Vector Bosons in New Phenomena Searches

Habilitationsschrift der Fakultät für Physik
der Ludwig-Maximilians-Universität München

Thomas Nunnemann

München, 24. August 2007

Vector Bosons in New Phenomena Searches

Thomas Nunnemann
Ludwigs-Maximilians-Universität München

München, 24. August 2007

Abstract

The production of the weak bosons, W and Z, is one of the dominant background contributions in searches for new phenomena and the Higgs boson at hadron colliders. In particular, their production with associated jets generates final state topologies which resemble those of many signal processes, including the production of leptoquarks, of scalar quarks and gluinos in supersymmetric models, and of the Higgs boson via WH and ZH. Therefore, an accurate description of vector boson production, based both on precision data and phenomenological modelling, is required.

First, the phenomenology of vector boson production at hadron colliders is reviewed including a discussion on the simulation of these processes. After summarizing the experimental data on total and differential cross sections for inclusive W and Z boson production obtained at the Tevatron, measurements of their associated production with jets are presented in detail and compared to QCD predictions. Finally, the relevance of W/Z + jets production as main background in searches for new phenomena and the Higgs boson is discussed. Using several exemplary searches both in final states with leptons and jets and with missing transverse momentum and jets, emphasis is put on reviewing methods to estimate and reject this background contribution.

Contents

Chapter 1

Introduction

At the current and future colliders, namely the Tevatron and the LHC, one of the main focus will be the search for the Higgs boson and for new phenomena beyond the standard model of particle physics. The production and decay of the Higgs boson and of new particles will often lead to signatures, which are similar to the ones of the production of the electroweak bosons Z and W. Therefore, in many cases the single or pair-production of the vector bosons, in particular their production in association with jets will dominate the background processes in searches for the Higgs boson and for new phenomena. Thus, a precise knowledge of vector boson production, both from precision data and phenomenological modelling, is mandatory.

Despite the success of the standard model (SM) of electroweak [1–3] and strong [4–6] interactions which is in excellent agreement with the experimental results, one of its key features, the electroweak symmetry breaking via the Higgs mechanism [7–10], which generates the masses of the Z and W bosons, has not yet been verified experimentally. As a consequence of this mechanism which introduces a doublet of complex scalar fields, a single neutral scalar particle, namely the Higgs boson remains after the symmetry breaking. Assuming the validity of the standard model, global fits to the electroweak data prefer a relatively low mass for the Higgs boson, $m_H = 85^{+39}_{-28}\,\mathrm{GeV}/c^2$ [11], while direct searches at the LEP collider set a lower bound on the mass of $114.4\,\mathrm{GeV}/c^2$ [12]. At low masses, $m_H \lesssim 140\,\mathrm{GeV}/c^2$, the standard model Higgs boson will dominantly decay via $H \to b\bar{b}$. For the main production channel, which is the gluon-gluon fusion process $gg \to H$ this will lead to signatures which are irreducible from QCD production of $b\bar{b}$ pairs. Therefore, at the Tevatron the highest sensitivity for low mass Higgs bosons is obtained from the associated production of the Higgs boson with the weak bosons, i.e. WH and ZH. The dominating background in these channels is the associated production of a Z or W boson with a $b\bar{b}$ pair.

While the standard model is in excellent agreement with all experimental data, as noted before, it contains several unsolved problems, e.g. the origin of the fermion mass spectrum, the structure of the quark and lepton families etc. Also, the SM is arguably *unnatural* as it contains the infamous hierarchy problem, the fact that the electroweak scale M_Z and the Planck scale M_P, at which quantum effects of gravity become strong, are about 17 orders of magnitude apart. Another aspect of the hierarchy problem is that it would generate radiative corrections to the bare Higgs mass which would require an unnatural large fine-tuning. Supersymmetry [13], which is the symmetry to

interchange fermions and bosons and which introduces a super-partner for each fermion and boson of the SM, provides an elegant solution to the fine-tuning problem, as the radiative corrections from SM boson and fermion loops are automatically cancelled through the corresponding loops of their super-partner fermion or boson.

If kinematically allowed, the super-partners of quarks q and the gluon g, scalar quarks (squarks) $\tilde{q}$ and the gluino $\tilde{g}$, would be produced with sizable cross sections at hadron colliders. The (cascade-) decays of squarks and gluinos would lead to event topologies containing several high energetic jets and missing transverse energy due to the escaping lightest supersymmetric particle[1]. For this topology, the main SM background is from Z boson production in association with jets, where the Z is decaying invisibly into neutrinos, $Z \to \nu\bar{\nu}$.

Although it will not be further discussed in the following, it should be noted that the production of W bosons in association with jets and in particular with a $b\bar{b}$ pair is also the dominating or a major background for single and pair-production of top quarks.

This work is organized as follows: First the phenomenology and the experimental data on inclusive vector boson production is discussed and results on total and differential cross section measurements are reviewed. The following chapter focuses on the associated production of vector bosons with jets and discusses both the case of flavour-inclusive jets and the case of heavy-flavour jets. An exemplary review of those searches for the Higgs boson and new phenomena at the Tevatron follows, for which the production of W or Z bosons in association with jets constitutes the major background process. Here, the primary interest is not to discuss the details of the phenomenology of the Higgs boson or of particular models for new physics beyond the standard model, but to focus on the estimation and rejection of the Z/W + jet background for various exemplary analyses.

[1]Here, it is assumed that R-parity is conserved. In supersymmetric models without R-parity conservation the lightest supersymmetric particle is not stable.

Chapter 2

Phenomenology of vector boson production

At hadron colliders the weak bosons, W and Z, are mainly produced through quark-antiquark annihilation. The phenomenological description of their production closely follows the Drell-Yan mechanism [14] for the production of a lepton pair $\ell^+\ell^-$ with large invariant mass $M_{\ell^+\ell^-} = \sqrt{(p_{\ell^+} + p_{\ell^-})^2}$, which is sketched in Fig. 2.1.

2.1 Inclusive production and decay

The cross section $\sigma_{AB\to\ell^+\ell^-}$ for producing a lepton pair in the collision of hadrons A and B is obtained by folding the underlying cross section for the quark-antiquark annihilation $\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-}$ with the parton distribution functions (PDFs) $f_{A,q}(x, M^2_{\ell^+\ell^-})$ and $f_{B,\bar{q}}(x, M^2_{\ell^+\ell^-})$ and summing over all possible quark-antiquark combinations in the hard interaction:

$$\sigma_{AB\to\ell^+\ell^-} = \sum_q \int \mathrm{d}x_A \mathrm{d}x_B f_{A,q}(x_A, M^2_{\ell^+\ell^-}) f_{B,\bar{q}}(x_B, M^2_{\ell^+\ell^-}) \hat{\sigma}_{q\bar{q}\to\ell^+\ell^-} \ . \tag{2.1}$$

The integration is performed over the momentum fractions of the partons x_A and x_B and the sum over q includes the antiquark (quark) distribution in hadron A (B) as well. At the Tevatron $p\bar{p}$ collider at a centre-of-mass energy $\sqrt{s} = 1.96\,\mathrm{TeV}$, pure valence-quark scattering amounts only to about 25% of the full cross section even for relatively large $M_{\ell^+\ell^-} \sim M_Z$ (Z boson mass), with the rest involving quarks from the proton sea.

Originally, Eq. (2.1) was derived within the parton model where the parton distribution functions only depend on the momentum fraction x and Bjorken scaling applies [15, 16]. The factorization of the hadron-hadron cross section into a fundamental parton-parton cross section and the PDFs of the scattering hadrons is still valid after inclusion of perturbative QCD corrections, as the singularities which occur in the corrections to the lowest-order processes can be absorbed into the scale-dependent PDFs. It has been shown that the factorization theorem generally holds for inclusive hard-scattering processes in hadron-hadron collisions [17, 18].

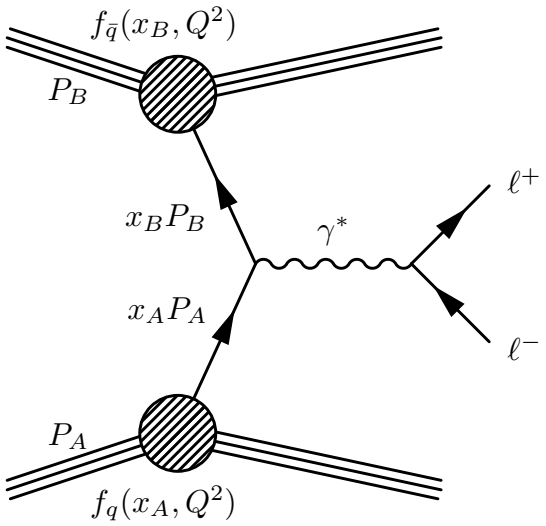

Figure 2.1: The production of a lepton pair in the Drell-Yan model.

The leading-order cross section for the quark-antiquark annihilation into a lepton pair $\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-}$ in Eq. (2.1) via an intermediate photon can be derived from the fundamental $e^+e^- \to \mu^+\mu^-$ cross section (modified with an additional colour factor 1/3) and is given by [19]

$$\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-} = \frac{4\pi\alpha^2}{9\hat{s}} Q_q^2 \,, \tag{2.2}$$

where α is the electromagnetic coupling constant, Q_q is the electric charge of the quark and $\hat{s} = (x_A P_A + x_B P_B)^2 = x_A x_B s$ is the centre-of-mass energy squared of the parton-parton collision.

At large energies s-channel Z exchange contributes to the Drell-Yan cross section and Eq. (2.2) must be supplemented by the Z exchange and the $\gamma^* Z$ interference terms:

$$\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-} = \\ \frac{4\pi\alpha^2}{9\hat{s}}\Big(Q_q^2 - Q_q \frac{\sqrt{2}G_F M_Z^2}{4\pi\alpha} g_V^\ell g_V^q \Re(\chi(\hat{s})) + \frac{G_F^2 M_Z^4}{8\pi^2\alpha^2}({g_V^\ell}^2 + {g_A^\ell}^2)({g_V^q}^2 + {g_A^q}^2)|\chi(\hat{s})|^2\Big) \,, \tag{2.3}$$

where G_F is the Fermi constant, g_V and g_A are the vector and axial couplings of the Z boson to leptons and quarks, respectively. The propagator term $\chi(\hat{s})$ is given by

$$\chi(\hat{s}) = \frac{\hat{s}}{\hat{s} - M_Z^2 + iM_Z\Gamma_Z} \tag{2.4}$$

with the total decay width Γ_Z.

For $M_{\ell\ell} \sim M_Z$ the cross section is dominated by Z exchange. As the Z decay width $\Gamma_Z = 2.50\,\text{GeV}$ is small compared to its mass, one can treat the Z boson as an effectively stable particle, and thus factorize the lepton pair production cross section into the Z production cross section $\hat{\sigma}_{q\bar{q}\to Z}$ and the branching fraction $Br(Z \to \ell^+\ell^-)$ for the decay $Z \to \ell^+\ell^-$. In the narrow width approximation the Breit-Wigner distribution

in $|\chi(\hat{s})|^2$ can be related to the δ function and the Z production cross section is given by

$$\hat{\sigma}_{q\bar{q}\to Z} = \frac{\sqrt{2}\pi}{3} G_F M_Z^2 ({g_V^q}^2 + {g_A^q}^2) \delta(\hat{s} - M_Z^2) \ . \tag{2.5}$$

The branching fraction into a fermion pair $f\bar{f}$ is the ratio of the partial width $\Gamma(Z \to f\bar{f})$ to the total width Γ_Z. The partial width can be expressed at leading order (and neglecting the mass of the decay products) as

$$\Gamma(Z \to f\bar{f}) = N_C \frac{G_F M_Z^3}{6\sqrt{2}\pi} ({g_V^f}^2 + {g_A^f}^2) \ , \tag{2.6}$$

where the colour factor N_C is 1 for leptons and 3 for quarks, respectively. Thus the branching fraction of the Z to decay into a $\ell^+\ell^-$ pair is only $Br(Z \to \ell^+\ell^-) \approx 3.4\%$ (for each charged lepton flavour), whereas the hadronic decay is dominating, as $Br(Z \to q\bar{q}) \approx 70\%$.

In analogy, the production of the W boson at a hadron collider can be described within the Drell-Yan mechanism. Within the narrow width approximation the cross section for the subprocess $q\bar{q}' \to W$ can be written as

$$\hat{\sigma}_{q\bar{q}'\to W} = \frac{\sqrt{2}\pi}{3} G_F M_W^2 |V_{qq'}|^2 \delta(\hat{s} - M_W^2) \ , \tag{2.7}$$

where $|V_{qq'}|$ denotes the corresponding element of the Cabbibo-Kobayashi-Maskawa (CKM) matrix [20, 21].

The partial width for the decay of the W boson in a fermion pair $f\bar{f}'$ is given by

$$\Gamma(W^- \to f\bar{f}') = N_C \frac{G_F M_W^3}{6\sqrt{2}\pi} \ , \tag{2.8}$$

where N_C is the colour factor defined above. Thus the branching fraction of the leptonic decay modes $W^- \to e^-\bar{\nu}_e,\ \mu^-\bar{\nu}_\mu,\ \tau^-\bar{\nu}_\tau$ is 1/9 each, and the hadronic decay modes add up to 6/9.

Despite the large contribution of the hadronic decay modes of the Z and W bosons, those can usually not be detected at a hadron collider due to very large background from ordinary QCD dijet production. However, it should be noted that the hadronic decay mode was observed by the UA2 collaboration as a statistically significant excess in the dijet mass distribution [22, 23]. At the Tevatron a significant signal for $Z \to b\bar{b}$ was detected by the CDF and D0 collaborations [24, 25]. For the Z boson decays into a $b\bar{b}$-pair, the signal can be enhanced over the QCD background by requiring b-tags, which exploit characteristic features of B meson decays (cf. Section 4.2.2).

2.2 Perturbative QCD corrections

Without any perturbative and non-perturbative corrections the leading order $2 \to 1$ processes $q\bar{q} \to Z/\gamma^*$ and $q\bar{q}' \to W$ would result in the production of vector bosons with vanishing transverse momentum. In reality, fixed target data on Drell-Yan lepton-pair production [26] as well as the measurements of W and Z/γ^* production at the

Tevatron [27–34] show that the cross sections are concentrated at moderate, but non-vanishing transverse momenta $0\,\text{GeV} < p_T \ll M_V$ and that they exhibit long tails up to large $p_T \gg M_V$ (cf. Section 3.4.3). Here M_V denotes the invariant mass of the lepton-pair or the vector boson mass.

The region of small transverse momenta can be well modelled with the assumption of an intrinsic transverse momentum k_T of the parton relative to the direction of the parent hadron. A Gaussian intrinsic k_T proportional to $\exp(-bk_T^2)$ will generate a differential distribution, which is exponentially decreasing with increasing p_T as

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2 p_T} \propto \exp(-\frac{b}{2}p_T^2) \; . \tag{2.9}$$

Fits to the transverse momentum spectra of Drell-Yan lepton-pairs produced in pN collisions [35] suggest an average $\langle k_T \rangle = 760\,\text{MeV}$ [19]. Data on Z boson production at the Tevatron indicate a slightly larger intrinsic k_T. Using the parton-shower event generator Pythia [36] (cf. Section 2.3) to match the measured transverse momentum spectrum of the Z boson an increased value for $\langle k_T \rangle = 1.3\,\text{GeV}$ is preferred.

The bosons can also be produced with large transverse momenta, namely in $2 \to 2$ processes which are present at higher orders. The perturbative corrections to the Drell-Yan process at next-to-leading-order (NLO) in QCD are shown in Fig. 2.2. In addition to the virtual loop corrections, the real corrections lead to the $2 \to 2$ scattering processes $q\bar{q}' \to Vg$ and $qg \to Vq'$, where V denotes the vector boson. For example, the matrix-element for $q\bar{q}' \to Vg$ is of the form

$$|\mathcal{M}^{q\bar{q}'\to Vg}|^2 \propto \frac{t^2 + u^2 + 2M_V^2 s}{tu} \; , \tag{2.10}$$

where s, t, and u denote the usual Mandelstam variables of the $2 \to 2$ scattering process. Thus, for very large $p_T \gg M_V$ these real correction will lead to a transverse momentum distribution [19]

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2 p_T} \sim \frac{\alpha_s(p_T^2)}{p_T^4} \; , \tag{2.11}$$

which is by far harder (i.e. having a higher rate at large p_T) than the Gaussian distribution originating from the intrinsic k_T of the partons.

The next-to-next-to-leading order (NNLO) calculation of the inclusive cross section of the Drell-Yan process has been available for several years now [37,38]. More recently, also the rapidity distribution (cf. Section 3.4) has been calculated to NNLO [39].

For relatively small transverse momenta of the vector boson $p_T \ll M_V$ (but still keeping $p_T \gg k_T$ to allow a perturbative description) higher-order terms in the perturbation series beyond the NNLO expansion cannot be neglected, as the emission of multiple soft gluons become important. The leading contributions of these emissions to the cross section are of the form [19]

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_T^2} \propto \frac{1}{p_T^2}\alpha_s^n \ln^{2n-1}\frac{M_V^2}{p_T^2} \; . \tag{2.12}$$

These leading logarithms can be summed to all orders such that the transverse momentum spectrum of the vector boson is given by

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_T^2} \simeq \frac{\mathrm{d}}{\mathrm{d}p_T^2}\exp\left(-\frac{\alpha_s}{2\pi}C_F \ln^2\frac{M_V^2}{p_T^2}\right) \; . \tag{2.13}$$

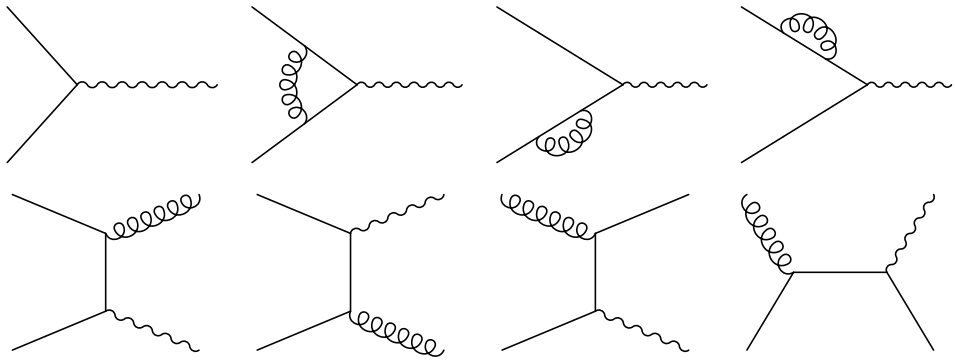

Figure 2.2: The leading-order and next-to-leading-order Feynman diagrams for the Drell-Yan process. The boson decay is not shown.

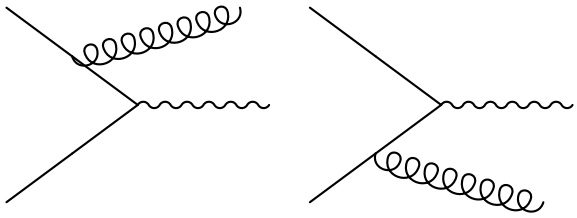

Figure 2.3: Leading-order diagrams for the production of a vector boson in association with a jet. The diagrams are equivalent to the graphs for $q\bar{q} \to Vg$ in Fig. 2.2.

This technique is commonly called resummation. The simplistic model as described above has been largely extended in more complete analyses [40–42]: The resummation is expressed as a Fourier transform in the two-dimensional impact-parameter space to properly treat momentum conservation. Corrections to account for fixed order perturbative results and non-perturbative parameters are added.

2.2.1 Associated production of vector bosons with jets

The transverse momentum of the vector boson must be balanced by the hadronic rest due to momentum conservation. In fact, the $2 \to 2$ processes shown in Fig. 2.2 generate an additional parton, which hadronizes into a jet (cf. Section 4.1), in association with the boson. It is instructive to redraw these diagrams as depicted in Fig. 2.3. The final-state gluon can also be thought of as a parton being radiated off the initial quark lines (initial state radiation, ISR). The corresponding diagrams in Fig. 2.2 and Fig. 2.3, respectively, are identical with respect to their representation of a Feynman calculus. Depending on the transverse momentum and the rapidity of the additional parton, one of the two pictures is more illustrative: At low p_T or forward/backward rapidities, where the outgoing parton is produced at a small angle relative to the incoming parton, the process resembles more the picture of initial state radiation. At large p_T or central rapidities, the representation of V+jet production as $2 \to 2$ process is more instructive.

The matrix-element for $q\bar{q}' \to Vg$ (given in Eq. (2.10)) diverges if $t \to 0$ or $u \to 0$, which corresponds to a final state gluon which is collinear to one of the incoming partons

or which has vanishing momentum. A similar behaviour can be seen for the $qg \to Vq'$ sub-process. Thus, to obtain a meaningful and finite result for the cross section of V+jet production a cut on the minimal transverse momentum for the additional parton needs to be applied. It should be noted that the divergences in V+jet production, which correspond to the real NLO corrections to inclusive vector boson production, exactly cancel with the virtual corrections to the inclusive process [19].

The addition of further partons to the final state lead to multijet production in association with a vector boson. As noted before these processes are of particular interest as they constitute one of the most important backgrounds to a large variety of standard model and beyond the standard model processes.

The inclusive vector boson production cross section can be formally decomposed into its exclusive multijet components

$$\sigma_V = \sum_{n_{excl}=0}^{\infty} \sigma_{V+n_{excl}\,\text{jets}} \ , \tag{2.14}$$

where

$$\sigma_{V+n_{excl}\,\text{jets}} = \sum_{l=n_{excl}}^{\infty} a_{n,l}\alpha_s^l \ . \tag{2.15}$$

Note that the associated production of n jets is at least of order n in the strong coupling constant. Here a jet corresponds to either a quark or gluon at parton level. The inclusive jet multiplicity n_{incl} is obtained by summing all contributions with $n_{excl} \geq n_{incl}$:

$$\sigma_{V+n_{incl}\,\text{jets}} = \sum_{n_{excl}=n_{incl}}^{\infty} \sigma_{V+n_{excl}\,\text{jets}} \ . \tag{2.16}$$

The coefficients $a_{n,l}$ in the above expansion in general depend on the particular choice for the definition of a *jet*, in particular the jet algorithm, cone size, and kinematic cuts, e.g. minimal jet transverse energy and separation cuts. However, the sum of the coefficients at each fixed order l in α_s, $\sum_{n=l}^{\infty} a_{n,l}$, must be independent of the jet definition as it corresponds to the perturbative expansion of the total inclusive V production cross section.

The leading term to the cross section of $V + n$ jet production, both defined in an exclusive or inclusive way, is

$$\sigma_{V+n\,\text{jets, LO}} = a_{n,n}\alpha_s^n \ , \tag{2.17}$$

thus, in particular it is proportional to the n-th power of the strong coupling constant α_s. These leading order contributions have been calculated from the tree-level matrix-element for all parton configurations which lead to a $V + n$ jet final state. Note again, that *jet* denotes any (anti-)quark or gluon. For high jet multiplicities these calculation get increasingly complicated as the number of Feynman diagrams which need to be calculated has a factorial growth. Currently, final states with up to 6 partons in association with a W or Z bosons can be calculated at leading order [43]. These calculations do not make explicit use of Feynman diagrams, but instead use an algorithm based on an recursive evaluation of the scattering matrix [44], which has the

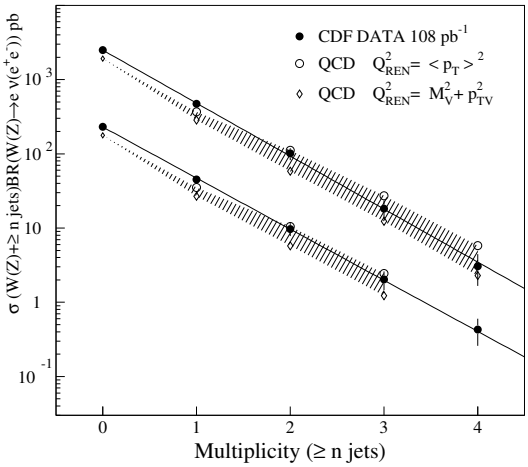

*Figure 2.4: $\sigma \cdot Br$ for $W^{\pm} \to e^{\pm}\nu + \geq n$ jets (top) and $Z \to e^{+}e^{-} + \geq n$ jets (bottom) versus inclusive jet multiplicity [47]. The solid lines are exponential fits to the data. The data are compared to LO calculations (*VECBOS*) using different renormalization scales [48].*

advantage that its complexity increases only with the power of the number of outgoing partons instead of a factorial growth. Next-to-leading order calculations have been performed for final states with up to two jets [45].

