136] Z. Bern, L. J. Dixon and D. A. Kosower, *Ann. Rev. Nucl. Part. Sci.* **46**, 109 (1996) [arXiv:hep-ph/9602280].
- [137] F. Cachazo, P. Svrcek and E. Witten, *JHEP* **0410**, 074 (2004) [arXiv:hep-th/0406177].
- [138] A. Brandhuber, B. J. Spence and G. Travaglini, *Nucl. Phys. B* **706**, 150 (2005) [arXiv:hep-th/0407214].
- [139] A. Brandhuber, S. McNamara, B. J. Spence and G. Travaglini, *JHEP* **0510**, 011 (2005) [arXiv:hep-th/0506068].
- [140] Z. Bern, L. J. Dixon and D. A. Kosower, *Phys. Rev. D* **71**, 105013 (2005) [arXiv:hep-th/0501240].
- [141] D. Forde and D. A. Kosower, arXiv:hep-ph/0509358.
- [142] A. Ferroglia, M. Passera, G. Passarino and S. Uccirati, *Nucl. Phys. B* **650**, 162 (2003) [arXiv:hep-ph/0209219].
- [143] C. Anastasiou and A. Daleo, arXiv:hep-ph/0511176.
- [144] W. T. Giele and E. W. N. Glover, *JHEP* **0404**, 029 (2004) [arXiv:hep-ph/0402152].
- [145] W. Giele, E. W. N. Glover and G. Zanderighi, *Nucl. Phys. Proc. Suppl.* **135**, 275 (2004) [arXiv:hep-ph/0407016].
- [146] F. del Aguila and R. Pittau, *JHEP* **0407**, 017 (2004) [arXiv:hep-ph/0404120].
- [147] A. van Hameren, J. Vollinga and S. Weinzierl, *Eur. Phys. J. C* **41**, 361 (2005) [arXiv:hep-ph/0502165].
- [148] T. Binoth, J. P. Guillet, G. Heinrich, E. Pilon and C. Schubert, *JHEP* **0510**, 015 (2005) [arXiv:hep-ph/0504267].
- [149] R. K. Ellis, W. T. Giele and G. Zanderighi, arXiv:hep-ph/0508308.
- [150] R. K. Ellis, W. T. Giele and G. Zanderighi, *Phys. Rev. D* **72**, 054018 (2005) [arXiv:hep-ph/0506196].
- [151] G. U. Flanagan, FERMILAB-THESIS-2005-58
- [152] C. Balazs, J. Huston and I. Puljak, *Phys. Rev. D* **63**, 014021 (2001) [arXiv:hep-ph/0002032].
- [153] S. Berge, P. Nadolsky, F. Olness and C. P. Yuan, *Phys. Rev. D* **72**, 033015 (2005) [arXiv:hep-ph/0410375].
- [154] S. Berge, P. M. Nadolsky, F. I. Olness and C. P. Yuan, *AIP Conf. Proc.* **792**, 722 (2005) [arXiv:hep-ph/0508215].
- [155] S. Moretti, M. R. Nolten and D. A. Ross, arXiv:hep-ph/0503152.