Fig. 2.4 shows the inclusive jet multiplicity in W and Z boson production measured by CDF in $p\bar{p}$ collisions at $\sqrt{s} = 1.8\,\text{TeV}$ (Tevatron Run I) [46, 47]. Jets are defined using a cone algorithm (cf. Section 4.1) with cone size $R = 0.4$ and are required to have a transverse energy $E_T > 15\,\text{GeV}$ and pseudo-rapidities $\eta < 2.4$. The data are compared to the tree-level predictions of the VECBOS Monte Carlo program [48] using two different choices for the renormalization scale, $Q^2_{\text{ren}} = M_V^2 + p_{T,V}^2$ and $Q^2_{\text{ren}} = \langle p_T \rangle^2$, with $\langle p_T \rangle$ being the average transverse momentum of the jets. As expected the leading-order predictions have a sizable dependence on the renormalization scale which increases with the number of jets, i.e. with the power in α_s.

The cross section as function of inclusive jet multiplicity n can be well parameterized by an exponential, $\sigma_n = \sigma_0 \times a^n$, with $a = 0.20$. Certainly, the value of a depends considerably on the jet definition and the kinematic cuts applied. This geometrical relation, sometimes called 'Berends scaling', has been suggested before by tree-level calculations [48–50]. Note, however, that this relation is only an approximation, which can largely degrade depending on the specific kinematic cuts. For example it has been shown in Ref. [48] that the ratio σ_{n+1}/σ_n decreases with n in case of a large jet separation cut $\Delta R(j, j)$. This can be easily understood when considering the phase space for the additional jet, which quickly decreases with increasing number of jets. Eventually, it even becomes impossible to add an extra jet. Evidently, this limit is being reached earlier with increasing jet-jet separation.

The tree-level predictions for the production cross sections of the associated production of vector bosons with jets obviously have several limitations. They neglect

program	matrix-element	full event generator	merging/matching	functionality w.r.t. W/Z production
PYTHIA	LO	yes	ME correction for first branching	inclusive production
HERWIG	LO	yes	ME correction for hardest branching	inclusive production
MC@NLO	NLO	yes (based on HERWIG)	n.a.	inclusive production
RESBOS	resum-mation	no (only boson kinematics)	n.a.	p_T spectrum of W/Z
ALPGEN	LO	no (but interf. to HERWIG or PYTHIA)	MLM, ME-PS match. (all parton multipl.)	W/Z + jets (incl. large multipl.)
SHERPA	LO	yes	CKKW, ME-PS match. (all parton multipl.)	W/Z + jets (incl. large multipl.)
MCFM	NLO	no (only parton level)	n.a.	NLO corrections to integral rates and shapes

Table 2.1: Monte-Carlo programs which can be used for the simulation of Z and W production in hadron collision. For each program, the table lists the following properties: the order (in α_s) of the matrix-element calculation, if the program is a full event generator or can be interfaced to one, the algorithm (if any) which is implemented to match/merge matrix-element (ME) calculations with parton showers (PS), and the main functionality of the program from the perspective of the simulation of W and Z boson production. See text for details.

higher order corrections, which might be substantial. For example, the exclusive and inclusive jet multiplicity are assumed to be identical. The calculations have a substantial scale dependence, especially at large jet multiplicities n as the cross section is proportional to α_s^n. Finally these parton-level calculations need to be complemented with an evaluation of the full hadronic final state, which involves a consistent merging of the matrix-element calculation with a parton shower evolution.

2.3 Simulation and event generation

In this section, methods and tools (with emphasis on available computer codes) for the calculation of cross sections and the event generation for vector boson production are reviewed. The focus will be on QCD effects, in particular in the production of W and Z bosons together with jets. Table 2.1 gives an overview of simulation programs which will be discussed in more detail below.

2.3.1 Fixed order calculation tools

At parton level, inclusive and differential cross sections for vector boson production in association with jets can be evaluated using various computer codes. The VECBOS program [48], which was already mentioned before, can be used to calculate cross sections for up to four additional partons together with the W or Z boson. This program was extensively used during Tevatron Run I, in particular for background studies preceding the observation of the top quark [51, 52]. Nowadays, several multi-purpose tree-level programs are available: COMPHEP [53, 54] (up to two additional partons), GRACE/GR@PPA [55], MADGRAPH/MADEVENT [56, 57] (both including up to four additional partons), AMEGIC++ [58] (limited to up to about four additional partons[1]) and ALPGEN [43] (up to six additional partons). It has been demonstrated that the cross section predictions of these programs agree [59]. To facilitate the calculation, the partons (i.e. gluons and quarks) are typically treated as massless in the matrix-element evaluation. For the explicit calculation of vector boson production in association with heavy flavour quarks (e.g. Wc, $Wc\bar{c}$, $Wb\bar{b}$, etc.) the quarks are treated as massive partons, which gives an improved description near the kinematic boundaries and allows the generation without applying cuts on the minimum transverse momentum and angular separation of the heavy flavour quarks.

The parton-level codes have been interfaced to parton shower event generators, which generate full events including initial and final state radiation as well as hadronization. With the exception of AMEGIC++ which is part of the SHERPA event generator [60], PYTHIA [36] or HERWIG [61, 62] can be used for the subsequent event generation starting from the parton level configuration. The matching of matrix-element calculations with parton showers will be further discussed below.

2.3.2 Higher order calculation codes

The next-to-leading order correction to $V + n$ parton production consists of real radiation leading to a $V + (n+1)$ parton final state and virtual loop corrections. Whereas the former can be evaluated using the tree-level codes, provided that the additional partons fulfil angular and momentum cuts, the full next-to-leading order calculation has only been performed for final states with up to two partons in association with the vector boson. These calculations are implemented in the computer code MCFM [45,63].

Fig. 2.5 shows the transverse momentum distribution for the leading jet in W + 1 jet and W + 2 jet events in proton-antiproton collisions at $\sqrt{s} = 2\,\text{TeV}$ evaluated with MCFM at leading and next-to-leading order. In contrast to the leading-order calculation which does not differentiate between the inclusive and exclusive definition of jet multiplicities, at next-to-leading order these definitions differ. A real correction might generate an additional jet, provided that the additional parton is created with sufficient energy and separation. Therefore it is essential to specify if a NLO correction is based on an inclusive or exclusive definition of jet multiplicities. Fig. 2.5 shows that the NLO correction (exclusive definition) significantly softens the transverse momentum distribution of the leading jet, in particular for W + 1 jet production. At high p_T a jet is more likely to radiate a parton which passes the kinematic cuts to count

[1] The limitation is only practical due to common limitations on computer memory resources.

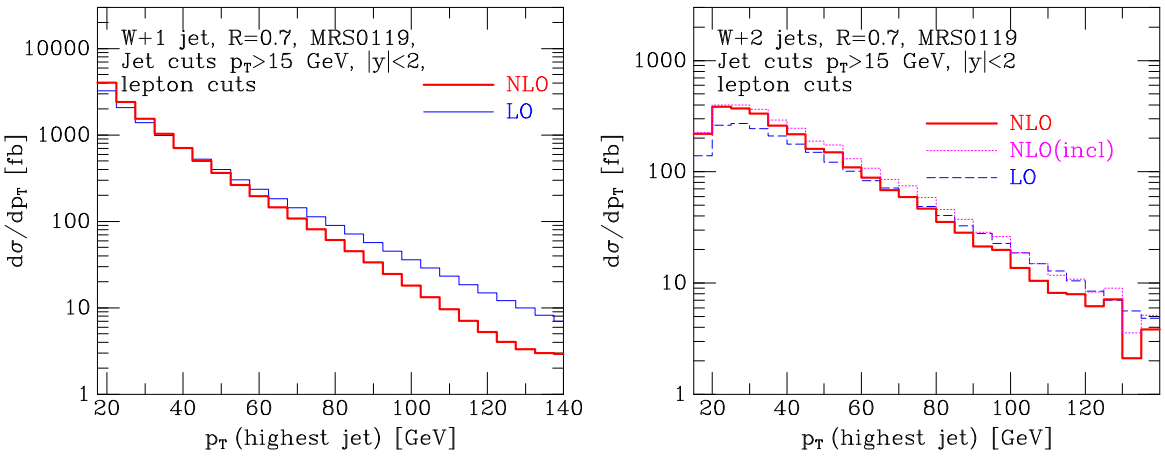

Figure 2.5: The transverse momentum distribution of the leading jet in W + 1 jet (left) and W + 2 jet (right) events in $p\bar{p}$ collisions at $\sqrt{s} = 2\,\text{TeV}$ predicted at LO and NLO by MCFM *[45]. Shown is the exclusive NLO calculation in both cases, and for W + 2 jets in addition the inclusive NLO prediction (i.e. where additional partons from real emissions are allowed). A renormalization/factorization scale of $\mu = 80\,\text{GeV}$ has been chosen.*

as an extra jet, so that the event is removed from the original sample. In case of the inclusive definition of jet multiplicities, the NLO correction softens the leading jet p_T distribution only to a considerably smaller amount. Here no event is removed from the original sample but a radiated parton can decrease the energy of the leading jet.

In case of $Wb\bar{b}$ and $Zb\bar{b}$ production, it has been shown that the next-to-leading order corrections can be rather large (depending on kinematic cuts) and that they considerably modify the shape of the b jet p_T and dijet mass $M_{b\bar{b}}$ distributions [64, 65]. It should be noted that MCFM does only provide the NLO result for the assumption of massless b quarks. However, a recent study demonstrated that the inclusion of massive b quarks have only a modest effect for not too low $M_{b\bar{b}}$ [66].

At next-to-next-to-leading order (NNLO), computer codes have been made public for inclusive Z and W production [37] as well as their rapidity distribution [39]. The NNLO corrections, especially on kinematic shapes, were found to be relatively small.

2.3.3 Resummed calculations

As described in Section 2.2 resummed calculations are well suited to describe the transverse momentum spectrum in vector boson production. The program RESBOS is publicly available [67] and incorporates the resummed calculations in form of grid files as provided by the LEGACY code [42, 68], which are matched to the NLO perturbative QCD calculation at large boson momentum. RESBOS, which also includes non-perturbative parameters, gives a good description of the Z boson p_T spectrum (cf. Section 3.4.3).

2.3.4 Parton shower models

In contrast to fixed order calculations which evaluate differential cross sections only at the parton level and resummed calculations which only describe the kinematics of the vector boson (and its decay leptons) while ignoring the hadronic rest, parton showers in combination with fragmentation models simulate full events at particle level. Thus, parton shower models are a crucial ingredient in event generators as PYTHIA [36, 69], HERWIG [61, 62], and SHERPA [60].

Similar to resummed calculations, parton showers provide an all-order approximation in the soft and collinear region. In the leading-logarithmic picture, the shower can be regarded as a sequence of $1 \to 2$ splittings of a mother parton a into two daughters b, c: $a \to bc$. Perturbative QCD includes the following fundamental $1 \to 2$ branchings: $q \to qg$, $g \to gg$, and $g \to q\bar{q}$. Parton shower algorithms might in addition include QED branchings, e.g. the model implemented in PYTHIA [36, 69] includes photon radiation off quarks and leptons as well, i.e. $q \to q\gamma$ and $\ell \to \ell\gamma$.

In the following, the main aspects of parton shower models are shortly described as well as some concepts which are also relevant for the succeeding discussion on the matching of fixed-order matrix-element calculation with parton showers. A detailed description of the parton shower models can be found in Refs. [36,60,61] and references therein. Here, the notation of Ref. [36] is being used.

The kinematics of each branching is defined by two variables, the evolution scale Q^2 and the relative momentum fraction z. Q^2 is related to the virtual mass of the parton a to branch or the transverse momentum scale of the branching. The different models differ in the exact definition of Q^2. Assuming a branching $a \to bc$, parton b will carry a fraction z of a's momentum and c the fraction of $(1 - z)$. It is convenient to normalize the scale Q^2 to the QCD scale Λ and define an evolution variable $t = \ln(Q^2/\Lambda^2)$. The branching of partons is described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [70–72] thus that the differential probability $\mathrm{d}\mathcal{P}_a$ for a branching $a \to bc$ in terms of the two variables t and z is given by

$$\mathrm{d}\mathcal{P}_a = \sum_{b,c} \frac{\alpha_s(t)}{2\pi} P_{a\to bc}(z)\mathrm{d}t\mathrm{d}z \ . \tag{2.18}$$

Here, the sum runs over all possible branchings and the functions $P_{a\to bc}(z)$ denote the appropriate DGLAP splitting kernels. For simplicity, only QCD branchings in Eq. (2.18) are considered, so that the vertex coupling can be specified with α_s.

The parton shower is developed with respect to the evolution variable t. In general the shower history can either follow decreasing or increasing t. Consider a hard interaction at scale Q^2_{hard}: The final-state evolves from the high scale t_{hard} down to a lower cut-off. The initial-state shower connects the parton within a hadron at a low scale to the hard interaction following increasing values of t. For a given t, the differential branching probability (with respect to a differential range $\mathrm{d}t$) can be obtained by integrating Eq. (2.18) over the allowed z region:

$$\mathcal{I}_{a\to bc}(t) = \int_{z_-(t)}^{z_+(t)} \mathrm{d}z \frac{\alpha_s(t)}{2\pi} P_{a\to bc}(z) \ . \tag{2.19}$$

A essential ingredient of parton showers (and also resummed calculations) is the concept of the Sudakov form factor [73], which describes the probability $\mathcal{P}_{\text{no-branching}}(t_0, t)$ for a parton a at a reference scale t_0 *not* to have branched when it is evolved to the scale t. This Sudakov form factor is derived by exponentiating the differential no-branching probability within a range $\mathrm{d}t$, given by $(1 - \sum_{b,c} \mathcal{I}_{a\to bc}(t)\mathrm{d}t)$, so that

$$\mathcal{S}_a(t) := \mathcal{P}_{\text{no-branching}}(t_0, t) = \exp\left\{ - \int_{t_0}^{t} \mathrm{d}t' \sum_{b,c} \mathcal{I}_{a\to bc}(t') \right\} . \tag{2.20}$$

The actual probability for a branching of parton a at the scale t is then directly given by the derivative of the Sudakov form factor $\mathcal{S}_a(t)$:

$$\frac{\mathrm{d}\mathcal{P}_a}{\mathrm{d}t} = -\frac{\mathrm{d}\mathcal{P}_{\text{no-branching}}(t_0, t)}{\mathrm{d}t} = \left(\sum_{b,c} \mathcal{I}_{a\to bc}(t) \right) \mathcal{S}_a(t) . \tag{2.21}$$

That means that the naive expectation for the branching probability, which is given by the first factor in the equation, $\sum_{b,c} \mathcal{I}_{a\to bc}(t)$, is suppressed to take into account potential branchings which could have happened before. This effect is often referred to as Sudakov suppression.

For our discussion the initial-state shower is of high relevance as a branching in the initial state can be regarded as a radiation of a quark or a gluon in the initial state which will lead to the associated production of jets as discussed in Section 2.2.1.

In principle the initial-state parton shower could be regarded as a probabilistic evolution of parton densities from the initial scale Q_0^2 to the scale of the hard interaction, Q^2_{hard}, following the DGLAP approach, where all possible cascades leading to a defined set of partons taking part in the hard scattering are considered. In practise this *forward* evolution scheme is too difficult to implement and an exact treatment of the kinematics is apparently not possible.

Instead, Monte Carlo implementations of initial-state parton showers commonly use a *backward* evolution scheme. Consider a branching $a \to bc$, such that a parton a with momentum fraction x_a is resolved into a parton b at $x_b = z x_a$ and the second daughter c at $x_c = (1 - z)x_a$. In the backward evolution, when t is decreased, the parton b may be 'unresolved' into parton a with a relative probability $\mathrm{d}\mathcal{P}_b = \mathrm{d}f_b(x_b, t)/f_b(x_b, t)$ which is derived using the DGLAP evolution equations:

$$\mathrm{d}\mathcal{P}_b = \frac{\mathrm{d}f_b(x_b, t)}{f_b(x_b, t)} = |\mathrm{d}t| \sum_{b,c} \int \frac{\mathrm{d}x_a}{x_a} \frac{f_a(x_a, t)}{f_b(x_b, t)} \frac{\alpha_s(t)}{2\pi} P_{a\to bc}\left(\frac{x_b}{x_a}\right) . \tag{2.22}$$

As seen before, the probability for no branching exponentiates and the probability that a parton b remains at x_b when evolving from t_1 to $t_2 < t_1$ can be expressed in form of a Sudakov-type form factor

$$\mathcal{S}_b(x_b, t_1, t_2) := \exp\left\{ - \int_{t_2}^{t_1} \mathrm{d}t' \sum_{a,c} \int \frac{\mathrm{d}x_a}{x_a} \frac{f_a(x_a, t')}{f_b(x_b, t')} \frac{\alpha_s(t')}{2\pi} P_{a\to bc}\left(\frac{x_b}{x_a}\right) \right\} \tag{2.23}$$

$$= \exp\left\{ - \int_{t_2}^{t_1} \mathrm{d}t' \sum_{a,c} \int \mathrm{d}z \frac{\alpha_s(t')}{2\pi} P_{a\to bc}(z) \frac{x_a f_a(x_a, t')}{x_b f_b(x_b, t')} \right\} . \tag{2.24}$$

The main difference with respect to the original Sudakov form factor $\mathcal{S}_a$ given in Eq. (2.20) is that the probability is proportional to the ratio of the distribution functions of partons a and b. This can be intuitively understood: The probability for parton b to originate from parton a is proportional to the density of a, $f_a(x_a, t')$, and inversely proportional to the density of b, $f_b(x_b, t')$. The latter is due to the fact that the overall probability is divided between all partons of type b.

In analogy to Eq. (2.21), the probability that a parton b, which either takes part in a hard scattering or branches at scale t', was produced in a previous branching at lower scale t can be derived directly from the form factor $\mathcal{S}_b(x_b, t', t)$:

$$\frac{\mathrm{d}\mathcal{P}_b}{\mathrm{d}t} = -\frac{\mathrm{d}\mathcal{S}_b(x_b, t', t)}{\mathrm{d}t} = \left(\sum_{a,c} \int \mathrm{d}z \frac{\alpha_s(t')}{2\pi} P_{a\to bc}(z) \frac{x_a f_a(x_a, t')}{x_b f_b(x_b, t')} \right) \mathcal{S}_b(x_b, t', t) \ . \quad (2.25)$$

The modification to the Sudakov form factor with the ratio of parton densities (cf. Eq. (2.24)) is not generally adopted in initial state parton shower algorithms, although this procedure correctly describes the non-branching probability. Whereas Pythia uses this modified Sudakov form factor, Herwig uses an identical Sudakov for initial and final state showers. Nevertheless, Herwig includes the ratio of parton densities in the calculation of differential branching probabilities similar to Eq. (2.25), although this is formally not identical to the derivative of the Sudakov form factor.

2.3.5 Combining matrix-element computations with parton shower models

As discussed before, fixed-order matrix-element calculations and parton shower algorithms have their strengths in different kinematic regimes. The perturbative calculation of a matrix-element is a very good approximation at large transverse momenta and large parton-parton separation. Also, it accounts for interference effects in the hard interaction, e.g. amplitudes of different processes leading to the same final state are correctly added. As matrix-element calculations turn to be increasingly complex at higher orders, they have a practical limit at large parton multiplicities. Contrary to these calculations, the complexity of parton showers only mildly increases with parton multiplicity as the branching probability is evaluated separately for each parton leg. The leading-log approximation implemented in parton showers provides a good description in kinematic regions dominated by soft and collinear radiation, but only poorly describes parton radiation at large momenta and large angles. Thus, the kinematic domains of the parton shower algorithms and matrix-element calculations, respectively, are complementary, which suggests a combination of both approaches to optimally describe processes leading to multi-parton final states.

Already several years ago, various techniques have been developed to combine matrix-element and parton-shower methods. *Matching* algorithms introduce an intermediate scale at which the transition from one method to the other takes place [74–76]. These methods are well suited to describe exclusive jet topologies, but might suffer from discontinuities around the transition scale.

Less general are *merging* strategies, in which the parton shower is modified using weights obtained from matrix-element calculations. This method has been explicitly

applied to multijet production in electron-positron annihilation (merging of $e^+e^- \to q\bar{q}$ with $e^+e^- \to q\bar{q}g$) [75, 77] and vector boson production at hadron colliders (merging of $q\bar{q}' \to V$ with $q\bar{q}' \to Vg$ and $qg \to Vq'$) [78, 79]. It is implemented for these processes in the parton shower event generators PYTHIA [69] and HERWIG [62]. Whereas in PYTHIA only the first branching is corrected, in HERWIG any emission, which could be the hardest, is modified. The correction is applied using a weight corresponding to the ratio of rates predicted by the matrix-element calculation and the parton shower, respectively. As these models only take into account matrix-element corrections for one additional parton, they give a fair description of the emission of one hard jet in association to a vector boson (with possible additional soft jets) but fail to correctly describe the emission of multiple hard jets, in particular they underestimate the jet multiplicity in vector boson production (see Section 4).

Recently, more general matching schemes have been developed which combine matrix-element calculations for various parton multiplicities with parton showers and which explicitly avoid double counting[2] by means of weighting procedures or veto algorithms. These models address in particular the associated production of multiple jets with vector bosons at hadron colliders and attempt to give a full description of inclusive W and Z boson production in the entire kinematic region, i.e. from zero jet emission up to multiple jet emission without discontinuities in the transition regions. In contrast to these schemes, the merging algorithms implemented in PYTHIA and HERWIG only attempt to correctly describe the emission of a single hard jet.

Of particular interest for the event generation at Tevatron and LHC experiments are the Catani-Krauss-Kuhn-Webber (CKKW) matching scheme [80, 81], which is implemented in the SHERPA Monte Carlo program [60], and the Mangano (MLM) scheme, which is implemented in the ALPGEN event generator [43]. Using the example of W boson production in proton-(anti-)proton collisions, $p^{(}\bar{p}^{)} \to W + n$ jets, the general strategy of these matching procedures can be described as follows [82]:

1. A matching scale Q_0 and a jet algorithm is defined, and the cross sections for $W + n_i$ jet production is evaluated for each n_i up to a maximal value $n_{\max}$ using matrix-elements.

2. A parton configuration is chosen based on the relative contribution of the $W + n_i$ cross section and the kinematic configuration is given by the matrix-element.

3. A weight is calculated which accounts for Sudakov-type suppressions and for the choice of the scale used in the calculation of the running strong coupling constant α_s and the Sudakov-type suppressions. Based on this weight and a random number the parton configuration is either accepted or rejected. In case of the latter, step 2 is repeated.

4. A parton shower is developed starting with the accepted parton configuration. Radiations above the matching scale are vetoed to avoid double counting. As in the case of the MLM scheme this step can be (partly) combined with the previous step.

[2] The same phase space could be filled by both matrix-element and parton shower.

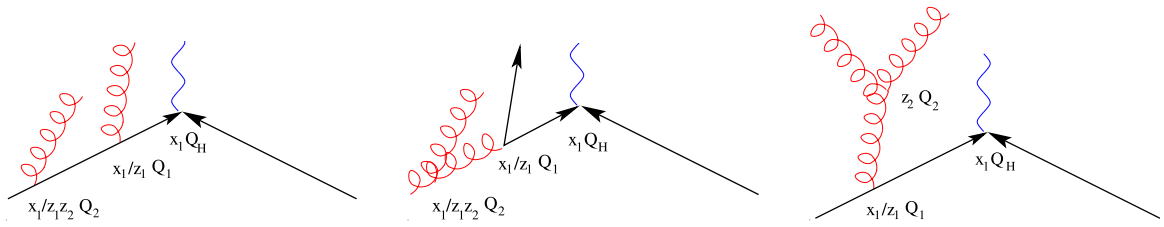

Figure 2.6: Three possible configurations for a V + 2 jet event [81]. The hard interaction takes place at the scale Q_H with an momentum fraction x_1 for the left incoming parton. The other branchings are labelled with the corresponding nodal scales and parton momentum fractions.

Whereas the vetoing of additional jets can be understood intuitively, the weights defined in step 3 are not obvious. They can be motivated by reminding that the parton shower provides a good description when the scale of the branching is small and there is a strong ordering in the scales of subsequent branchings. This behaviour of the parton shower can be approximated by the matrix-element calculation provided that the strong-coupling constant is chosen according to the scale of each branching and Sudakov form factors are added to all quark and gluon lines to account for the non-vanishing branching probability for this leg.

The CKKW algorithm, which has been initially introduced for multi-parton production in electron-positron annihilation and proven here to be accurate to the precision of next-to-leading-logarithms (NLL) [80], has been extended to hadron-hadron collisions by one of the original authors [81]. This scheme uses a k_T-measure to assign partons to jets [83] and step 3 of the procedure can be sketched as follows: A *shower history* is reconstructed by clustering the initial and final state partons of a configuration given by the matrix-element using the k_T-algorithm. Nodal scales $Q_1 > Q_2 > ... > Q_n > Q_0$ are defined as the scales at which 1, 2, ... , n jets are resolved. A coupling constant weight is defined as $\alpha_s(Q_1) \times ... \times \alpha_s(Q_N)/\alpha_s(Q_0)^n$, i.e. relative to the fixed scale Q_0, which is used in the evaluation of the matrix-element in step 1. For each quark and gluon line a Sudakov factor (in NLL-approximation) is added and the corresponding weight is given by the product of all factors. This weight is modified for the configurations given by the matrix-element corresponding to the highest parton multiplicity to implicitly take into account the possibility to have extra soft jets [84]. Fig. 2.6 shows possible configurations for a V + 2 jet event and denotes the nodal scales for the parton branchings. The procedure to determine the weights for this example is detailed in Refs. [81, 84].

The CKKW scheme is implemented in the SHERPA event generator and, in a modified version, also in the ARIADNE program [85], which is based on the dipole model [86, 87] which describes $2 \to 3$ partonic splittings instead of the $1 \to 2$ branchings in conventional parton shower models. An alternative implementation of the CKKW scheme combines matrix-element calculations obtained with MADGRAPH with the parton showers of HERWIG and PYTHIA, which are also used to provide a numerical approximation of the Sudakov weighting procedure [88].

The MLM matching scheme, which is implemented in ALPGEN (since version 2),

adopts some of the original ideas suggested by CKKW, but uses a parton shower veto algorithm instead of a rigorous Sudakov weighting. A tree structure is build from the initial parton configuration (step 3) in analogy with the CKKW procedure, but using the colour-flow extracted from the matrix-element instead of a k_T measure [89]. The reweighting in α_s is adopted, but instead of using additional weights defined by formal Sudakov factors, an algorithm based on parton shower trial and veto is employed. The particles at the end of the parton shower, given by either HERWIG or PYTHIA, are clustered using a cone jet algorithm instead of the k_T-measure and the jets are matched to the partons from the matrix-element calculation. For all but the highest multiplicity parton configurations, an exclusive matching is applied, where events with additional jets are vetoed. This veto replaces the Sudakov reweighting used in the CKKW scheme. For the highest parton multiplicity, an inclusive matching allows the production of extra jets.

Instead of combining tree-level matrix-element calculations for various parton multiplicities with parton showers, another approach is to match NLO calculations with a parton shower, which has been achieved with the event generator MC@NLO [90], which utilizes the parton shower model of HERWIG. This matching procedure has only been implemented for a limited number of processes including the production of vector bosons. The advantage of MC@NLO compared to the tree-level generators discussed above is that integrated cross sections obtain the higher order correction as well. As the NLO matrix-element for vector boson production does only include the emission of a single additional parton, MC@NLO is not well suited to describe the production of multiple parton emission at large transverse momenta.

2.3.6 Event generators for W/Z + jet production

As discussed in the previous sections, the event generators SHERPA and ALPGEN, provide both matrix-element calculations for large parton multiplicities as well as a matching scheme to combine these matrix-elements with parton shower models. Therefore, these programs are very well suited for the simulation of vector boson production in association with jets. In this section, first systematic studies for the validation of the event generators are reviewed and finally the predictions of SHERPA and ALPGEN are compared to those of the parton shower generator PYTHIA.

The ALPGEN program is discussed as a generator for $W+$ jet production and in particular for the associated production with heavy flavour jets, $Wb\bar{b}+$ jets, in Ref. [89]. Fully showered final states are analysed using HERWIG as parton shower model. The systematics in the mapping of the parton level prediction given by the matrix-element calculation to the configuration after the parton shower are studied. When the parton configuration after the parton shower is clustered into jets using a cone algorithm of size $R_{\text{jet}} = 0.4$ and when it is compared to the matrix-element prediction using a parton separation of the same size $\Delta R = 0.4$, one observes that the parton shower results in an energy loss out of the cone, which is depicted in the left panel of Fig. 2.7. This suggests that a jet cone should be used which is wider than the parton separation for the configurations which are fed into the parton shower. This is demonstrated in the right panel of Fig. 2.7. Here, parton-level events with separation $\Delta R = 0.4$ are showered and subsequently clustered into jets of cone size $R_{\text{jet}} = 0.7$. The resulting jet

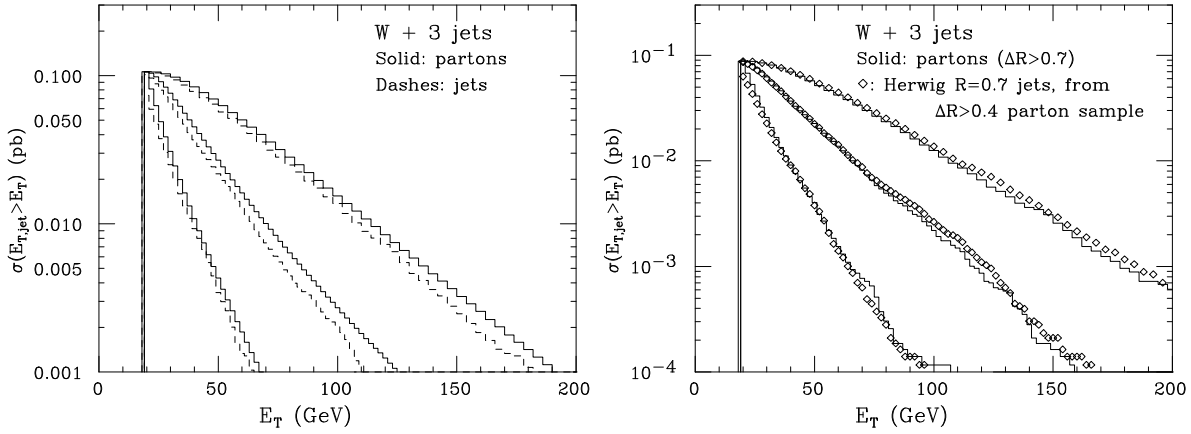

Figure 2.7: Inclusive transverse momenta distributions of jets in W + 3 jet production at the parton level and after shower evolution calculated with ALPGEN *[43] and* HERWIG *[61] for the parton shower. Left: parton level jets are defined by an isolation cut $\Delta R > 0.4$, and showered jets by a cone $R_{\text{jet}} = 0.4$. Right: for the parton level prediction $\Delta R > 0.7$ is chosen; the showered jets with $R_{\text{jet}} = 0.7$ are based on partons generated with $\Delta R > 0.4$ [89].*

transverse energy distributions agree well with the parton-level prediction if the latter has been generated using a parton separation of $\Delta R = 0.7$.

The SHERPA event generator was studied and validated for $W/Z+$ jet production at the Tevatron [91] and LHC [92]. It was verified that the predictions for inclusive observables ($p_T(W)$, $p_T(e)$, $\eta(W)$, $\eta(e)$) and for differential jet rates show only very little dependence on the choice of the matching scale which was varied between 10 GeV and 50 GeV. Also it was demonstrated that the inclusion of higher multiplicities does not significantly change more inclusive variables, which can be interpreted as a verification of the modified matching procedure for the highest multiplicity bin.

The parton level predictions of SHERPA were also compared to the NLO calculations given by MCFM [91]. Here, the event generation was stopped before the parton shower, i.e. the matrix-element was reweighted with the ratios of the strong coupling constant and the ratios of the Sudakov form factors according to the CKKW scheme. The fixed-order LO and NLO results were obtained using M_W for the factorization and renormalization scales. Fig. 2.8 shows a comparison of the leading and trailing jet transverse momenta as predicted by SHERPA compared to the LO and NLO calculation. One can notice that at NLO the jet transverse momentum distribution is significantly softer than the LO prediction, since the probability that a parton is emitted (corresponding to the real NLO correction) which fulfils the criteria for an additional jet is increasing with larger jet momenta. Consequently, the fraction of events removed from the exclusive sample is increasing with jet p_T. For not too large jet p_T, SHERPA agrees well with the NLO calculation, demonstrating that the α_s and Sudakov weighting can *mimic* higher order corrections. At very high transverse momenta, the rate predicted by SHERPA is higher than the NLO prediction, as the scale used for the relevant branching is considerably larger than the fixed scale used in the NLO calculation, which is given by M_W. In fact increasing the scale in the NLO calculation results in

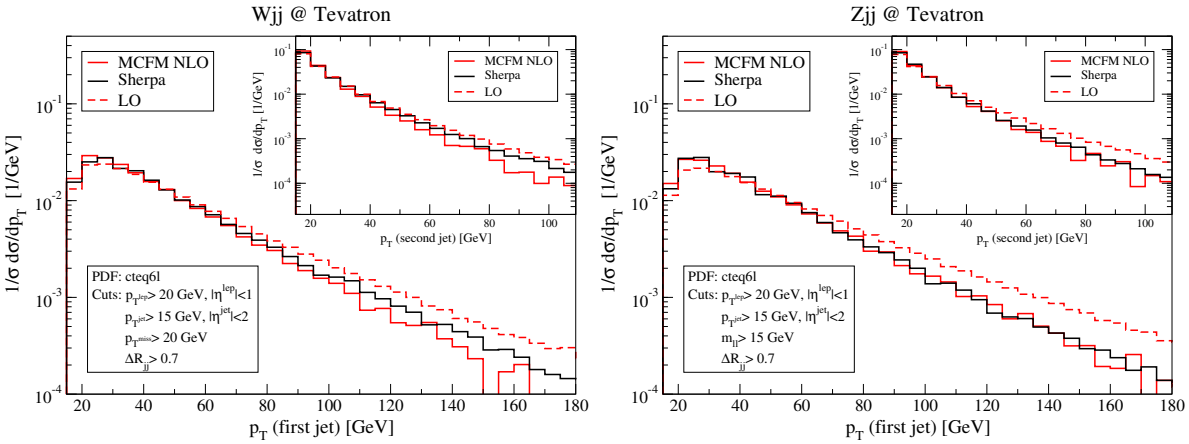

Figure 2.8: The transverse momentum p_T of the first and second jet in exclusive W + 2jet (left) and Z + 2jet (right) production at the Tevatron Run II, predicted by the Sherpa *event generator [91] and compared to LO and NLO given by* Mcfm *[45, 65].*

an improved agreement.

To be able to relate generator predictions to experimental data, kinematic distributions at the particle level, i.e. including fragmentation and hadronization, need to be simulated. In the following, kinematic distributions for Z+ jet production at the Tevatron obtained with Alpgen, Sherpa, and Pythia are compared. A similar comparison, but not including Pythia, was presented in Ref. [82]. Here, the difference of Alpgen and Sherpa to the traditional partons shower generator Pythia, which does not include a full matching scheme, but only matrix-element corrections to the first parton shower branching, is also explicitly demonstrated. In this study inclusive Z/γ^* production (mass range: $60 < M_{ll} < 130\,\text{GeV}$) is simulated and final state particles are clustered using the Pxcone implementation [93] of the cone algorithm using a cone size of $R_{\text{jet}} = 0.5$. For all event generators the underlying event which includes multiple parton interactions was simulated. In addition to the hard process (e.g. $q\bar{q} \to Z$), possible additional interactions between partons from the beam remnants were included. The parameter settings have been chosen to match the values used in the Monte Carlo event generation for the DØ experiment. Generator defaults have been selected with the exceptions noted below:

- Pythia: Version v6.323 is used with CTEQ6L1 PDF sets [94]. The inclusive $2 \to 1$ process is selected (`MSEL 11`) and CDF's Tune A is used for the underlying event simulation [95, 96]. About 600,000 events were processed.

- Alpgen: Version v6.205 with CTEQ6L1 PDF sets is used. Matrix-elements with up to 3 massless partons (g, u, d, s, c) in association with the Z/γ^* are included requiring a minimal transverse momentum of 8 GeV (`ptjmin 8`) and a parton separation $\Delta R > 0.4$ (`drjmin 0.4`). The CKKW scale prescription is selected (`ickkw 1`) and Pythia is used for the parton shower, fragmentation, hadronization and simulation of the underlying event (with Tune A). For the jet

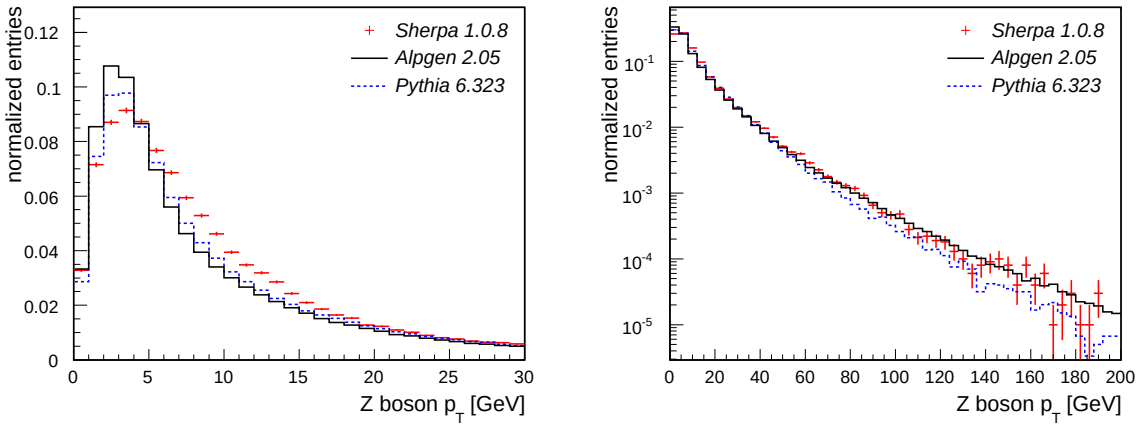

Figure 2.9: Z boson transverse momentum spectrum in $p\bar{p}$ scattering at $\sqrt{s} =$ 1.96 TeV as predicted by Sherpa, Alpgen, *and* Pythia *(left: linear scale, right: logarithmic scale). See text for the parameters used in the generation. For the* Sherpa *prediction statistical errors are shown. The* Alpgen *and* Pythia *distributions are based on significantly larger statistics.*

veto algorithm, jets are defined using a minimum $E_T = 8\,\text{GeV}$, a cone size of $R_{\text{jet}} = 0.4$, and a parton-jet separation of $\Delta R = 1.5 R_{\text{jet}}$ (`etclus 8`, `rclus 0.4`). About 1.5 million events were simulated.

- Sherpa: Version v1.0.8 with CTEQ6L PDF sets[3] is used. Matrix-elements for up to 3 massless partons (g, u, d, s, c) are included and $Q = 15\,\text{GeV}$ is chosen as CKKW matching scale. The underlying event is simulated using the Amisic generator internal to Sherpa. About 100,000 events were processed.

All distributions have been simulated for proton-antiproton collisions at centre-of-mass energies $\sqrt{s} = 1.96\,\text{TeV}$, i.e. the situation at Tevatron Run II, and they have been normalized to unity to allow direct shape comparisons. The verification of the event generators with experimental data will be discussed in the following chapters.

Fig. 2.9 presents the transverse momentum spectrum of the Z boson. At very large $p_T(Z)$ the event generators Alpgen and Sherpa predict a higher rate as Pythia as final states with multiple hard radiated partons, which are absent in Pythia, have an increasing contribution. At low $p_T(Z)$ Pythia and Alpgen agree and both differ significantly from Sherpa. This can be understood as in this kinematic region the shape of the transverse momentum distribution is sensitive to details of the parton shower model and to assumptions on the intrinsic k_T of the partons within the proton.

The exclusive jet multiplicities for lower transverse energy cuts $E_T > 15\,\text{GeV}$ and $E_T > 25\,\text{GeV}$, respectively, are shown in Fig. 2.10. As the parton shower in Pythia includes a correction to the first branching, one expects that the rate of one jet events

[3]The CTEQ6L set differs from CTEQ6L1 only in the detail that for the PDF fit a NLO evolution for the running of α_s is used instead of the LO one. Both PDF sets are based on LO matrix-elements.

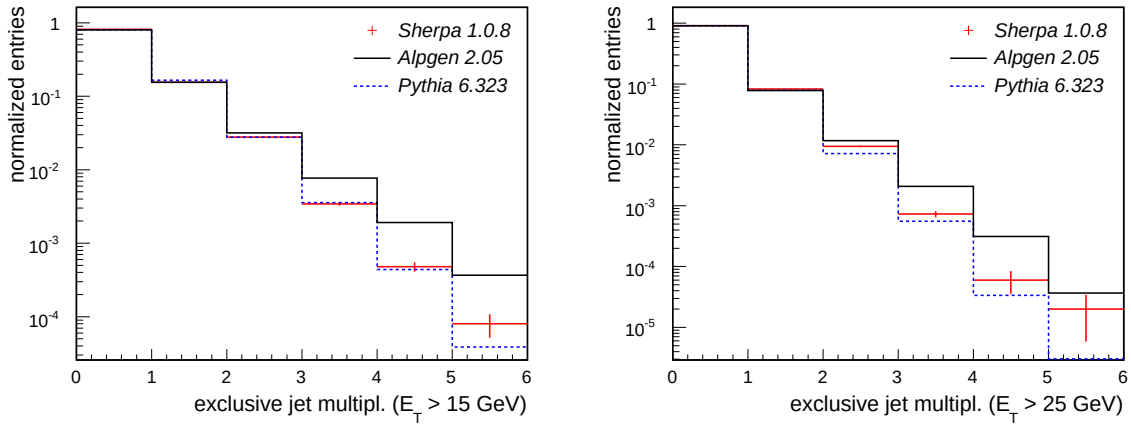

Figure 2.10: Exclusive jet multiplicity in Z boson production in $p\bar{p}$ scattering at $\sqrt{s} = 1.96\,\text{TeV}$ as predicted by Sherpa, Alpgen, *and* Pythia. *The jets are required to be within pseudo-rapidities of $|\eta| < 2.5$ and to have a minimal transverse energy of $E_T > 15\,\text{GeV}$ (left) or $E_T > 25\,\text{GeV}$ (right), respectively.*

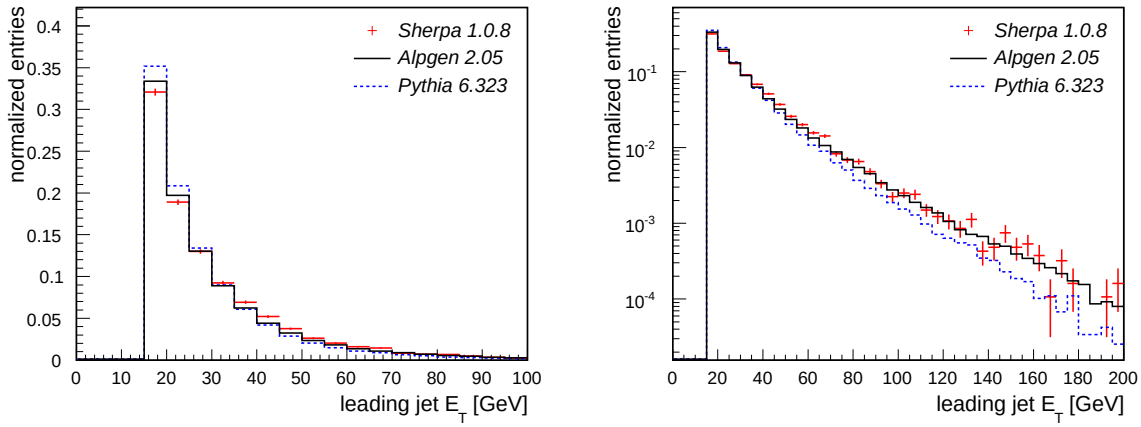

Figure 2.11: Leading jet E_T in $Z+ \geq 1$ jet production in $p\bar{p}$ scattering at $\sqrt{s} = 1.96\,\text{TeV}$ as predicted by Sherpa, Alpgen, *and* Pythia *(left: linear scale, right: logarithmic scale).*

is well described, but for larger jet multiplicities the rate is underestimated when comparing to simulations including matrix-elements for higher parton multiplicities. This trend can be observed when comparing the rates predicted by Alpgen and Pythia. It is interesting to notice that the jet multiplicity as generated by Sherpa falls significantly below the Alpgen prediction. As both generators include matrix-elements up to the same parton multiplicity, this observation is in contrast to the naive expectation, but it can be understood with the predicted jet transverse energy distributions.

Figs. 2.11 and 2.12 depict the E_T distribution of the leading jet in $Z + 1$ jet events and the second jet in $Z + 2$ jet events, respectively. One can notice that the Alpgen

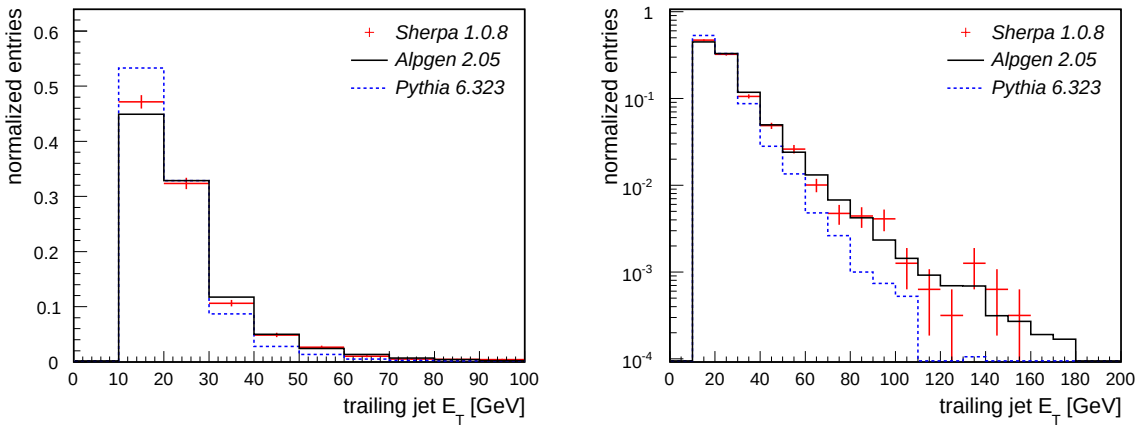

Figure 2.12: Trailing (second hardest) jet E_T in $Z+ \geq 2$ jet production in $p\bar{p}$ scattering at $\sqrt{s} = 1.96\,\text{TeV}$ as predicted by Sherpa, Alpgen, *and* Pythia *(left: linear scale, right: logarithmic scale).*

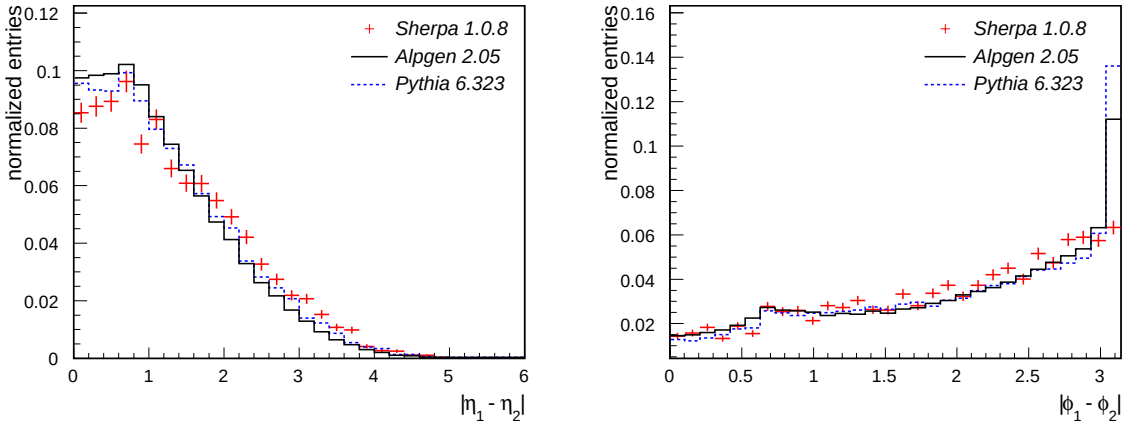

Figure 2.13: Separation of leading and trailing jets in $Z+ \geq 2$ jet production in $p\bar{p}$ scattering at $\sqrt{s} = 1.96\,\text{TeV}$ as predicted by Sherpa, Alpgen, *and* Pythia. *Shown is the difference of the jets' pseudorapidties η_i (left) and polar angle ϕ_i (right).*

and Sherpa distributions agree at large E_T, i.e. the kinematic region dominated by the hard matrix-element, but that at low E_T Alpgen predicts a higher rate. In this rather soft region, the prediction significantly depends on the details of the parton shower and the assumptions of the underlying event model. Therefore, against the common believe, the measurement of jet multiplicities alone is not a meaningful test of hard parton radiation at large E_T. The figure also demonstrates that the jet E_T spectrum as predicted by Pythia is significantly softer, an effect which is even more pronounced for sub-leading jets.

It is also interesting to study the correlations between jets in events with at least two

jets. The most basic quantities are the jet-jet separation in pseudo-rapidity, $|\eta_1 - \eta_2|$, and in polar angle, $|\phi_1 - \phi_2|$, both shown in Fig. 2.13. The most striking difference is the enhancement of the rate at $|\phi_1 - \phi_2| \approx \pi$ seen in PYTHIA and ALPGEN. This can be attributed to the underlying event model implemented in PYTHIA and also used in the ALPGEN simulation, which can generate a pair of additional low-p_T jets balanced in azimuth.

2.4 Conclusions for new phenomena searches

The production of W and Z bosons, as well as Drell-Yan lepton-pairs is a major background in searches for new phenomena and the Higgs boson. In particular, vector boson production with associated jets generates final state topologies which resemble those of many signal processes, which will be discussed in detail in the last chapter.

The phenomenology of the Drell-Yan process is in principal well understood. QCD corrections for the inclusive reaction are available up to NNLO. The higher order corrections lead to the production of jets in association with the vector bosons, reaching large jet multiplicities n_{jet} as the radiation of additional partons is only suppressed with $\mathcal{O}(\alpha_s)$. Due to the complexity of the final state, cross sections for $n_{\text{jet}} \geq 3$ can only be calculated at LO resulting in predictions with substantial uncertainties.

The traditional parton shower event generators, as PYTHIA and HERWIG, which are long established and tuned to experimental data, are well suited to simulate the inclusive production of vector bosons, but fail to accurately describe the multiplicity and transverse momentum spectrum of associated jets. Event generators which consistently combine matrix-element calculations up to large parton multiplicities with parton shower radiation, as SHERPA and ALPGEN, aim to correctly describe the associated production of vector bosons with jets up to high multiplicities. Although the predictions of these Monte Carlo programs differ in some important details and they still need to be confronted and verified in-depth with experimental data, one can conclude that they should be the primary choice for the simulation of W/Z + jet production, which is an essential ingredient in searches for a vast collection of new phenomena.

Chapter 3

Measurement of inclusive W, Z boson and Drell-Yan production at hadron colliders

Evidence for the production of the weak bosons W and Z was first reported in 1982/83 by the UA1 and UA2 collaborations at the CERN $p\bar{p}$-collider S$p\bar{p}$S. Nowadays at the Tevatron and soon at the LHC W and Z boson production are *standard candles* which are used both for precision measurements and calibration purposes. As the production cross sections can be calculated at next-to-next-to-leading order (NNLO) in the strong coupling, the precision of the theoretical prediction is high, thus allowing an alternative measurement of luminosities which is not based on the total inelastic cross section in $p\bar{p}$ or pp collisions. An additional motivation for precise measurements (and calculations) of W and Z boson production is given by their role as an important background for the measurement of and search for standard model and non-SM processes.

3.1 Predictions on W and Z production cross sections

As discussed in Section 2.2 the inclusive cross section of the Drell-Yan process, which is the basic mechanism for the production of high-mass lepton pairs and the weak bosons, can be calculated at NNLO in the strong coupling constant [37, 38]. Thus the precision of the theoretical prediction is expected to be very high. The uncertainty on the calculation due to the contribution of higher orders beyond NNLO can be estimated by varying the renormalization and factorization scales in the NNLO calculation. A simultaneous variation of these scales between $0.5M_V < Q < 2M_V$ (with M_V being the mass of the vector boson) results in modest uncertainties of about 0.3% and 0.5-1% on the W and Z boson production cross sections at the Tevatron and LHC, respectively (cf. Tab. 3.3).

A larger uncertainty on the theoretical prediction originates from the limited precision of the parton distribution functions (PDF) of the proton (and anti-proton). Thus the main uncertainty on the cross-section predictions is in fact of experimental nature as the PDFs are constrained through measurements of deep-inelastic electron, muon

or neutrino-proton scattering, jet production in proton-antiproton collisions, etc.

Since several years PDF sets[1] which parameterize the uncertainty on the parton distributions have been made available. Whereas the first published error on PDF sets where only based on deep-inelastic scattering data [97, 98], the recent global PDF fits by the CTEQ and MRST groups also provide an error evaluation [94, 99]. The CTEQ and MRST uncertainty estimates sets are based on the *Hessian* method [100]: A global χ^2 function is used not only to extract the best fit but also to parameterize the region close to the global minimum. The error matrix is diagonalized (*Hessian* matrix) and eigenvector PDF sets are defined. To account for non-symmetric dependencies, for each eigenvector a separate PDF error set is defined for each direction.

The PDF uncertainty on a cross section σ is then given by

$$\Delta\sigma = \frac{1}{2}\left[\sum_{k=1}^{n}\left[\sigma(a_i^+) - \sigma(a_i^-)\right]^2\right]^{\frac{1}{2}} , \tag{3.1}$$

where the sum is over all n eigenvectors and a_i^+, a_i^- denote the corresponding PDF error sets. The PDF uncertainty of any observable X can be evaluated in an analogue way. As both the global χ^2 function does in general not depend quadratically on the fit parameters and the observable can have a non-linear dependence, one also defines asymmetric uncertainties on the cross section as follows:

$$\begin{aligned}
\Delta^+\sigma &= \left[\sum_{k=1}^{n}\left[\max(\sigma(a_i^+) - \sigma(a_0), \sigma(a_i^-) - \sigma(a_0), 0)\right]^2\right]^{\frac{1}{2}} \\
\Delta^-\sigma &= \left[\sum_{k=1}^{n}\left[\min(\sigma(a_i^+) - \sigma(a_0), \sigma(a_i^-) - \sigma(a_0), 0)\right]^2\right]^{\frac{1}{2}} ,
\end{aligned} \tag{3.2}$$

where a_0 denotes the central PDF fit[2].

In the following, Z and W boson production cross sections and their PDF uncertainty based on the PDF error sets of CTEQ6.1M [101], CTEQ6.5M [102], and MRST2001 [99] are reported. Whereas MRST uses 15 eigenvectors, CTEQ defines 20 eigenvectors, leading to 30 or 40 PDF error sets, respectively.

Table 3.1 lists cross sections for Z and W boson production cross section for both $p\bar{p}$-scattering at Tevatron ($\sqrt{s} = 1.96\,\text{TeV}$) and pp-scattering at LHC ($\sqrt{s} = 14\,\text{TeV}$) calculated with various PDF sets and up to NNLO in the strong coupling constant using the code of Refs. [37, 38] augmented with the latest PDF sets. The numbers are given as the product of the cross section times the leptonic branching ratio (w.r.t. one flavour), $\sigma(Z) \cdot Br(Z \to ll)$ and $\sigma(W^\pm) \cdot Br(W^\pm \to \ell^\pm\nu)$, respectively. The contributions from W^+ and W^- production are added and $Br(Z \to ll) = 0.03366$ and $Br(W^\pm \to \ell^\pm\nu) = 0.108$ [103] are used. The NLO correction is about $\sigma_{\text{NLO}}/\sigma_{\text{LO}} \approx 1.33$ at the Tevatron and ≈ 1.15 at the LHC. The additional NNLO correction is $\sigma_{\text{NNLO}}/\sigma_{\text{NLO}} \approx 1.02$ (Tevatron) and ≈ 0.99 (LHC), respectively. Note that in Table 3.1

[1] A PDF *set* refers to a collection of PDFs for the various partons obtained in a combined fit.

[2] The somewhat complicated form of this definition is motivated by the observation that in some cases $\sigma(a_i^+)$ and $\sigma(a_i^-)$ deviate from the central value in the same direction.

				$\sigma(Z) \cdot Br(Z \to ll)$ / $\sigma(W^{\pm}) \cdot Br(W^{\pm} \to \ell^{\pm}\nu)$			
				TeV (1.96 TeV)		LHC (14 TeV)	
PDF set	Ref.	PDF order	ME order	Z [pb]	W [nb]	Z [nb]	W [nb]
CTEQ5L	[104]	LO	LO	182.7	1.979	1.603	17.65
CTEQ6L1	[94]	LO	LO	177.5	1.933	1.601	17.58
MRST2001	[105]	LO	LO	176.6	1.925	1.548	16.82
CTEQ5M	[104]	NLO	NLO	247.3	2.689	2.008	21.93
CTEQ6.1M	[101]	NLO	NLO	235.9	2.560	1.846	20.16
CTEQ6.5M	[102]	NLO	NLO	244.2	2.655	1.993	21.72
MRST2001C	[99]	NLO	NLO	240.3	2.610	1.887	20.35
MRST2004	[106]	NLO	NLO	240.3	2.613	1.894	20.46
MRST2004	[106]	NNLO	NLO	245.5	2.661	1.882	20.25
CTEQ5M	[104]	NLO	NNLO	253.6	2.748	1.999	21.76
CTEQ6.1M	[101]	NLO	NNLO	241.6	2.612	1.834	19.96
CTEQ6.5M	[102]	NLO	NNLO	250.6	2.715	1.984	21.55
MRST2001C	[99]	NLO	NNLO	246.4	2.665	1.878	20.18
MRST2004	[106]	NLO	NNLO	246.4	2.670	1.886	20.30
MRST2004	[106]	NNLO	NNLO	251.9	2.721	1.880	20.16

Table 3.1: Z and W boson production cross sections at Tevatron ($p\bar{p}$ scattering at $\sqrt{s} = 1.96\,\text{TeV}$) and LHC ($pp$ scattering at $\sqrt{s} = 14\,\text{TeV}$) for various PDF parameterizations calculated up to NNLO (based on the computer code of Refs. [37, 38]).

the only full NNLO calculations, i.e. calculations which use both the NNLO matrix-elements and an NNLO PDF set, are based on the MRST2004 (NNLO) PDF set and are thus considered to be the 'best' predictions.

The uncertainty on the calculation of the Z boson production cross section at the Tevatron is shown in Table 3.2. The scale and PDF uncertainties are evaluated at LO, NLO, and NNLO to demonstrate the following: At LO the underlying hard process is independent of the strong coupling. Thus the small change in the evaluated cross section is only due to the variation of the factorization scale. While the scale uncertainty is $> 1\%$ at NLO, it is considerably reduced to 0.3% at NNLO. The PDF uncertainty has been evaluated with both the assumptions of symmetric and asymmetric errors according to Equations (3.1) and (3.2). It is significantly larger than the scale uncertainty and it is mildly increasing at higher orders due to the increased contribution of diagrams with gluons in the initial state, the distribution functions of which are less well constrained. Notably, the allowed ranges calculated with the PDF error sets of CTEQ6.1M [101], CTEQ6.5M [102], and MRST2001 [99] differ considerably. The moderate reduction of the uncertainty from 3.7% to 2.9% with the CTEQ update can be understood as the theoretical and experimental input for CTEQ6.5M were improved and the fit parameters were more thoroughly studied. The uncertainty range evaluated using the MRST PDF sets is about a factor 2.5 smaller. Part of this difference can be attributed to the different ways how the CTEQ and MRST groups define their confidence region. They both aim to approximate a 90% C.L. interval using their error sets, but as this range cannot be rigorously defined, they do not agree on its defini-

$\sigma(Z)$	scale uncert.		PDF uncertainty							
TeV			CTEQ6.1M			CTEQ6.5M			MRST2001	
	[%]		symmetric / asymmetric + / − [%]							
LO:	+0.0	−0.6	±3.5	+3.5	−3.8	±2.7	+2.7	−2.8	±1.2	+0.9	−1.5
NLO:	+1.7	−1.1	±3.6	+3.6	−4.0	±2.9	+2.9	−2.9	±1.2	+1.0	−1.6
NNLO:	+0.3	−0.3	±3.7	+3.7	−4.1	±2.9	+3.0	−2.9	±1.2	+1.0	−1.6

Table 3.2: PDF and scale uncertainties on the Z boson production cross sections at the Tevatron. The PDF uncertainties are evaluated using the PDF error sets of CTEQ6.1M [101], CTEQ6.5M [102], and MRST2001 [99], respectively. The scale uncertainty is evaluated by varying simultaneously the factorization and renormalization scales between $0.5M_V < Q < 2M_V$.

	scale uncert.	PDF uncertainty					
		CTEQ6.1M		CTEQ6.5M		MRST2001	
	[%]	symmetric / asymmetric [%]					
TeV: $\sigma(Z)$	$^{+0.3}_{-0.3}$	±3.7	$^{+3.7}_{-4.1}$	±2.9	$^{+3.0}_{-2.9}$	±1.2	$^{+1.0}_{-1.6}$
TeV: $\sigma(W)$	$^{+0.2}_{-0.2}$	±3.8	$^{+3.7}_{-4.3}$	±3.1	$^{+3.2}_{-3.1}$	±1.2	$^{+1.0}_{-1.6}$
LHC: $\sigma(Z)$	$^{+1.1}_{-0.5}$	±4.8	$^{+4.8}_{-5.6}$	±3.8	$^{+4.4}_{-3.9}$	±1.7	$^{+1.4}_{-2.0}$
LHC: $\sigma(W)$	$^{+0.8}_{-0.3}$	±5.0	$^{+4.9}_{-5.7}$	±4.0	$^{+4.6}_{-4.0}$	±1.8	$^{+1.5}_{-2.3}$

Table 3.3: PDF and scale uncertainties on the Z and W boson production cross sections at NNLO at Tevatron (TeV) and LHC. See Tab. 3.2 and text for details.

tion. Whereas CTEQ estimates the region by allowing in their χ^2-test a difference of $\Delta\chi^2 = 100$ w.r.t. their best fit value, MRST uses a value of $\Delta\chi^2 = 50$. Therefore, one would expect that the uncertainty range given by CTEQ is roughly a factor $\sqrt{2} \sim 1.4$ larger, which can only explain a fraction of the observed difference. Given that the central values of the Z and W boson production cross sections calculated at the same order using various recent PDF sets scatter by about $\sim 3\%$ (cf. Table 3.1), a more conservative estimate of the uncertainty range appears to be appropriate. To give an example, we note that the cross sections calculated with CTEQ6.1M are outside the allowed uncertainty range of CTEQ6.5M.

Table 3.3 lists the scale and PDF uncertainties (calculated at NNLO) for both Z and W boson production cross sections at both the Tevatron and the LHC. In general the scale error is comparably small and the PDF uncertainty is about 3% at the Tevatron and 4% at the LHC.

A large fraction of the uncertainties in the cross section predictions cancel in the ratio

$$\mathcal{R} = \frac{\sigma_{W^\pm} \cdot Br(W^\pm \to \ell^\pm \nu)}{\sigma_Z \cdot Br(Z \to ll)} \,, \tag{3.3}$$

which is of particular relevance for the indirect measurement of the total W boson

	CTEQ5M	-6.1M	-6.5M	MRST2001C	-2004NLO	-2004NNLO
TeV:	10.83	10.81	10.83	10.82	10.83	10.80
LHC:	10.89	10.88	10.86	10.74	10.76	10.73

Table 3.4: Predictions for the ratio $\mathcal{R} = \sigma(W^{\pm}) \cdot Br(W^{\pm} \to \ell^{\pm}\nu) / (\sigma(Z) \cdot Br(Z \to ll))$ at Tevatron and LHC calculated at NNLO.

	scale uncert.	PDF uncertainty					
		CTEQ6.1M		CTEQ6.5M		MRST2001	
	[%]	symmetric / asymmetric [%]					
TeV: $\sigma(W)/\sigma(Z)$	$^{+0.10}_{-0.05}$	± 0.52	$^{+0.44}_{-0.67}$	± 0.51	$^{+0.54}_{-0.55}$	± 0.39	$^{+0.42}_{-0.44}$
LHC: $\sigma(W)/\sigma(Z)$	$^{+0.18}_{-0.29}$	± 0.38	$^{+0.43}_{-0.37}$	± 0.46	$^{+0.46}_{-0.46}$	± 0.32	$^{+0.31}_{-0.34}$

Table 3.5: PDF and scale uncertainties on production cross section ratio $\sigma(W)/\sigma(Z)$ at NNLO at Tevatron (TeV) and LHC. See Tab. 3.2 and text for details.

decay width Γ_W. It can be calculated from a measurement of $\mathcal{R}$, as

$$\mathcal{R} = \frac{\sigma_W}{\sigma_Z} \frac{\Gamma_Z}{\Gamma_{Z\to ll}} \frac{\Gamma_{W\to l\nu}}{\Gamma_W} \,. \tag{3.4}$$

Both the ratio σ_W/σ_Z and the leptonic W decay width $\Gamma_{W\to l\nu}$ can be calculated theoretically to high precision. The leptonic branching fraction of the Z boson $\Gamma_{Z\to ll}/\Gamma_Z$ has been measured very accurately at LEP. Tables 3.4 and 3.5 show the central values for the cross section ratio calculated with various PDF sets and its uncertainties, respectively.

3.2 Measurement of Z and W boson production cross sections at hadron colliders

Following the discovery of the weak vector bosons at the S$p\bar{p}$S collider, their production cross sections in proton-antiproton scattering at $\sqrt{s} = 546\,\text{GeV}$ and $\sqrt{s} = 630\,\text{GeV}$ were measured in both electron and muon decay channels by the UA1 and UA2 collaborations [107–109]. The most precise measurement was achieved by the UA2 collaboration in the e channel based on a data set of an integrated luminosity of $13\,\text{pb}^{-1}$. A statistical error of 2% and 6% was achieved for the W and Z production cross section, respectively. A systematic error of 6% was estimated which was largely dominated by the uncertainty on the luminosity.

During Run I at the Tevatron σ_W and σ_Z have been measured in $p\bar{p}$-collisions at $\sqrt{s} = 1.8\,\text{TeV}$ by the CDF [110–113] and DØ [114–116] collaborations in both the electron and muon decay channels. DØ also published a measurement of the W cross section in the τ decay channel [117]. The most precise measurement was obtained by DØ in the e channel [116]. Based on an integrated luminosity of $84.5\,\text{pb}^{-1}$ statistical errors as low as 0.4% (W) and 1.4% (Z) could be obtained with a systematic error of 5% which mainly reflects the uncertainty on the luminosity determination.

Measurements of the weak bosons' production cross sections at Tevatron's increased collision energy of $\sqrt{s} = 1.96\,\text{TeV}$ in Run II have been published by the CDF collaboration based on an initial dataset with an integrated luminosity of just $72\,\text{pb}^{-1}$ [118,119]. Detailed descriptions of the CDF and DØ experiments at the Tevatron Run II are given in Refs. [120, 121]. For the CDF measurement both the electron and muon decay channels have been used. The data are collected using common single-electron or single-muon triggers, respectively, for both W and Z selections.

W candidates are selected by a reconstructed high p_T electron or muon and large missing transverse energy, $\not\!\!E_{\text{T}}$, originating from the undetected neutrino. Electrons are identified with electromagnetic showers in the calorimeter which are matched to reconstructed tracks. Muons are identified based on track segments in the muon chambers which are matched with a track reconstructed in the central detector. For Z candidates looser lepton identification requirements are imposed on the second electron or muon to increase the total event reconstruction efficiencies.

The kinematic and geometric acceptance for the event selection was evaluated using the PYTHIA event generator [122] and a full simulation of the CDF detector based on the GEANT simulation program [123]. As the acceptance critically depends on the pseudorapidities of the decay leptons and therefore on the boson rapidity y, the acceptance was determined as function of y and applied on the NNLO calculation of the differential production cross section $\text{d}\sigma/\text{d}y$ [39]. This theoretical prediction depends on the input parton distribution functions and the corresponding uncertainty is evaluated similar to the prescription given above using the CTEQ6M error PDF sets. The resulting uncertainty on the acceptance ranges from 0.7%-2.1% depending on the channel. It is interesting to notice that this uncertainty constitutes the largest systematic error (besides the 6% uncertainty of the luminosity determination) in the cross section measurements. For the electron channels, the limited accuracy in the material description implemented in the detector simulation results in the second largest contribution to the uncertainty on the acceptance.

Efficiencies for lepton reconstruction, identification, and triggers are measured directly using $Z \to \ell^+\ell^-$ decays. A *tag-and-probe* method is employed in which standard identification cuts are applied to one lepton and efficiencies are measured using the other lepton candidate. The limited number of Z candidates results in statistical uncertainties on the measured efficiencies which range up to about 1%. Obviously, the accuracy should significantly improve with the increasing size of data samples.

The background consists mostly of QCD multijet production and weak boson production. For the latter, a W or Z decay of a certain type appears in another selection channel. As an example, the largest background in the $W \to \mu\nu$ channel is due to $Z \to \mu^+\mu^-$ production. This is a consequence of the limited muon acceptance which leads to 'fake' $\not\!\!E_{\text{T}}$ due to the non-detected muon. For the $Z \to \mu^+\mu^-$ channel, which has the smallest level of background contamination, cosmic-rays traversing the detector comprise the most important background contribution.

The measurement of $\sigma_W \cdot Br(W \to \ell\nu)$ in both electron and muon channels is used to confirm the e-μ universality with a precision of 1.2%. The measurements are combined taking correlated uncertainties into account to obtain the product of cross

section times branching ratio:

$$\begin{aligned} \sigma_W \cdot Br(W \to \ell\nu) &= 2775 \pm 10(\text{stat}) \pm 53(\text{syst}) \pm 167(\text{lum})\,\text{pb} \\ \sigma_{Z/\gamma^*} \cdot Br(Z/\gamma^* \to \ell^+\ell^-) &= 254.9 \pm 3.3(\text{stat}) \pm 4.6(\text{syst}) \pm 15.2(\text{lum})\,\text{pb} \ . \end{aligned} \tag{3.5}$$

The latter cross section is given for dilepton production in the mass range $66\,\text{GeV} < M_{\ell\ell} < 116\,\text{GeV}$ including contributions from both γ^* and Z boson exchange. The cross section for pure Z exchange over the entire mass range is 0.4% higher than the quoted measurement. Both measurements are in good agreement with the SM predictions reported in Section 3.1. A comparison of the predictions as a function of the centre of mass energy with this CDF result and previous measurements of $\sigma_W \cdot Br(W \to \ell\nu)$ and $\sigma_Z \cdot Br(Z \to \ell^+\ell^-)$ at the CERN S$p\bar{p}$S and Tevatron Run I is shown in Fig. 3.1.

When not including the luminosity uncertainty, the cross section measurements have a precision of 2.0% and 2.2%, respectively. Combining these numbers with the uncertainty on the theoretical cross section prediction of about 3%, one can conclude that these measurements could provide an alternative determination of the luminosity with a higher precision than the traditional method based on the total inelastic $p\bar{p}$ cross section.

As the uncertainty on the integrated luminosity cancels in the cross section ratio $\mathcal{R}$ (see Equations (3.3) and (3.4)) the total W decay width Γ_W can be derived with high accuracy. The combined result for $\mathcal{R}$ from both the electron and muon channels is

$$\mathcal{R} = 10.92 \pm 0.15(\text{stat}) \pm 0.14(\text{syst}) \ , \tag{3.6}$$

which can be used to extract the leptonic branching ratio $Br(W \to \ell\nu)$ and the total width of the W boson Γ_W:

$$\begin{aligned} Br(W \to \ell\nu) &= 0.1089 \pm 0.0022 \\ \Gamma_W &= 2079 \pm 41\,\text{GeV} \ . \end{aligned} \tag{3.7}$$

The extracted value for Γ_W is of the same precision as the current world limit (which does not yet include this measurement) of $\Gamma_W = 2141 \pm 41\,\text{GeV}$ [103].

At this time the DØ collaboration has only reported preliminary results on cross-section measurements of single W or Z production decaying into electron or muon final states [125–127]. However, the collaboration has published a cross section measurement for Z production decaying via $Z \to \tau^+\tau^-$ [128]. As the τ lepton can decay into various leptonic and hadronic final states, its identification is challenging and thus a precision in the cross section measurement which approaches the accuracy of the electron or muon channels cannot be achieved. Therefore, this measurement aims to test the SM since a deviation from the expected cross section value would be an indication of an anomalous production of $\tau^+\tau^-$ pairs in $p\bar{p}$ collisions. Furthermore, it verifies the capability to identify isolated τ leptons, which could be critical in the search for non-SM signals such as supersymmetric particles or heavy gauge bosons with enhanced couplings to the third lepton generation. The measurement yields

$$\sigma_Z \cdot Br(Z \to \tau^+\tau^-) = 237 \pm 15(\text{stat}) \pm 18(\text{syst}) \pm 15(\text{lum})\,\text{pb} \ , \tag{3.8}$$

which is in good agreement with the SM prediction.

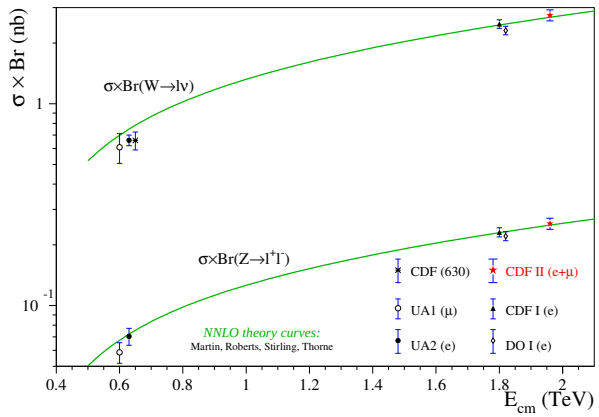

Figure 3.1: $W \to \ell\nu$ and $Z \to \ell\ell$ cross section measurements as a function of the $p\bar{p}$ centre-of-mass energy, E_{CM} [108, 109, 112, 116, 118]. The solid lines show the NNLO calculations based on Ref. [37] and the MRST PDF set [124].

3.3 Drell-Yan lepton pair production

As discussed in the previous chapter, Drell-Yan dilepton production probes the exchange of virtual photons besides the Z boson as well as their interference. Measurements of dilepton production cross sections help to constrain the parton distribution functions. Drell-Yan lepton pair production can also be used to test the electroweak sector of the standard model, e.g. the measurement of the forward-backward production asymmetry constrains the vector and axial-vector couplings of the Z boson to the quarks [129].

Dilepton production outside the Z resonance region is also an important background to many searches, e.g. lepton pairs at low invariant mass dilute the trilepton signal originating from the chargino-neutralino pair-production in supersymmetric models, which yield leptons with relatively low transverse momenta [130]. In Drell-Yan events a third reconstructed lepton could be found as part of the hadronic activity could be falsely identified as a lepton. The production of dilepton pairs at high invariant masses is predicted by numerous extensions of the SM. Searches for a deviation from the SM Drell-Yan background have been performed within various models, e.g. searches for additional gauge bosons, for quark-lepton compositness, and extra dimensions [131–134].

Whereas additional high mass gauge bosons could be detected as resonances in the dilepton spectrum, quark-lepton compositness and large extra dimensions could change the dilepton spectrum over the full mass range above a certain invariant mass threshold. Especially for the latter the exact knowledge of the SM cross section and of the uncertainty of its prediction is essential. As already discussed before, the Drell-Yan dilepton production has been calculated to NNLO in the strong coupling constant. Thus the precision of the cross section predictions is given by the PDF uncertainty as in the case of W and Z boson production.

Fig. 3.2 shows the K-factor and the PDF uncertainty as function of the dilepton mass $M_{\ell\ell}$ for $p\bar{p}$ scattering at $\sqrt{s} = 1.96\,\mathrm{TeV}$ (Tevatron Run II). The K-factor has been

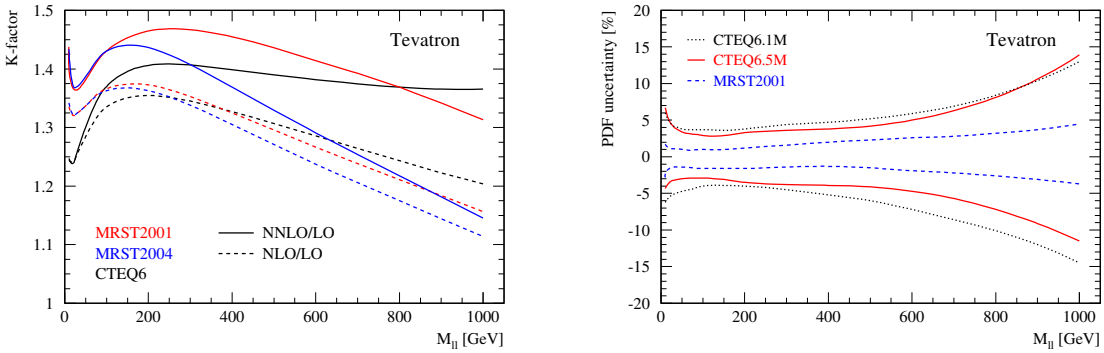

Figure 3.2: NLO and NNLO K-factors (left) and PDF uncertainty (right) for Drell-Yan lepton-pair production in $p\bar{p}$ scattering at $\sqrt{s} = 1.96\,TeV$ (Tevatron) as function of the dilepton mass $M_{\ell\ell}$. See text for further details.

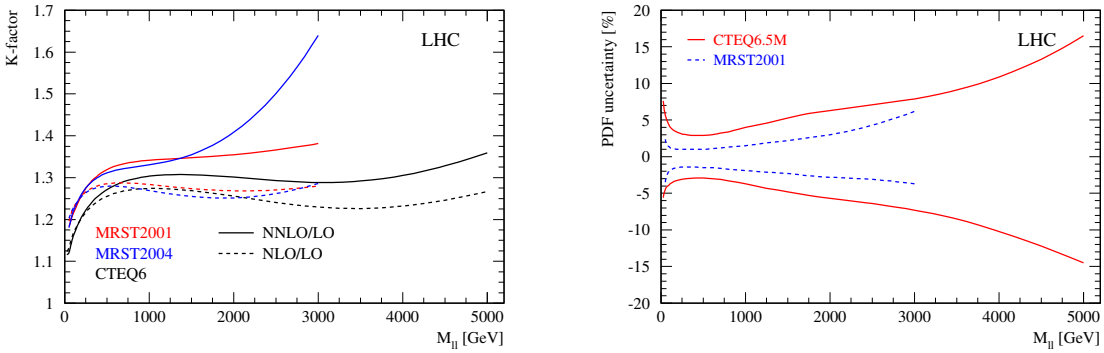

Figure 3.3: NLO and NNLO K-factors (left) and PDF uncertainty (right) for Drell-Yan lepton-pair production in pp scattering at $\sqrt{s} = 14\,TeV$ (LHC) as function of the dilepton mass $M_{\ell\ell}$. See text for further details.

evaluated at NLO and NNLO based on the program code of Refs. [37,38] and using the PDF sets of CTEQ6 [94], MRST2001 [105] and MRST2004 [106]. The K-factors are defined as $K_{NLO} = \sigma_{NLO}/\sigma_{LO}$ and $K_{NNLO} = \sigma_{NNLO}/\sigma_{LO}$, respectively. The order of the PDF fit has been matched to the order of the matrix-element, except for the NNLO calculation based on CTEQ6, which does not include NNLO distribution functions, and thus the NLO PDFs are used. For the K-factors labelled 'MRST2004', the LO cross section was evaluated using the LO fit of MRST2001 (as a LO fit is not included in the 2004 version). Therefore, it should be noted that this factor does not only reflect the higher order corrections, but also any changes between the 2001 and 2004 versions of the PDF fit.

Notably, the NNLO correction is positive over the entire mass range. Both K_{NNLO} and K_{NLO} decrease with increasing invariant mass for $M_{\ell\ell} \gtrsim 200\,\text{GeV}$, but whereas

K_{NNLO} appears to be relatively stable for CTEQ6, for MRST2001 K_{NNLO} decreases by 10% between $M_{\ell\ell} = 250\,\text{GeV}$ and $1000\,\text{GeV}$.

The PDF uncertainty of the cross section was evaluated using the PDF error sets of CTEQ6.1M [101], CTEQ6.5M [102], and MRST2001 [99] as described before for Z and W production. The relative uncertainty (calculated with CTEQ6.5M) is below 5% for $M_{\ell\ell} \lesssim 600\,\text{GeV}$, but rapidly grows for higher invariant masses.

Similarly, the K-factor and PDF uncertainty as function of $M_{\ell\ell}$ for pp scattering at $\sqrt{s} = 14\,\text{TeV}$ (LHC) is shown in Fig. 3.3. The predictions based on the MRST PDF sets are only shown up to $M_{\ell\ell} = 3\,\text{TeV}$ to ensure that the parton density is not evaluated within a non-defined kinematic region. Both K_{NLO} and K_{NNLO} are nearly constant for large $M_{\ell\ell} \gtrsim 500\,\text{GeV}$, except for MRST2004, where the steep rise is probably an artifact of the NNLO PDF fit in an extreme kinematic region. The PDF uncertainty is steadily increasing above $M_{\ell\ell} \gtrsim 500\,\text{GeV}$.

In the region of high invariant dilepton mass the Drell-Yan production cross section is most accurately measured in the $Z/\gamma^* \to e^+e^-$ decay channel, where one can benefit from the precise energy measurement in the electromagnetic calorimeter with its energy resolution improving approximately as $\Delta E/E \sim 1/\sqrt{E}$. In case of the muon decay channel, the energy/momentum is measured using tracking detectors, for which the resolution degrades rapidly at large transverse momenta, thus resulting in a relatively poor measurement of the invariant dilepton mass at high $M_{\ell\ell}$. The difference in resolution outweighs the advantage of the muon decay channel of having a considerably smaller background due to QCD multijet production. Therefore the majority of the measurements obtained at the Tevatron are based on the electron decay channel.

At the Tevatron Run I the differential cross section $\mathrm{d}\sigma/\mathrm{d}M_{\ell\ell}$ for high-mass Drell-Yan lepton pair production was measured in the electron decay channel by both the CDF [135] and DØ [136] collaborations using an integrated luminosity of $\sim 110 - 120\,\text{pb}^{-1}$. In addition, CDF extracted $\mathrm{d}\sigma/\mathrm{d}M_{\ell\ell}$ from a measurement of $\mathrm{d}^2\sigma/(\mathrm{d}M_{\ell\ell}\mathrm{d}y)$ in the dimuon channel for the central rapidity region $|y| < 1$ [113, 135]. Due to the steeply falling cross section with increasing $M_{\ell\ell}$ these measurements are statistically limited, e.g. the CDF analysis observed 14 dielectron events at $M_{\ell\ell} > 200\,\text{GeV}$ (with 0.3 estimated background events). Within the relatively large uncertainties of the cross section measurements, all results are in agreement with the NNLO prediction. High-mass dilepton production was also studied in various searches for physics beyond the standard model [137–139]. No indication of a deviation from the SM was observed.

In Run II the CDF collaboration analysed Drell-Yan e^+e^- production with a small dataset of an integrated luminosity of $72\,\text{pb}^{-1}$ to measure the forward-backward asymmetry A_{FB} and to derive the Z-quark couplings therewith [129]. Unfortunately this analysis was not used to extract the Drell-Yan cross section $\mathrm{d}\sigma/\mathrm{d}M_{ee}$. At this time the DØ collaboration only reported preliminary results on A_{FB} based on $180\,\text{pb}^{-1}$, but this analysis also derived $\mathrm{d}\sigma/\mathrm{d}M_{ee}$ in the mass range $70\,\text{GeV} < M_{ee} < 400\,\text{GeV}$ [140]. This measurement is shown in Fig. 3.4 together with the NNLO prediction, which is in good agreement with the data. Compared to the DØ result obtained in Run I [136] the background contribution from multijet QCD production could be significantly reduced using a tight track match condition for the electron candidates defined through energy deposits in the electromagnetic calorimeter. Nevertheless the precision of the cross-section measurement could not exceed the CDF Run I result [135].

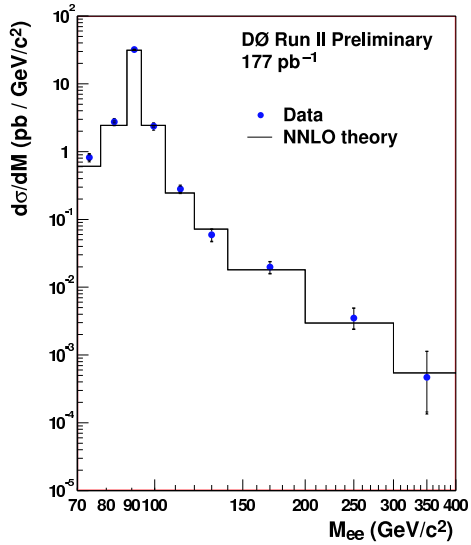

Figure 3.4: Drell-Yan production cross section $\mathrm{d}\sigma/\mathrm{d}M_{ee}$ *in* $p\bar{p}$ *scattering at* $\sqrt{s} = 1.96$ *TeV as function of dielectron mass* M_{ee} *[140]. The histogram shows the NNLO prediction [37, 38].*

Both, the CDF and DØ collaborations studied high-mass lepton-pair production in numerous searches for new phenomena beyond the SM. The CDF collaboration published a search for Z' bosons using the dielectron mass and angular distribution [134] and presented preliminary limits for several BSM models based on the dielectron mass spectrum measured with an integrated luminosity of $1.3\,\mathrm{fb}^{-1}$ [141]. The DØ collaboration searched for Randal-Sundrum gravitons, which are excitations in extra dimension models which decay in fermion-antifermion or diboson pairs, in e^+e^-, $\mu^+\mu^-$, and $\gamma\gamma$ final states [131]. For the e^+e^- and $\gamma\gamma$ channels a preliminary update was reported based on a largely increased integrated luminosity of $1\,\mathrm{fb}^{-1}$ [142]. Furthermore, limits on large extra dimensions were published based on the dimuon invariant mass distribution [132]. All measurements found no indication for a deviation from the SM prediction. Despite the numerous searches based on measurements of the dilepton invariant mass spectrum, a Drell-Yan cross section measurement at the Tevatron Run II has not been published yet.

3.4 Differential measurement of W and Z boson production

In addition to accurate measurements and predictions for the *total* production cross sections of the weak bosons, a precise determination of the *differential* distributions is of great importance as well, both for an in-depth understanding of the production processes as well as for the correct prediction of event rates (e.g. background contributions) after kinematic selections. Some of these measurements are also sensitive to the parton distribution functions and can therefore provide additional constraints to PDF fits.

3.4.1 Forward-backward charge asymmetry in W boson production

The $W^{\pm}$ rapidity[3] distribution, y_W, in $p\bar{p}$ collisions is sensitive to the u and d quark distribution functions in the proton, as the dominating production processes at Tevatron collision energies are $u\bar{d} \to W^+$ and $d\bar{u} \to W^-$ with quarks and antiquarks predominantly originating from their valence distribution in the proton and antiproton, respectively. The momentum fractions $x_{1,2}$ of the struck quark and antiquark can be inferred from the rapidity of the W boson, since

$$x_{1,2} = \frac{M_W}{\sqrt{s}} e^{\pm y_W} \ . \tag{3.9}$$

As u quarks carry on average a higher momentum fraction than d quarks, W^+ bosons are primarily boosted along the proton (forward) direction and W^- bosons along the antiproton (backward) direction, thus resulting into a charge asymmetry in the W boson rapidity distribution. At hadron colliders W bosons can only be identified with high purity in leptonic decays, where y_W cannot be directly determined as the longitudinal momentum of the neutrino is unmeasured. However, the rapidity y_l or pseudorapidity[4] η_ℓ of the charged lepton from the W boson decay, which can be accurately measured, is correlated with y_w. The production and decay of the W boson is proceeding via the $V - A$ coupling of the weak interaction. Therefore, the spin of the W boson is aligned with the direction of the incoming antiquark (thus antiproton) and the charged lepton is preferentially emitted opposite to the boost of the decaying boson. Thus, the lepton charge asymmetry, defined as

$$A(\eta_\ell) = \frac{\mathrm{d}\sigma(\ell^+)/\mathrm{d}\eta_\ell - \mathrm{d}\sigma(\ell^-)/\mathrm{d}\eta_\ell}{\mathrm{d}\sigma(\ell^+)/\mathrm{d}\eta_\ell + \mathrm{d}\sigma(\ell^-)/\mathrm{d}\eta_\ell} \ , \tag{3.10}$$

which can be directly measured, is a convolution of the asymmetries in the W boson production and decay.

At the Tevatron Run I the CDF collaboration measured the lepton charge asymmetry in both $W \to e\nu$ and $W \to \mu\nu$ decays [143]. No result was obtained with the DØ detector as its central solenoid magnet, which is essential for the charge determination, was only added before Run II.

At the increased centre-of-mass energy of Run II, measurements were performed by CDF in the electron channel [144] and by DØ in the muon decay channel [145]. These channels have complementary advantages: The CDF and DØ detectors have a better acceptance for electrons, especially at large rapidities. However, electrons are more prone to bremsstrahlung, leading to an increased charge misidentification.

CDF's electron charge asymmetry measurement is based on an integrated luminosity of $170\,\mathrm{pb}^{-1}$. As the acceptance of the central drift chamber only extends to $\eta \approx 1$, at larger pseudorapidities tracks are reconstructed using the silicon detectors

[3]The rapidity y is defined by $y = \frac{1}{2}\log\frac{E+p_L}{E-p_L}$, where E and p_L denote the particle's energy and longitudinal momentum with respect to the proton direction.

[4]The pseudorapidity η is defined by $\eta = -\log\tan(\theta/2)$, where θ is the angle between the particle and the proton direction. If the particle's mass is much smaller than its energy, the pseudorapidity η approximates the true rapidity y.

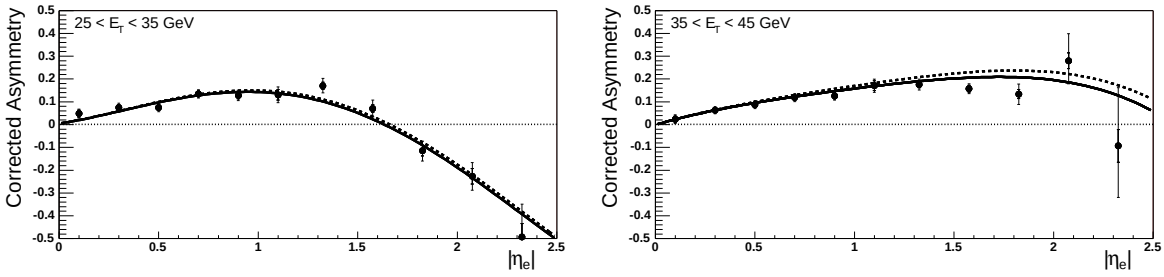

Figure 3.5: The electron asymmetry, $A(|\eta_e|)$, in $W \to e\nu$ production measured by the CDF collaboration in Run II [144]. The data are compared to resummed NLO calculations using the CTEQ6.1M (solid) and MRST02 (dashed) PDFs [68, 99, 101]. The left plot is for electron transverse energies $25\,\mathrm{GeV} < E_T < 35\,\mathrm{GeV}$, and the right one for $35\,\mathrm{GeV} < E_T < 45\,\mathrm{GeV}$.

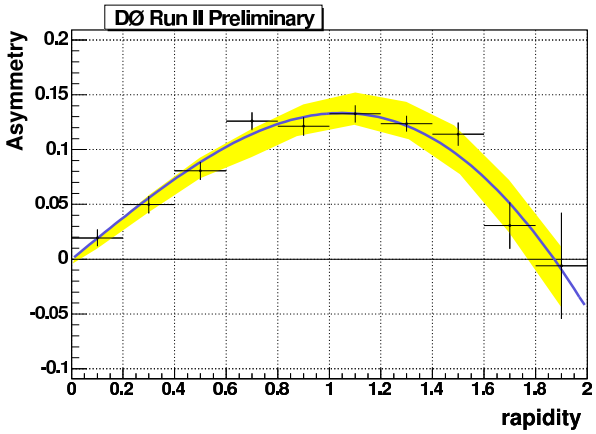

Figure 3.6: The muon asymmetry, $A(|\eta_\mu|)$, in $W \to mu\nu$ production measured by DØ [145]. The data are compared to the resummed NLO predictions using the MRST02 PDFs (blue line) [99, 146]. The yellow band corresponds to the envelop obtained with the CTEQ6.1M error PDF sets [101].

and a calorimeter-seeded algorithm. The electron charge is determined from the curvature of the reconstructed track. Bremsstrahlung and resolution effects can lead to a misidentification of the charge, which dilutes the measured asymmetry. A residual misalignment of the detector might bias the charge identification, resulting in a shift in the observed charge asymmetry. The charge misidentification probability is determined using $Z \to e^+e^-$ events in which one electron track is used as a charge tag. The measured probability, which exceeds $\sim 2\%$ for $|\eta| > 1.5$, is used to correct the raw asymmetry. The precision on the measured charge misidentification probability is statistically limited by the limited number of Z events and as a consequence is the dominating systematic uncertainty in the asymmetry measurement.

The CDF collaboration has measured the electron asymmetry $A(\eta_\ell)$ in two regions of electron transverse momentum, which probe different ranges of W boson rapidity and thus of parton momenta x. The data are shown in Fig. 3.5 together with predictions

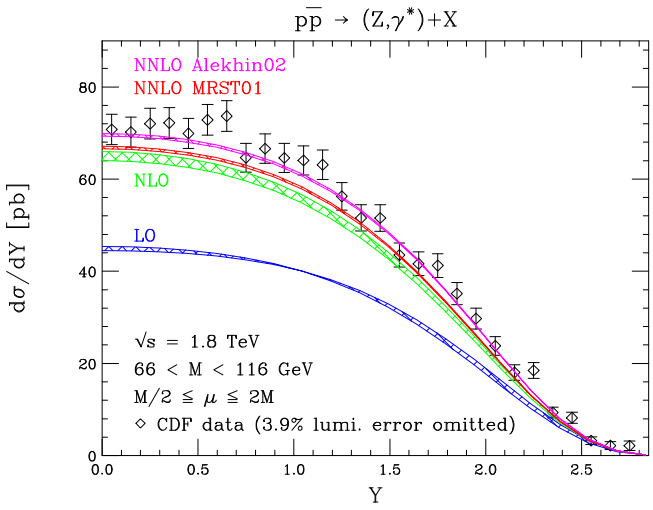

Figure 3.7: The dilepton rapidity distribution for $p\bar{p} \to Z/\gamma^*$ production at $\sqrt{s} = 1.8\,\text{TeV}$ calculated up to NNLO [39], compared with CDF Run I data [147].

from resummed NLO calculations based on the CTEQ6.1M and MRST02 PDFs [68, 99, 101].

The DØ collaboration presented preliminary results for $A(\eta_\ell)$ in the muon channel based on an integrated luminosity of $230\,\text{pb}^{-1}$. The DØ detector provides a wide muon acceptance within $|\eta| \lesssim 2$ and track reconstruction up to very forward $|\eta| \sim 3$ using the silicon microstrip tracker. Due to the vanishing bremsstrahlung of the muons, the charge misidentification probability is only $5 \cdot 10^{-5}$, which has been determined using $Z \to \mu^+\mu^-$ data events. Full simulation verified that the misidentification probability is not significantly increasing at large η_ℓ. Although a huge uncertainty of 100-500% was estimated for the charge misidentification, this contributes only little to the final systematic error, which is dominated by the efficiency ratio of the muon identification for the two charges.

Fig. 3.6 shows the DØ measurement of $A(\eta_\ell)$ together with the resummed NLO prediction based on the MRST02 and CTEQ6.1M PDFs [99, 101, 146]. The precision of the data and the accuracy of the prediction due to the PDF uncertainty is of comparable size, indicating that future, improved measurements based on larger integrated luminosities will provide significant additional constraints on global PDF fits.

3.4.2 Boson rapidity distribution in $p\bar{p} \to Z/\gamma^*$ production

The rapidity distribution of the produced Z bosons can be directly measured using the leptonic decays. As the momentum fractions x of the struck partons can be inferred from the Z boson's rapidity y (cf. Eq. (3.9)), the measurement is in principal sensitive to the quark PDFs of the proton. In particular the high x region could be probed with Z bosons produced at large rapidities. The boson rapidity distribution has recently been calculated up to NNLO in QCD [39]. The prediction for the differential production cross section $\mathrm{d}\sigma/\mathrm{d}y$ for the production of Z bosons in $p\bar{p}$-collisions at $\sqrt{s} = 1.8\,\text{TeV}$

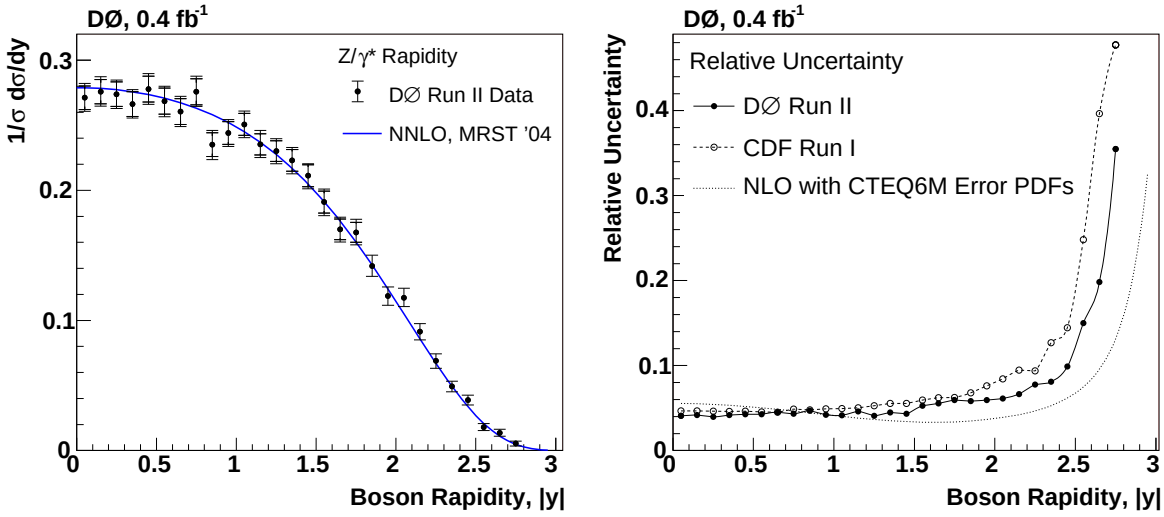

Figure 3.8: Left: the differential production cross section of the Z boson as function of its rapidity y normalized to the integrated cross section [149]. The solid line shows the NNLO prediction [39] based on the MRST2004 NNLO parton distribution functions [106]. Right: relative uncertainties of this measurement compared to the CDF Run I result [147] and the PDF uncertainty on the theoretical prediction obtained with the CTEQ6M error PDF sets [94].

is shown in Fig. 3.7 together with the CDF result obtained using $Z \to e^+e^-$ events collected in Run I [147]. The CDF collaboration presented a preliminary update of this measurement based on 1.1fb^{-1} taken during Run II [148].

The DØ collaboration published a measurement of the rapidity distribution for $p\bar{p} \to Z/\gamma^* \to e^+e^- + X$ events based on a data set of 400 pb^{-1} collected in Run II [149]. The challenge of this measurement was the reconstruction and identification of electrons over a wide range in pseudo-rapidity η up to values of $|\eta| \sim 3.2$ to reach large acceptances for the interesting kinematic region of large $y(Z)$. The electrons were reconstructed as electromagnetic showers in the fiducial regions of the central and endcap calorimeters. At least one electron candidate was required to have a spatial track match. For events with two central electron candidates, both needed to have a track match. The electron reconstruction and identification efficiencies were measured as function of η using the *tag-and-probe* method (see Section 3.2). The statistical precision on these efficiencies are the main contribution to the systematic uncertainty, which is smaller than the statistical error. The result for the normalized differential cross section $1/\sigma \times \mathrm{d}\sigma/\mathrm{d}y$ is depicted in Fig. 3.8. The NNLO prediction [39] based on the MRST2004 NNLO PDFs [106] is in good agreement with the data. Also shown in the figure is a comparison of the total uncertainty of this measurement with the one of the previously best published result obtained by CDF in Run I [147] together with the PDF uncertainty of the prediction. Although the uncertainty on the measurement was significantly improved, it has not yet the sensitivity to provide additional constraints on the proton PDFs.

3.4.3 Transverse momentum distribution of the electroweak bosons

As already discussed in Section 2.2, QCD corrections to vector boson production manifest themselves in the production of additional quarks or gluons in the final state which gives rise not only to the production of associated jets, but also to a substantial transverse momentum p_T of the produced boson. For the Z boson the p_T distribution can be measured inclusively without studying the details of the hadronic recoil. In this section measurements of the transverse momentum spectrum are reviewed. In the following chapter analyses on the production of jets in association with the weak bosons will be discussed in detail.

Besides being of interest for QCD studies, an accurate determination of the weak boson's transverse momentum distribution is needed for a precise measurement of the W boson mass (in particular if the mass is extracted from the transverse momentum spectrum of the charged lepton from the boson decay, which directly depends on the assumed W boson p_T [150]). Also a correct description of the kinematics of this major background in searches for new phenomena and the Higgs boson is mandatory.

In Run I CDF and DØ have both measured the $p_T(Z)$ distribution using the $Z \to e^+e^-$ selection [30,151]. For these measurements it is essential to study the dependence of the acceptance and the efficiencies as function of $p_T(Z)$. In particular the efficiency for the electron isolation requirement, which is needed to suppress the background from W+jet and multijet production, shows a distinct dependency on the transverse momentum of the Z boson. With increasing $p_T(Z)$ (but still at not too large $p_T(Z)$) the hadronic recoil is more likely to spoil the isolation of at least one of the two electrons. At very high $p_T(Z)$ the Z boson and thus the two electrons are boosted away from the recoil, so that the isolation efficiency recovers again.

For the measurement of the differential cross section $\mathrm{d}\sigma/\mathrm{d}p_T(Z)$ the data need to be corrected to take into account the finite energy resolution of the calorimeter. Resolution effects can be significant at low $p_T(Z)$, where the relative uncertainty on the measured value is large. For this correction DØ applied a Bayesian method based on statistical inference and resolution-smeared ansatz p_T distributions [31]. The smearing correction was found to be as large as 18.5% in the first bin corresonding to $0\,\mathrm{GeV} < p_T(Z) < 1\,\mathrm{GeV}$, but decreases rapidly so that for $p_T(Z) > 5\,\mathrm{GeV}$ the correction was always $\lesssim 5\%$. The DØ and CDF collaboration present their results on $\mathrm{d}\sigma/\mathrm{d}p_T(Z)$ with bin sizes at low $p_T(Z)$ as small as 1 GeV and 0.5 GeV, respectively. These values are considerably smaller than the resolution on $p_T(Z)$, which is about $\sim 1.5\,\mathrm{GeV}$. Therefore, the bin purity (i.e. the percentage of events reconstructed in the same bin as the true value) of the data points at low $p_T(Z)$ is relatively small and the result is largely sensitive to the assumptions on the ansatz p_T distribution and its uncertainty.

Fig. 3.9 shows the DØ measurement of $\mathrm{d}\sigma/\mathrm{d}p_T(Z) \times Br(Z \to e^+e^-)$ based on an integrated luminosity of $111\,\mathrm{pb}^{-1}$ compared to resummed calculations using different parameterizations and a fixed order calculation (cf. Section 2.2). The cross section peaks around 3 GeV and rapidly decreases towards $p_T(Z) = 0\,\mathrm{GeV}$. The calculation at fixed $\mathcal{O}(\alpha_s^2)$ fails to describe this shape but diverges at vanishing $p_T(Z)$. The resummed calculations include perturbative corrections in form of logarithmic terms beyond the fixed order expansion. Physically, these corrections correspond to the emission of mul-

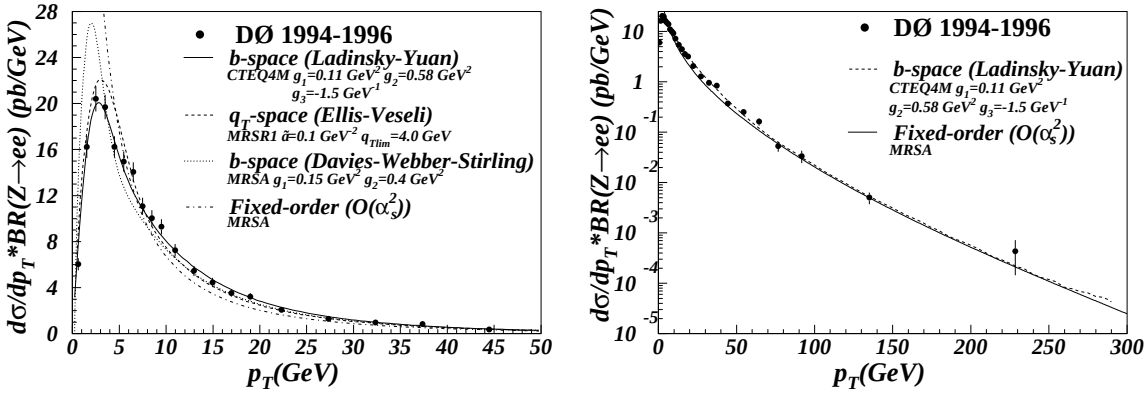

Figure 3.9: The differential production cross section of the Z boson as function of its transverse momentum p_T [30] compared with resummation calculations as well a a fixed-order calculation [42, 152, 153].

tiple soft partons which suppress the rate at low $p_T(Z)$. The resummation can be carried out either in p_T (q_T) space [152] or impact parameter (b) space [42,153]. In addition, phenomenological form factors are applied for the transition to the region of very low $p_T(Z) \lesssim \Lambda_{QCD}$, where perturbative QCD breaks down. While for the calculations based on the models of Ellis-Veseli [152] and Davies-Webber-Stirling [153] the default parameters for the form factor parameterization were used in the predictions shown in Fig. 3.9, the measurement was used to constrain one of the three non-perturbative parameters (g_2) in the model of Ladinsky-Yuan [42]. The data has hardly any sensitivity to constrain the other two parameters (g_1, g_3) simultaneously. The model of Ladinsky-Yuan is implemented in the event generator RESBOS [67].

The CDF result on $\mathrm{d}\sigma/\mathrm{d}p_T(Z)$ is in good agreement with both the DØ measurement and the RESBOS prediction. As demonstrated in Fig. 3.10 a good description of the Z boson's transverse momentum spectrum is also obtained with the SHERPA event generator [60, 91] which is based on the CKKW matching algorithm (cf. Section 2.3). The figure shows that at large $p_T(Z)$ processes with additional final-state partons above a scale Q_{cut} are of increasing importance. Therefore the parton shower event generators PYTHIA [36,69] and HERWIG [61,62], which only include corrections to describe a single hard gluon emission, underestimate the cross section at large $p_T(Z)$ (cf. Section 2.3).

In Run II the DØ collaboration has presented a preliminary measurement of the Z boson transverse momentum spectrum based on a sample of $Z \to e^+e^-$ events corresponding to an integrated luminosity of $0.98\,\mathrm{fb}^{-1}$ [34], which amounts to a nine-fold increase in statistics compared to the previous measurements. The systematic precision of the new measurement is limited by the uncertainty of the electron identification efficiency as function of $p_T(Z)$. The $p_T(Z)$ distribution was measured both for the inclusive sample of Z bosons, and for a sample of Z bosons produced at rapidities $|y| > 2$, which is shown in Fig. 3.11. The data is compared to the prediction of the RESBOS event generator which was interfaced to PHOTOS [155] to simulate photon final state radiation. For Z boson production at large rapidities, where one of the in-

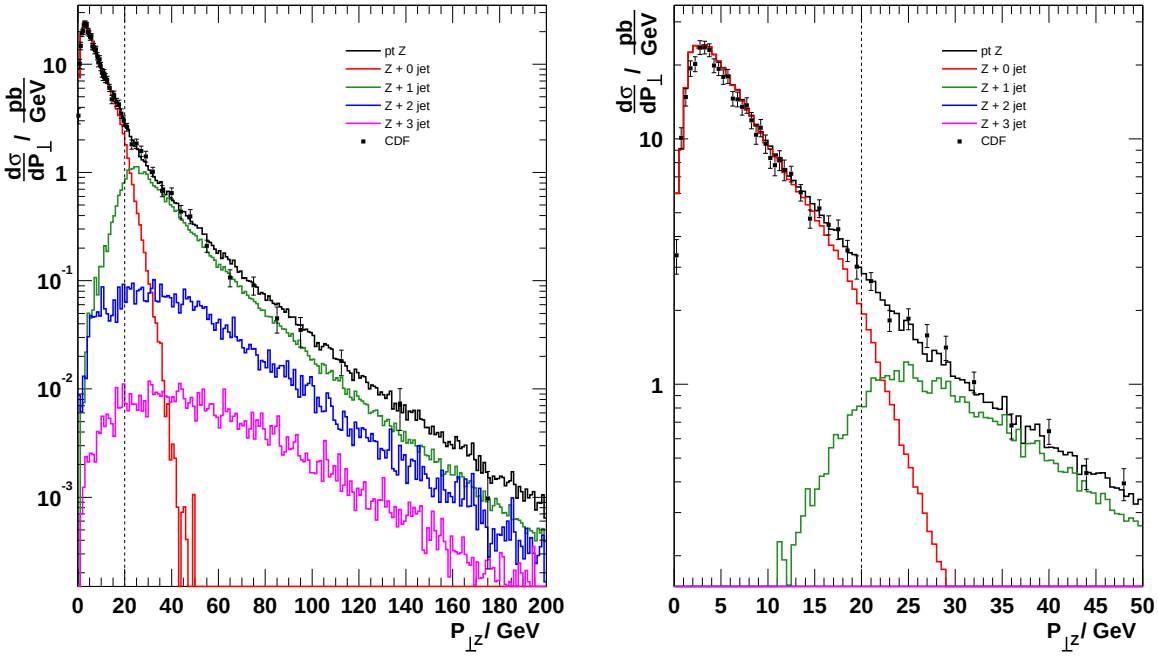

Figure 3.10: The transverse momentum p_T of the Z boson as predicted by the SHERPA *event generator [91] and compared to CDF Run I data [151]. The lines show the contributions of the different multiplicity processes and their sum. The dashed line indicates the separation cut at $Q_{\text{cut}} = 20\,\text{GeV}$.*

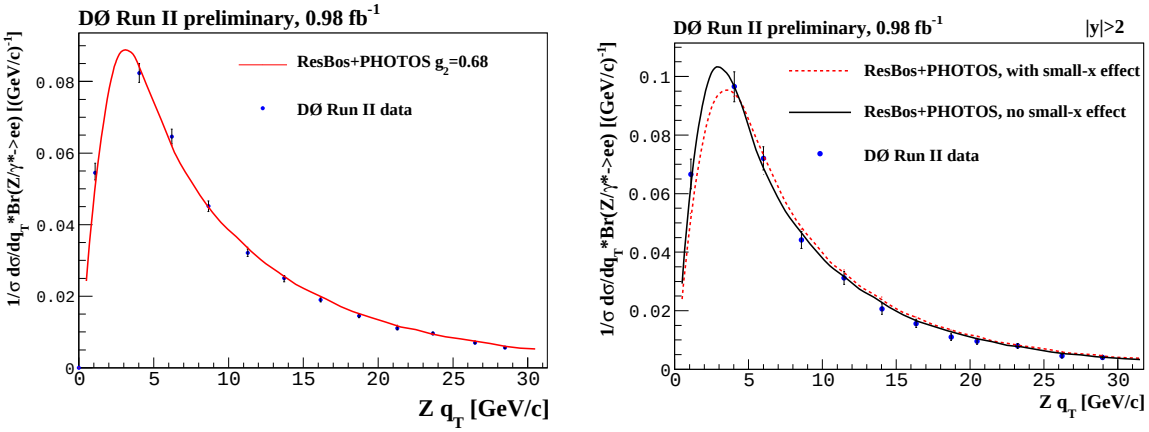

Figure 3.11: The Z boson p_T (q_T) distribution in $p\bar{p}$-collisions at $\sqrt{s} = 1.96\,\text{TeV}$ for inclusive Z production (left) and for Z bosons with rapidities $|y| > 2$ (right) [154]. The measurement is compared to the prediction obtained with the RESBOS *event generator [67] which was interfaced to* PHOTOS *[155] to simulate photon final state radiation. The distribution for Z-bosons produced at large rapidities is also compared to a model which includes a correction to account for the suggested small-x broadening effect [156].*

coming partons has a low momentum fraction x, it was suggested that the form factor of the Ladinsky-Yuan model implemented in RESBOS should be modified to address the broadening of the p_T distribution observed in semi-inclusive deep inelastic scattering [156]. In Fig. 3.11 the distribution for $|y| > 2$ is also compared to the RESBOS prediction, when including this modified form factor. The data at these large rapidities prefer the unmodified calculation. The transverse momentum spectrum of the Z boson was measured up to $p_T(Z) \approx 250\,\text{GeV}$. At high $p_T(Z)$ the data was found to be well described by the NNLO QCD prediction [39].

The DØ collaboration also published measurements of the W boson transverse momentum spectrum in Run I using $W \to e\nu$ events. The first result was based on an integrated luminosity of only $13\,\text{pb}^{-1}$ and was not corrected for detector resolution [29]. The second measurement, which was based on $85\,\text{pb}^{-1}$, was fully corrected [32]. While the understanding of the $p_T(W)$ distribution is an important ingredient to precisely measure the W boson mass, its direct measurement suffers from both the moderate energy resolution of the hadronic remnant and sizable backgrounds from QCD multijet productions. Integrated over all $p_T(W)$ this background was estimated to be 2% [32], but it largely increases as function of $p_T(W)$ [29]. The advantage of the $p_T(W)$ compared to the $p_T(Z)$ measurement is the about ten-fold increase in cross section times branching ratio. The measurement of $\mathrm{d}\sigma/\mathrm{d}p_T(W)$ was found to be in good agreement with the resummed calculation of Ladinsky-Yuan [32].

DØ has also published a measurement of the ratio of differential cross sections for W and Z boson production as a function of p_T in Run I [33]. In the ratio part of the experimental uncertainties as well as non-perturbative and radiative corrections in the phenomenological prediction cancel. The ratio can thus be calculated reliably using perturbative QCD without the need of resummation and non-perturbative form factors [157]. Good agreement between the data and the prediction was observed. Therefore, it was suggested to use the $p_T(Z)$ distribution to determine the $p_T(W)$ spectrum rather than measuring the latter directly, as the improvement in the systematic uncertainty outweighs the statistical penalty for the large integrated luminosities which are collected in Run II.

3.5 Conclusions for new phenomena searches

The total cross sections for the inclusive production of W and Z bosons are predicted with high precision, e.g. $\sim 3\%$ for $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$. Also, measurements at Run II reach precisions of $\sim 2\%$ (excluding the uncertainty on the luminosity determination). Thus, simulated background data from inclusive W or Z boson production can be normalized to both, theoretical predictions and measurements with minimal uncertainty. By contrast, Drell-Yan production at large invariant dilepton masses $M_{\ell\ell}$ obtains a substantial uncertainty due the precision of parton distribution functions which is deteriorating at very high parton momenta x. This uncertainty reduces the sensitivity of searches for numerous extensions of the standard model which predict the production of additional lepton-pairs at large $M_{\ell\ell}$ or new processes interfering with the SM production.

Differential distributions in inclusive vector boson production are well modelled

by QCD. For an accurate description of the boson's transverse momentum spectrum, corrections beyond the lowest order are required, including higher-order corrections at large $p_T(V)$ and resummation techniques with non-perturbative form factors at low $p_T(V)$. Although these calculations only model the boson kinematics, i.e. they cannot be used to simulate full W or Z boson events, their prediction can be utilized to reweight events simulated with full event generators. A correct simulation of the kinematic distributions is required to accurately predict background rates after signal selections.

Chapter 4

Associated production of vector bosons with jets

Higher order QCD corrections to W and Z boson production do not only manifest themselves in the generation of high transverse momenta of the bosons, but these QCD effects become directly visible in the production of jets in association with the weak bosons. Studying these processes is not only interesting from the perspective of testing perturbative QCD, but also to constrain a major background to many SM or non-SM physics signals, which will be discussed in the next chapter. The phenomenology of W/Z + jets production as well as the simulation of these processes using event generators was discussed in Chapter 2. This chapter covers measurements on W/Z + jet production performed at the Tevatron: first focusing on flavour inclusive jets (i.e. reconstructed jets without any flavour identification), secondly on heavy flavour jets in association with a weak boson.

4.1 Associated production of vector bosons with jets of any flavour

The CDF and DØ collaborations have measured the properties of jets accompanying W or Z bosons in various analyses performed during both Run I and Run II. The experimental challenges for these studies include, amongst others, the reconstruction of the jets, the corrections to their energy scale, and the determination of the background contribution.

4.1.1 Experimental Challenges

Quarks and gluons produced in hard parton collisions hadronize into an ensemble of particles which are boosted along the initial quark/gluon direction and thus form a collimated stream which is called a *jet*. The energy and direction of these particle jets can be approximated by reconstructed jets which are based on the energy deposits in the calorimeter[1]. The measurements discussed below use a fixed-cone clustering

[1] Algorithms which cluster tracks to reconstruct jets (*track-jets*) exist as well, but are not used to measure jet energies at the Tevatron.

algorithm in the plane given by the pseudorapidity η (or rapidity y) and the azimuth angle ϕ with radius $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ (or $R = \sqrt{\Delta y^2 + \Delta\phi^2}$) [158, 159]. The cone size typically ranges from $R = 0.3$ to $R = 0.7$. Whereas smaller cone sizes result in significant out-of-cone energy leakage due to fragmentation, larger cones suffer from an increased contribution of noise and underlying event. Thus non-optimal cone sizes result in a degradation of the jet energy resolution. For the study of multijet production one preferentially uses relatively small cone sizes to avoid an increased jet merging due to phase space limitations at high multiplicities.

The implementations of the cone algorithm in Run I [158] and Run II [159] differ in some important details, e.g.:

- Recombination scheme: Instead of combining the particles (or calorimeter towers) as massless objects, their four-vector is now combined. As a consequence the cone is defined in the $(y,\ \phi)$-plane instead of the $(\eta,\ \phi)$-plane. The new procedure implies Lorentz invariance under boosts along the beam axis.

- Seeds: The jet finding is started using calorimeter towers above a certain energy thresholds as seeds. The mid-points between proto-jets (jets which have been reconstructed at an initial stage in the algorithm) are included as additional seeds. This modification reduces the algorithm's sensitivity to soft radiation (infrared safety).

- Splitting and merging: In the Run I algorithm two jets are merged if their distance satisfies $\Delta R_{ij} < 1.3R$, which corresponds to the jet separation resolution. The new procedure to split or merge overlapping jets is based on the percentage of transverse energy shared by the less energetic jet.

The cone algorithm used in Run I is often denoted as iterative cone algorithm or simply as Run I algorithm. The new algorithm is usually called midpoint or Run II cone algorithm. The exact implementations by CDF and DØ slightly differ. Note also that some analyses performed by CDF in Run II are still based on the iterative cone algorithm.

The measured jet energy E_j^{meas} must be corrected for several effects to correspond to the particle or true jet energy E_j^{ptcl} [158, 160–162]. This jet energy scale correction can be applied as follows [160]:

$$E_j^{\text{ptcl}} = \frac{E_j^{\text{meas}} - E_0(R, \eta, \mathcal{L})}{\mathcal{R}_{\text{jet}}(R, \eta, E)\, S(R, \eta, E)} \ , \tag{4.1}$$

with the various corrections described below, which are almost entirely derived from data.

- $E_0(R, \eta, \mathcal{L})$ parameterizes an offset which includes the underlying event from multi-parton-interactions, noise, pile-up from previous beam crossings, and additional $p\bar{p}$-collisions in the same bunch crossing (minimum bias). The offset grows with increasing cone size R and luminosity $\mathcal{L}$. E_0 is determined from energy densities measured in minimum bias events.

- $\mathcal{R}_{\text{jet}}(R, \eta, E)$ is the energy response of the calorimeter to jets. It is determined from γ + jets events by requiring transverse momentum conservation. The response to photons and electrons is calibrated using $Z \rightarrow ee$ events. $\mathcal{R}_{\text{jet}}$ is typically less than one, due to energy lost in uninstrumented regions between calorimeter modules and due to non-perfect compensation (equal response to hadronic and electromagnetic interacting particles).

- $S(R, \eta, E)$ parameterizes the fraction of the jet energy deposited outside the cone with radius R. This showering correction is derived using the energy density within jets obtained from both data and simulation.

The jet energy correction procedure applied by CDF [161] is slightly different. As an additional step CDF also attempts to correct the jet energy to the one of the parent parton. One should note that a correction to the parton level cannot be unambiguously defined as it depends on the higher-order corrections included in the parton-level prediction.

Another potential problem might arise, when comparing the energy of reconstructed jets after jet energy scale corrections with the one of corresponding particle jets obtained using event generators. Whereas the reconstructed jet energy is corrected for the underlying event, this additional particle flow originating from multi-parton interactions (soft collisions of spectator quarks) is simulated by full-fledged event generators (e.g. PYTHIA, HERWIG, SHERPA). At low jet energies the inclusion of the underlying event can result in a substantial fractional corrections, as it can exceed 1 GeV for large cone sizes. Due to the steeply falling energy spectra of produced jets, this results in sizable changes in the multiplicity of jets above a minimal E_T.

Backgrounds in measurements of W + jets production can be substantial, especially for large jet multiplicities and large jet transverse momenta. In Z + jets events their contribution is much smaller as they are largely suppressed by the requirement of a reconstructed dilepton system consistent with the Z mass. The backgrounds to W + n jet production, where the W boson is identified via its decay into $e\nu$ or $\mu\nu$, can be classified in three categories: processes with a real (i.e. primary) electron or muon (mostly $t\bar{t}$, $Z \rightarrow ee/\mu\mu$, $W \rightarrow \tau\nu$, WW production), QCD multijet production with a fake (or secondary) lepton, and jet promotion. For the latter a jet from an additional $p\bar{p}$ interaction in the same bunch crossing (minimum bias) is reconstructed in the same event, thus promoting the jet multiplicity from $(n-1)$ to n.

Fig. 4.1 shows the background contributions as determined by CDF for their W + jets sample in Run II, where the W boson decays in the electron channel [163, 164]. The total background fraction is found to range from about 10% at a jet multiplicity of one and low jet E_T to $\sim 80\%$ at large jet transverse momenta. The background is dominated by QCD multijet production and at large jet multiplicities in addition by top pair production. Whereas the contribution of $t\bar{t}$ production and electroweak processes is estimated using Monte Carlo simulation, the large QCD multijet background, which cannot be reliably modelled, is estimated from data. Template distributions as function of missing transverse energy are created for signal and background and fitted to the observed spectrum. The template for the QCD multijet background is defined using a data sample which fulfils all selection requirements except one or more electron identification cuts.

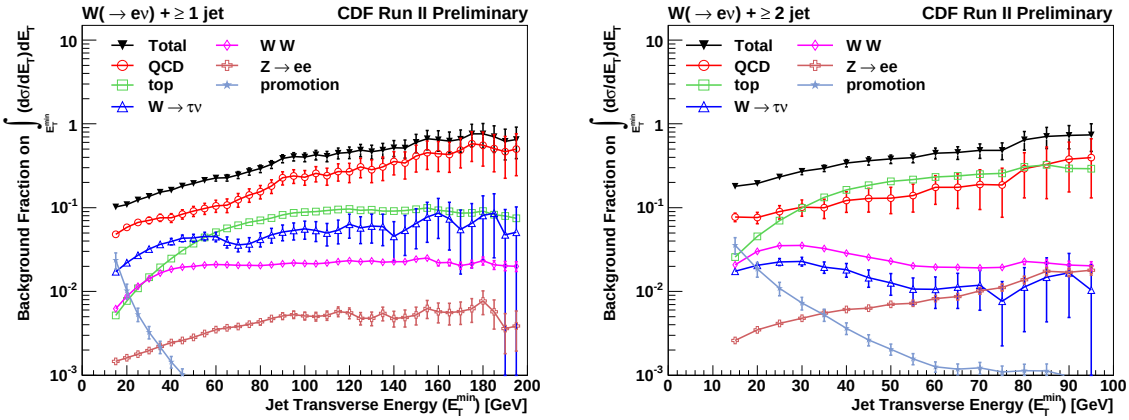

Figure 4.1: Background fraction as a function of the leading jet minimum transverse energy in CDF's $W(\rightarrow e\nu)+ \geq 1$ jet (left) and ≥ 2 jets (right) samples [163, 164].

4.1.2 Run I measurements

Properties of Jets in Z and W boson events

The CDF collaboration analysed the properties of jets in Z + jets [46] and W + jets [47, 165] events based on the electron decay channels $Z \rightarrow ee$ and $W \rightarrow e\nu$. The data sets were collected using high electron triggers and included integrated luminosities of $106\,\mathrm{pb}^{-1}$ and $108\,\mathrm{pb}^{-1}$, respectively.

Jets were reconstructed using the iterative cone algorithm with cone size $R = 0.4$ and jet energy scale corrections were applied. Jets with $E_T \geq 15\,\mathrm{GeV}$ and $|\eta| < 2.4$ were selected. The measured jet multiplicities were corrected for jet promotion. In the Z + jets analysis the probability to reconstruct an extra jet from an additional $p\bar{p}$ interaction in the same bunch crossing or from photons which are counted as hadronic interaction were 3% to 5% and 2% to 3%, respectively. For the W + jet analysis this probability was determined to be 1%.

Backgrounds with falsely reconstructed W or Z bosons are found to be dominated by jets faking electrons (including secondary electrons from heavy-flavour decays). They also comprise some contribution from the production of the other boson or a different decay channel. For W + jets a large contribution to the background at high jet multiplicities is from top pair production. The background contribution from QCD multijet production is estimated using a data selection without electron isolation requirement (and without $\not\!\!E_{\mathrm{T}}$ requirement in case of W + jets) and by extrapolating from the background dominated kinematic region into the signal region. In case of Z + jet production this background is small, i.e. consistent with zero, and upper limits (1σ) of 1.1% for jet multiplicity $n \geq 0$ to 4.0% for $n \geq 3$ were found. For W + jets the QCD multijet background was estimated to increase from $(2.9 \pm 0.9)\%$ for $n \geq 0$ to $(27 \pm 11)\%$ for $n \geq 4$. At large jet multiplicities ($n \geq 4$) top pair production has a similar contribution, which was estimated to be $(26 \pm 11)\%$.

The raw jet multiplicity distributions were corrected for efficiencies and acceptance.

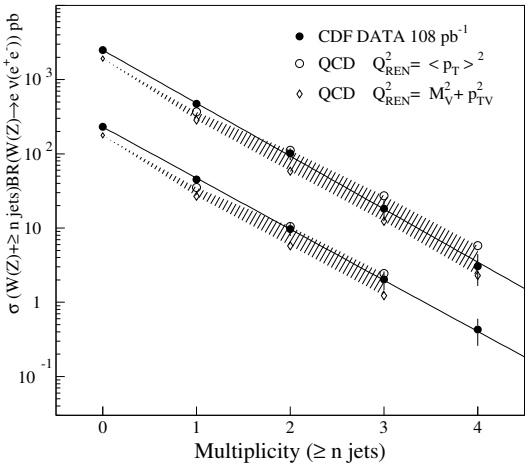

*Figure 4.2: $\sigma \cdot Br$ for $W^{\pm} \to e^{\pm}\nu + \geq n$ jets (top) and $Z \to e^{+}e^{-} + \geq n$ jets (bottom) versus inclusive jet multiplicity [47]. The solid lines are exponential fits to the data. The data are compared to LO calculations (*VECBOS*) using different renormalization scales [48].*

The cross sections for $Z/W + n$ jet production as function of jet multiplicity were derived using the corrected ratio of the number of events with $\geq n$ jets over the total number of events multiplied with the previously measured cross sections for inclusive W and Z boson production [112].

Several systematic uncertainties on the number of reconstructed jets were studied, including uncertainties on the jet energy scale and on the rate of jets from additional $p\bar{p}$ interactions. The combined jet counting uncertainty was estimated to range from 11% at $n \geq 1$ to 23% at $n \geq 4$ for Z + jets and 10% to 30% for W + jets. This dominates the total uncertainty on the cross section measurement.

The measured cross sections for W + jets and Z + jets production as function of inclusive jet multiplicity are shown in Fig. 4.2. The cross sections fall exponentially with jet multiplicity n with a suppression factor $\sigma(n)/\sigma(n-1) \sim 0.2$. The results for Z + jets and W + jets production show the same general features. The measurements are compared to LO QCD predictions obtained with the VECBOS program [48] (cf. Section 2.2.1) using a two-loop α_s evolution and two different choices for the renormalization scale, $Q^2_{\text{ren}} = M_V^2 + p_{T,V}^2$ and $Q^2_{\text{ren}} = \langle p_T \rangle^2$, with $\langle p_T \rangle$ being the average transverse momentum of the jets. The CTEQ3M PDFs [166] with the factorization scale set to the renormalization scale were used. The parton-level prediction of VECBOS was interfaced to the HERWIG event generator [61] to simulate initial state radiation, fragmentation, and hadronization. The addition of a parton shower simulation to the LO matrix-element calculation mimics part of the higher order corrections. Thus, an improved description of kinematic distributions, compared to the parton level predictions of VECBOS alone, is expected. Nevertheless this procedure does not include a correction to the integrated LO cross sections and does not provide a method to combine the VECBOS calculations for different jet multiplicities (cf.

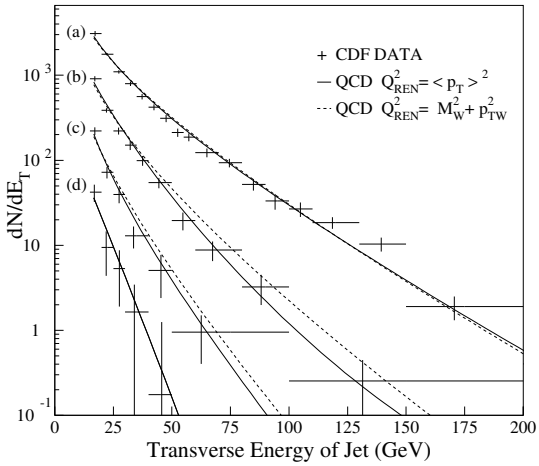

Figure 4.3: Transverse energy distribution of jets in W + jets events measured by CDF in Run I, for (a) the highest E_T jet in ≥ 1 jet events, (b) the second highest E_T jet in ≥ 2 jet events, (c) the third highest E_T jet in ≥ 3 jet events, and (d) the fourth highest E_T jet in ≥ 4 jet events [47]. The lines show the LO predictions based on VECBOS *[48] interfaced to* HERWIG *[61].*

matching, see Sect. 2.3.5). The generated events are passed through a full simulation of the CDF detector and finally the same reconstruction and selection is applied as for real experimental data. The comparison at reconstructed level ensures that data and simulation can be directly compared, but compromises the comparison to any other model.

The simulation using the scale $Q^2_{\text{ren}} = M^2_V + p^2_{T,V}$ predicts about the correct shape for the jet multiplicity and similar cross section ratios $\sigma(n)/\sigma(n-1)$ are obtained, but the normalization is about a factor 1.6 too low. For inclusive vector boson production (i.e. $n \geq 0$) a factor of about 1.35 is expected to account for the higher-order terms which are absent in the calculation (see Section 3.1). The K-factor for larger multiplicities (calculated up to $n \geq 2$ in NLO) is of similar size. The softer scale $Q^2_{\text{ren}} = \langle p_T \rangle^2$ predicts higher cross sections (as α_s increases with decreasing scales), but the shape of the multiplicity distribution is less well modelled. This scale is also less motivated from phenomenological arguments: It is considerably smaller than the vector boson mass (for the bulk of the phase space) and α_s diverges towards vanishing jet transverse momentum. Within the substantial uncertainty of the LO QCD calculation, the prediction is in agreement with the measured multiplicity distributions.

To further test the QCD prediction several kinematic distributions of the jets were studied. The measured distributions were compared to full simulation at detector level without applying any efficiency or acceptance corrections. Fig. 4.3 shows the transverse energy spectrum of the first to fourth highest E_T jet in W + jets events. The data is compared to the QCD prediction, which is normalized separately for each jet multiplicity. Within its uncertainty the shapes of the E_T distributions are well described. Furthermore, it was shown that the QCD prediction can reproduce the

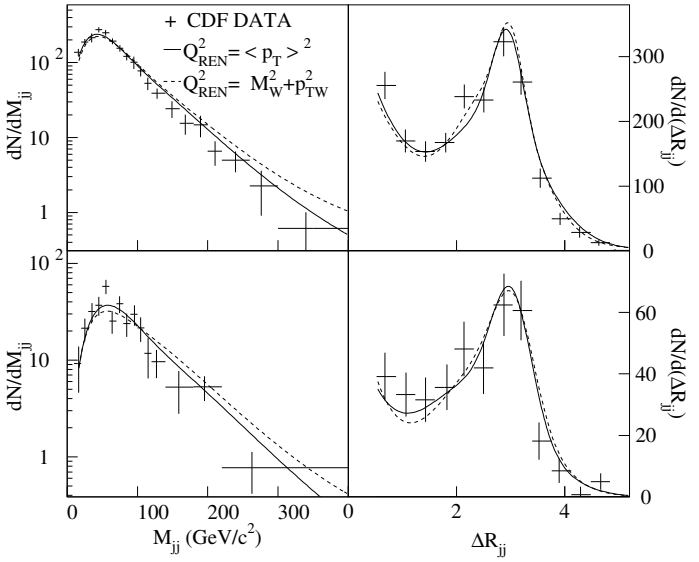

Figure 4.4: Dijet mass M_{jj} and jet separation $\Delta R_{jj} = \sqrt{\Delta\eta_{jj}^2 + \Delta\phi_{jj}^2}$ for the two leading jets for W + ≥2 jet events (top) and W + ≥3 jet events (bottom) [47]. The lines show the LO predictions based on VECBOS *[48] interfaced to* HERWIG *[61].*

main features of jet-jet correlation, e.g. the dijet mass distribution M_{jj} and the jet separation in $\eta - \phi$ space, $\Delta R_{jj} = \sqrt{\Delta\eta_{jj}^2 + \Delta\phi_{jj}^2}$, which is demonstrated in Fig. 4.4 for $W + \geq 2$ jet and $W + \geq 3$ jet events. Nevertheless, some deviations, e.g. at high M_{jj}, are observed and the scale uncertainty is substantial. As already noted, the simulation is lacking a mechanism to properly match the parton-level matrix-element calculation with the parton shower and to combine samples based on different parton multiplicities. As the data are not fully corrected, a comparison to advanced event generators as ALPGEN [43] or SHERPA [60], which both incorporate a matching prescription (cf. Section 2.3.5), is not possible.

Study of the strong coupling

The probability to produce a jet in association with a Z or W boson depends on the value of the strong coupling constant α_s. At leading order the cross section for $V + 1$ jet events is proportional to α_s.

The DØ collaboration attempted to extract α_s from the ratio $\mathcal{R}_{10}$ of $W + 1$ jet and $W + 0$ jet cross sections measured using $W \to e\nu$ decays collected during Run IA [167]. The jets were identified using the cone algorithm with $R = 0.7$ and were required to have a minimum transverse momentum of 25 GeV. $\mathcal{R}_{10}$ was determined with an accuracy of about 10% and compared with NLO QCD predictions [168]. These predictions, which were computed using different choices for the PDFs, were below the data by over one standard deviation. The study concluded that the sensitivity of the measurement

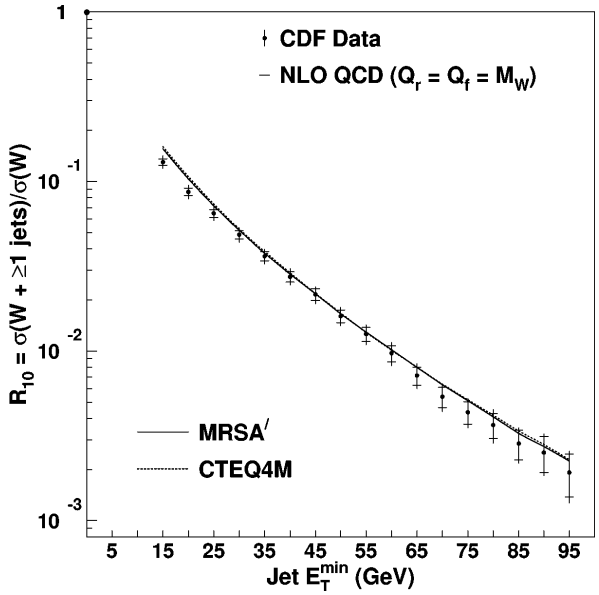

Figure 4.5: CDF Run I measurement of $\mathcal{R}_{10} = \sigma(W+ \geq 1jets)/\sigma(W)$ as function of jet $E_T^{\min}$ [169]. The data are compared to NLO QCD predictions [168] using the MRSA′ [170] and CTEQ4M [171] parton distribution functions. The W boson mass is chosen as renormalization and factorization scale and $\alpha_s(M_Z)$ is given by the PDF set (MRSTA′: 0.113, CTEQ4M: 0.116).

to α_s is largely reduced once the variation of α_s is also taken into account in the PDFs.

A similar measurement of the ratio $\mathcal{R}_{10} = \sigma(W+ \geq 1\,\text{jet})/\sigma(W)$ was performed by the CDF collaboration based on the full Run I data set [169]. The jet reconstruction and energy correction follows closely the $W+n$ jet multiplicity analysis discussed above. The ratio $\mathcal{R}_{10}$ was measured as function of the minimum jet $E_T^{\min}$ between $15\,\text{GeV} \geq E_T^{\min} \geq 95\,\text{GeV}$. The uncertainty in the jet energy scale and in the determination of the QCD multijet background were found to dominate the systematic uncertainty, which was estimated to range from 8% at $E_T^{\min} = 15\,\text{GeV}$ and 19% at $E_T^{\min} = 95\,\text{GeV}$. The measurement of $\mathcal{R}_{10}$ is shown in Fig. 4.5 together with the NLO QCD prediction [168] obtained using the MRSA′ and CTEQ4M PDFs. In the NLO calculation two partons are merged into a single jet if their distance is $\Delta R < 0.52$, which corresponds to the minimal separation of jets reconstructed with a cone of $R = 0.4$. A variation of the separation ΔR by $\pm 30\%$ was found to change the prediction for $\mathcal{R}_{10}$ by 10%. Instead of correcting the measurement for migration effects due to the limited resolution in the reconstructed jet energy and direction, the prediction was smeared to match the resolution observed in data. The NLO theory describes the data well above $E_T^{\min} \geq 30\,\text{GeV}$. For low $E_T^{\min}$ soft gluon radiation which is not included in the fixed order calculation could be significant. Similar to the findings of the preceding DØ analysis, the measurement of $\mathcal{R}_{10}$ has little sensitivity on the strong coupling constant and is consistent with $0.105 \geq \alpha_s(M_Z) \geq 0.130$. The slight disagreement between data and theory observed by DØ was not confirmed.

This CDF result can also be confronted with their preceding cross section measurement for $W + \geq n$ jets production [47], which is based on the same data set and

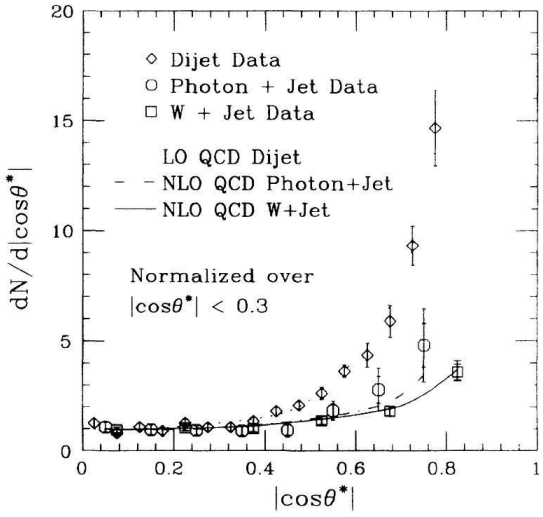

Figure 4.6: $\mathrm{d}N/\mathrm{d}|\cos\theta^*|$ *for* W *+ jet events measured by CDF in Run I [172], compared to dijet [173] and photon + jet data [174] and (N)LO QCD predictions.*

uses the same jet reconstruction and $E_T^{\min} = 15\,\mathrm{GeV}$, albeit a direct comparison of the measured values for $\mathcal{R}_{10}$ is impeded as the corrections applied are not identical. In Ref. [47] the jet rate at $E_T^{\min} = 15\,\mathrm{GeV}$ was measured to be $\mathcal{R}_{10} = 0.189 \pm 0.021$ and the LO QCD prediction was estimated to be 0.15 using a scale of $Q = \sqrt{M_W^2 + p_{T,W}^2}$. In Ref. [169] $\mathcal{R}_{10} = 0.130 \pm 0.012$ is obtained for $E_T^{\min} = 15\,\mathrm{GeV}$ compared to a NLO prediction of 0.156 (at $Q = M_W$). Ref. [169] did not address possible reasons for this difference in $\mathcal{R}_{10}$.

Further studies

The CDF collaboration measured the angular distribution in W + jet events using a $W \to e\nu$ sample corresponding to an integrated luminosity of $23\mathrm{pb}^{-1}$ recorded during the initial phases of the Tevatron Run [172]. The shape of the angular distribution gives an indication of the spin of the exchanged particle in the scattering process. As W + jet events are predominantly produced by quark exchange processes (cf. Sct. 2.2.1), one expects that the W angular distribution is approximately of the form $\mathrm{d}N/\mathrm{d}\cos\theta^* \sim (1 - |\cos\theta^*|)^{-1}$ due to the spin-$\frac{1}{2}$ propagator term. Here, the polar angle θ^* with respect to the proton beam direction is given in the centre-of-mass frame of the system consisting of the W boson and the highest E_T jet.

As the longitudinal momentum of the neutrino p_z^ν from the W decay cannot be inferred from momentum conservation, it is impossible to fully reconstruct the four-momentum of the W boson. However, a constraint to the W mass leaves only two solutions for p_z^ν, of which the one is selected which is preferred due to the W polarization (cf. Sct. 3.4.1). From Monte Carlo studies it was found that in $\sim 73\%$ of the cases the correct solution was chosen.

The measurement of $\mathrm{d}N/\mathrm{d}|\cos\theta^*|$ is given in Fig. 4.6 and compared to the NLO

QCD prediction [168], which is in good agreement. Also shown in the figure are previous measurements of CDF on the angular distributions in γ + jet and dijet events [173,174]. The γ + jet production is dominated by quark exchange as well (mostly $gq \to \gamma q$) and thus has a similar distribution of $\mathrm{d}N/\mathrm{d}|\cos\theta^*|$. In contrast to W + jet and γ + jet, dijet production is dominated by the gluon propagator and thus has an angular distribution of the form $\mathrm{d}N/\mathrm{d}\cos\theta^* \sim (1 - |\cos\theta^*|)^{-2}$, which gives a steeper rise towards the forward/backward region.

The DØ collaboration reported a study on colour coherence effects in W + jet events [175]. Initial-to-final-state colour interference effects were observed by comparing the distributions of soft particles (as measured in the multiplicity distribution of calorimeter towers above an energy threshold) around the colourless W boson and the leading jet in the hemisphere opposite to the W boson. Soft particle radiation is enhanced around the jet in the event plane (i.e. the plane which is defined by the direction of the W boson or jet and the proton beam) with respect to the transverse plane. This enhancement is consistent with the predictions of the Pythia event generator [176], which accounts for colour coherence effects at the perturbative and non-perturbative level by means of the angular ordering approximation in the parton shower and the string fragmentation, respectively. Also an analytic perturbative QCD calculation [177] based on the modified leading logarithmic approximation (MLLA) and local parton-hadron duality (LPHD) is able to describe this enhancement.

4.1.3 Run II cross section measurements

The large luminosities being collected during Tevatron Run II and the recent progress in both NLO calculations and in the development of event generators for the associated production of vector bosons with jets (cf. Sct. 2.3) allows to measure W/Z + jet production and to test the QCD models with unprecedented precision. The DØ collaboration published a measurement of $Z + \geq n$ jet cross sections [178] and the CDF collaboration presented preliminary results on both Z+ jet and W + jet production [164, 179].

The DØ collaboration measured the jet multiplicity produced in association with a Z boson based on a data sample of about 14,000 $Z \to e^+e^-$ candidates collected using a single electron trigger and an integrated luminosity of $0.4\,\mathrm{fb}^{-1}$ [178]. Two electrons were required to be reconstructed in the central region of the calorimeter corresponding to $|\eta| \leq 1.1$. The background from jets which are misidentified as electrons was suppressed by using quality criteria based on the shower profile of the electron candidate. Jets were reconstructed using the Run II cone algorithm with a cone radius $R = 0.5$ and jet energy scale corrections (cf. Sct. 4.1.1) were applied. The jets were required to have a transverse energy $E_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$.

The electron selection efficiencies were derived from data using the *tag-and-probe* method (cf. Sct. 3.2). All efficiencies were found to be independent of the jet multiplicity except for the efficiency to find a track matched to the reconstructed electron cluster, which was taken into account when correcting the data. The kinematic and geometric acceptance was determined using an inclusive and fully simulated Pythia sample, which was reweighted in order to correctly describe the observed p_T distribution of the Z boson. The jet reconstruction and identification efficiency was obtained

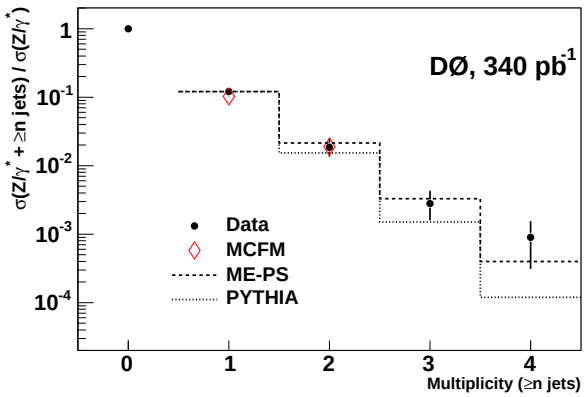

Figure 4.7: Jet inclusive $Z/\gamma^* + \geq n$ jet cross sections normalized to the total inclusive Z/γ^* cross sections for jet multiplicities n up to 4 measured by DØ in Run II [178]. The data are compared to a calculation which matches the LO matrix-elements with a parton shower (ME-PS, dashed line) [88] and to the prediction of Pythia *(dotted line) [36], both normalized to the $Z/\gamma^* + \geq 1$ jet cross section. The diamonds represent a NLO calculation given by* Mcfm *[45] for jet multiplicities up to 2.*

from full detector simulation, which was corrected to account for differences between data and simulation. This correction was determined from the efficiency to find a recoiling jet opposite in azimuth to the Z boson, which was measured both in data and simulation and parameterized as a function of $p_T(Z)$. The jet efficiencies were then calculated from matching particle jets based on generator-level particles to reconstructed calorimeter jets.

Backgrounds from QCD multijet production in which jets are wrongly identified as electrons were estimated to be $3-5\%$ depending on the jet multiplicity based on the dielectron invariant mass spectrum observed in data compared to the Drell-Yan spectrum obtained in simulation. Backgrounds from $t\bar{t}$ production, $Z \to \tau^+\tau^-$, and $W \to e\nu$ were found to be $< 1\%$.

The measured cross section as function of jet multiplicity was fully corrected to particle level, including corrections for the jet reconstruction and identification efficiency, and migration effects due to the finite resolution of the jet energy measurement. These corrections were determined using Monte Carlo samples, which were reweighted to correctly describe the observed jet multiplicity and jet E_T distributions. Two independent event generators were used: Pythia [36] and an implementation of the CKKW scheme which combines matrix-element calculations obtained with Madgraph with the parton shower of Pythia [88] (cf. Section 2.3.5). The latter is denoted ME-PS in the following.

Fig. 4.7 shows the corrected cross section normalized to the inclusive Z/γ^* cross section

$$R_n = \frac{\sigma(Z/\gamma^* + \geq n\,\text{jets})}{\sigma(Z/\gamma^*)} \tag{4.2}$$

for the mass region $75\,\text{GeV} < M_{ee} < 105\,\text{GeV}$. The systematic uncertainties in this

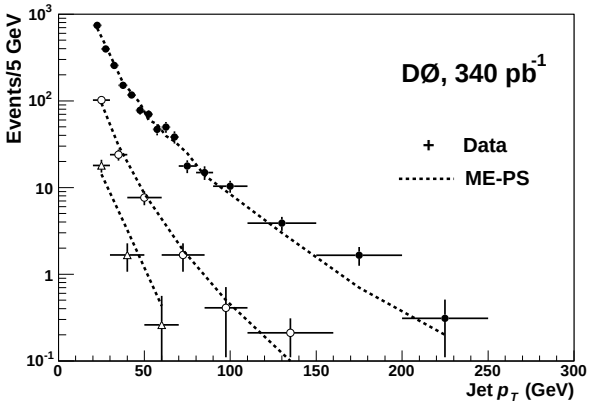

Figure 4.8: Transverse momentum distribution of jets in Z + jets events measured by DØ in Run II, for the highest p_T jet in ≥ 1 jet events (dark circles), the second highest p_T jet in ≥ 2 jet events (open circles), the third highest p_T jet in ≥ 3 jet events (open triangles) [178]. The dashed lines show the ME-PS prediction [88], which was normalized to the data. The data are uncorrected and the ME-PS prediction is after full detector simulation.

measurement are dominated by contributions from jet energy scale corrections, jet reconstruction and identification, the unsmearing method to determine the migrations in the jet E_T distributions, and the variation of electron related efficiencies with jet multiplicity. R_n falls nearly exponentially with a suppression factor of $\sim 1/8$ per jet.

The data are compared to the predictions of Pythia, the ME-PS simulation, and to NLO calculations obtained with the Mcfm code [45]. The NLO computation, which can be performed up to parton multiplicity $n = 2$, was obtained using M_Z as the renormalization and factorization scale and the CTEQ6M PDFs [94]. It gives a good description of the data, despite the fact that no correction to translate the parton level prediction of Mcfm to the particle level was applied. The Pythia and ME-PS predictions were normalized to the $Z/\gamma^* + \geq 1$ jet cross section ratio. The ME-PS prediction is in good agreement with the data, whereas Pythia underestimates the rate at high jet multiplicities, which is expected as the parton shower algorithm generates less hard additional partons compared to higher-order matrix-element calculations (cf. Section 2.3.6).

Fig. 4.8 shows the uncorrected jet p_T spectra of the first, second, and third highest p_T jet in Z/γ^*+ 1, 2, 3 jet events. The ME-PS Monte Carlo after full detector simulation is in good agreement with the data.

The CDF collaboration presented preliminary measurements of the inclusive jet cross section as function of the jet transverse momentum and of the jet multiplicity in $Z(\rightarrow ee)$ + jets events based on a data sample with an integrated luminosity of 1.1 fb^{-1} collected with a single electron trigger [179]. Jets were reconstructed using the midpoint algorithm with a cone size $R = 0.7$ and were required to have transverse momenta ${p_T}^{\text{jet}} > 30\,\text{GeV}$ and rapidities $|y_{\text{jet}}| < 2.1$. Z boson candidates were selected based on a reconstructed central electron fulfilling tight identification cuts and a second electron with loose identification requirements which can be either central or forward. The

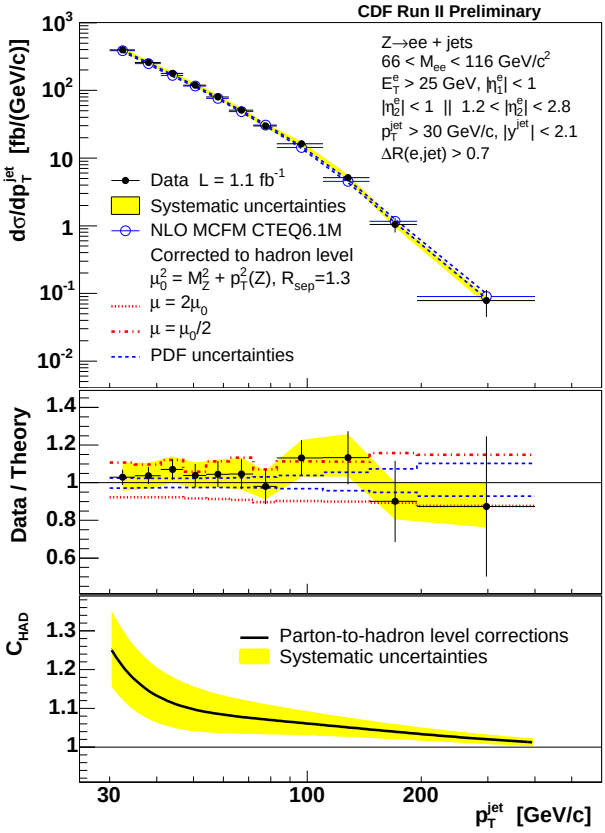

Figure 4.9: Transverse momentum distribution of jets in $Z + \geq 1$ jet production measured by CDF in Run II [179] compared to NLO QCD predictions obtained with MCFM *[45]. The middle plot shows the ratio data/theory with lines indicating the PDF and scale uncertainty on the prediction. The QCD prediction was corrected from parton to hadron (particle) level (bottom plot). The cross section is evaluated for the phase-space cuts stated in the figure.*

electrons were required to have a minimum separation from all jets of $\Delta R(e, \text{jet}) > 0.7$. No further isolation with respect to nearby reconstructed tracks or energy deposits in the calorimeter was imposed.

The dominating background from QCD multijet and W + jets processes was determined based on data. The electron fake rate, which denotes the probability for a jet to pass the electron selection criteria, was estimated using an inclusive jet sample and was applied to jets found in a data sample with exactly one reconstructed tight central electron. Other backgrounds which comprise additional electroweak processes and $t\bar{t}$ production were obtained from Monte Carlo simulation. The total background amounts to $10 - 14\%$ depending on jet multiplicity. The relatively large background contribution compared to other similar measurements is likely due to the loose electron isolation requirement.

Fig. 4.9 shows the jet transverse momentum distributions $\mathrm{d}\sigma/\mathrm{d}p_T^{\text{jet}}$ of the jets in $Z + \geq 1$ jet events measured between $30\,\text{GeV} < p_T^{\text{jet}} < 300\,\text{GeV}$. The cross section is unfolded to hadron (particle) level using an iterative bin-by-bin unsmearing method.

Instead of correcting the cross section to the full phase-space of Z boson production and decay, it is derived in a well-defined kinematic range which matches the detector acceptance and event selection to avoid additional systematic uncertainties due to the extrapolation to the full kinematic region. The systematic uncertainty on the cross section measurement is dominated by uncertainties in the jet energy scale (resulting in an error of 5% to 12% which is increasing with p_T^{jet}), the electron identification efficiency (5%), and the luminosity normalization (5.8%). The background estimation and the correction to the hadron level each account for 1% systematic uncertainty.

The CDF measurement of $\mathrm{d}\sigma/\mathrm{d}p_T^{\text{jet}}$ shown in Fig. 4.9 is compared to the NLO QCD prediction obtained with the MCFM program [45] which has been corrected from parton to hadron level using the PYTHIA event generator with a tuned parameter setting to correctly describe the underlying event ("Tune A") [96]. This parton-to-hadron level correction, which is defined as the ratio of the full PYTHIA prediction over the one without multi-parton interactions and fragmentation, is shown in the bottom panel of Fig. 4.9 including a systematic uncertainty which was derived using an alternative fit of PYTHIA parameters ("Tune DW") [180]. This correction is non-negligible especially at small p_T^{jet}, where it exceeds a value of about 1.2. For smaller cone sizes (cf. the previously discussed DØ analysis) this correction is smaller as the energy bias due to the underlying event is decreasing. The MCFM prediction was computed with the CTEQ6.1M PDFs [101] and with the renormalization and factorization scales set to $\mu_0^2 = M_Z^2 + p_T^2(Z)$. The middle panel of the figure displays the ratio of the CDF data to the calculation together with the PDF and scale uncertainties of the prediction. MCFM gives a good description of the jet transverse momentum distribution over the full range in p_T^{jet}.

Fig. 4.10 displays the $Z(\rightarrow ee) + \geq n$ jets production cross section corrected to hadron level as function of the jet multiplicity n. The NLO calculation, which is available only up to $n = 2$, is in good agreement with the data, whereas the LO prediction underestimates the data by a factor ∼1.35 nearly independent of jet multiplicity.

The CDF collaboration also reported a preliminary measurement of inclusive $W(\rightarrow e\nu) + n$ jet production based on an integrated luminosity of $320\,\mathrm{pb}^{-1}$ collected using a high E_T electron trigger [164]. Differential cross sections as function of jet transverse energy, dijet mass, and jet separation were presented.

The jets were reconstructed using the iterative seed-based cone algorithm (Run I algorithm) with a cone radius $R = 0.4$ and were required to have a minimal transverse energy $E_T^{\text{jet}} > 15\text{GeV}$ and pseudorapidities $|\eta| < 2.0$. As described in Section 4.1.1 and shown in Fig. 4.1 the background to W + jets production was determined to be about 10% at a jet multiplicity of one and low jet E_T and to increase to ∼ 80% at large jet transverse momenta. The cross sections were evaluated for a limited but well defined W phase space to match the detector acceptance and event selection: $E_T^e > 20\,\mathrm{GeV}$, $|\eta^e| < 1.1$, $\not\!\!E_{\mathrm{T}} > 30\,\mathrm{GeV}$, and W boson transverse mass $m_T^W > 30\,\mathrm{GeV}$. Within this kinematic region the W reconstruction efficiency was found to be $(60 \pm 3)\%$. The results are corrected to hadron level using the ALPGEN generator [43] interfaced to PYTHIA-TUNE A [96] (cf. Section 2.2.1).

Fig. 4.11 displays the corrected differential cross section $\mathrm{d}\sigma(W(\rightarrow e\nu) + \geq n\,\text{jets})/\mathrm{d}E_T^n$ as function of the E_T^n of the n^{th} jet in W boson events with ≥ 1, ≥ 2, ≥ 3, and ≥ 4 jets. The systematic error was found to be dominated by the uncertainty on the jet energy

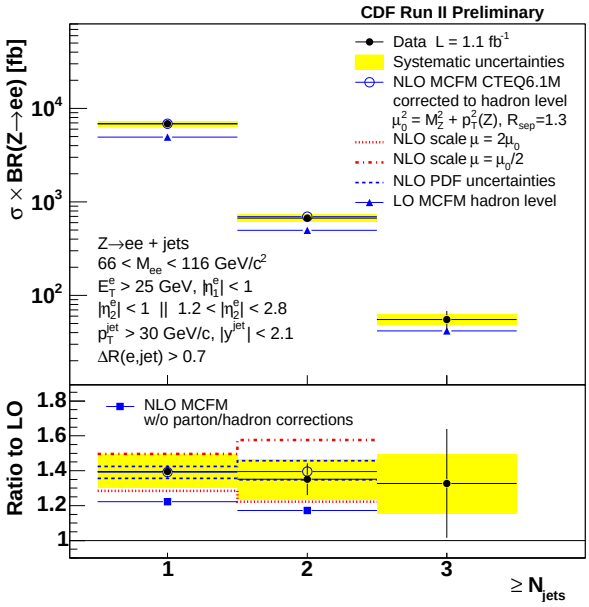

Figure 4.10: Cross section for $Z(\rightarrow ee) + \geq n$ jet production as function of jet multiplicity n [179] compared to LO on NLO QCD predictions [45]. The lower plot shows the data and the NLO calculation with and without parton-to-hadron correction normalized to the LO prediction, including scale and PDF uncertainties.

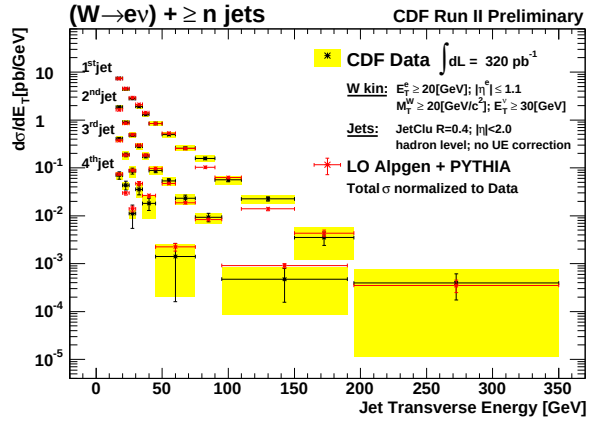

Figure 4.11: Differential cross section $\mathrm{d}\sigma(W(\rightarrow e\nu) + \geq n\,\mathrm{jets})/\mathrm{d}E_T^n$ as function of E_T^n of the n^{th} jet in W boson events with ≥ 1, ≥ 2, ≥ 3, and ≥ 4 jets [163, 164]. The data are compared to the predictions of Alpgen *(interfaced to* Pythia*) [36, 43], normalized to the measured inclusive cross section for each jet multiplicity. The cross section is evaluated for the phase-space cuts stated in the figure.*

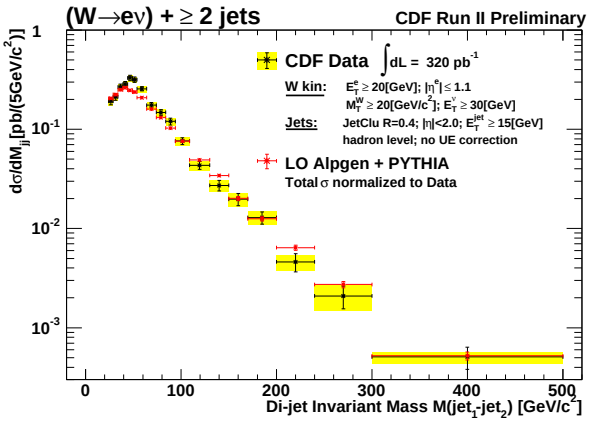

Figure 4.12: Differential cross section $\mathrm{d}\sigma(W(\to e\nu)+ \geq 2\text{jets})/\mathrm{d}M_{jj}$ *as function of the invariant mass* M_{jj} *of the two leading jets [163, 164]. The data are compared to the predictions of* ALPGEN *(interfaced to* PYTHIA*) [36, 43], normalized to the measured inclusive cross section.*

calibration at low jet E_T, while at large E_T the uncertainty on the background determination is the main contribution, which results in an error $> 50\%$ for $E_T > 150\,\mathrm{GeV}$. The measurement was compared to the prediction of ALPGEN interfaced to PYTHIA which was normalized to data for each multiplicity bin separately to allow a direct comparison of the shape of the distributions. ALPGEN gives a reasonably good description of the shape of the data. Due to the missing higher order corrections in the LO event generator ALPGEN the normalization (integrated cross section) is underestimated. It was also found that ALPGEN interfaced to PYTHIA is able to describe the topology of W + jet events as demonstrated with the differential cross section $\mathrm{d}\sigma(W(\to e\nu)+ \geq 2\,\text{jets})/\mathrm{d}M_{jj}$ as function of the invariant mass M_{jj} of the two leading jets, which is shown in Fig. 4.12.

4.1.4 Validation of event generators

As already mentioned before, it is essential to validate event generators with measurements of W/Z + jets production. Thus, their prediction could be taken to accurately estimate the background contribution from these events in measurements of rare standard model processes and searches for the Higgs boson or new phenomena. In the previous section it was shown that event generators which match fixed-order matrix-element calculations with a parton shower simulation give a good description of differential cross sections as function of various kinematic quantities, albeit they predict a too low rate due to the missing higher-order corrections in the matrix-element calculation. While the measurement of corrected cross sections is essential to obtain a universal result, the comparison of uncorrected data to fully simulated Monte Carlo events facilitates the comparison and even increases its precision as some of the systematic uncertainties cancel in the ratio of data to simulation at reconstruction level.

There is a qualitative difference in the validation of the event generators in the regions of low and high jet or vector boson transverse momenta. At low p_T the pre-

dictions of the event generators largely depend on non-perturbative assumptions and model parameters, e.g. assumptions being made for parton-shower cut-offs, in the fragmentation and for the underlying event. These model parameters can generally be tuned to data, which, for example, has been done for the underlying event at the Tevatron [96, 180]. At large transverse momenta the differential cross sections are primarily predicted by perturbative calculations and thus are less dependent on model assumptions, albeit they have a considerable uncertainty due to scale dependences etc. It is essential to validate the event generators in a kinematic region with negligible contribution from new physics processes to avoid a bias which would eventually cover a new signal. Subsequently these tested event generators can be used to predict the rate of standard model processes in the signal regions. In particular, a thorough validation of event generators at the Tevatron is of large importance for the initial modelling of the standard model backgrounds at the LHC.

The DØ collaboration reported a preliminary comparison between uncorrected $Z/\gamma^*(\rightarrow ee)$ + jet data with simulations obtained with the Pythia [181] and Sherpa [60] event generators (cf. Section 2.3.5). The data sample with an integrated luminosity of $950\,\mathrm{pb}^{-1}$, which was collected using a combination of single and dielectron triggers, included 50417 $Z \rightarrow ee$ events.

For the event generation Pythia 6.319 was employed including the CTEQ6L1 PDFs and the Tune-A [96] parameter set for the modelling of the underlying event. Sherpa 1.0.6 was used with CTEQ6L PDFs and default parameter settings for the underlying event. In the matrix-element calculation up to three partons in association with the Z boson were included and 20 GeV was chosen as matching scale between the matrix-element and parton shower domains. For both Monte Carlo samples zero bias events were overlaid to account for additional $p\bar{p}$ interactions in the same bunch crossing, pile-up, and noise. The samples were processed through a full DØ detector simulation based on Geant [123].

Z boson candidate events were selected by requiring two opposite charge electrons with $p_T > 25\,\mathrm{GeV}$, tight identification criteria, and invariant dielectron mass $70\,\mathrm{GeV} < M_{ee} < 100\,\mathrm{GeV}$, such that any background contributions are minimal. Jets are reconstructed based on the Run II cone algorithm with cone radius $R = 0.5$ and are required to have a transverse momentum $p_T > 15\,\mathrm{GeV}$.

Various kinematic and topological distributions were studied at reconstruction level to compare the measured data with the two Monte Carlo samples, including the transverse momenta of the Z boson and the three leading jets, the jet multiplicity and angular jet-jet correlations. As one example the p_T of the leading jet is plotted in Fig. 4.13. Hard parton emission (i.e. jet production at large p_T) can be well modelled with matrix-elements which include the radiated parton, whereas the parton shower approximation gives a good description at low p_T but underestimates the jet production at large p_T. Consequently, the parton shower event generator Pythia predicts a too soft p_T distribution for the leading jet, whereas Sherpa, which relies on the matrix-element at large p_T, is in good agreement with the data. A similar trend is seen in the p_T distributions of the second and third jet.

Both Sherpa and Pythia can model the angular separation (i.e. $\Delta\eta$ and $\Delta\phi$) between the two leading jets, with one limitation for Pythia since it predicts a peak at $\Delta\phi \approx \pi$ which is not observed in data. This enhancement for back-to-back jets

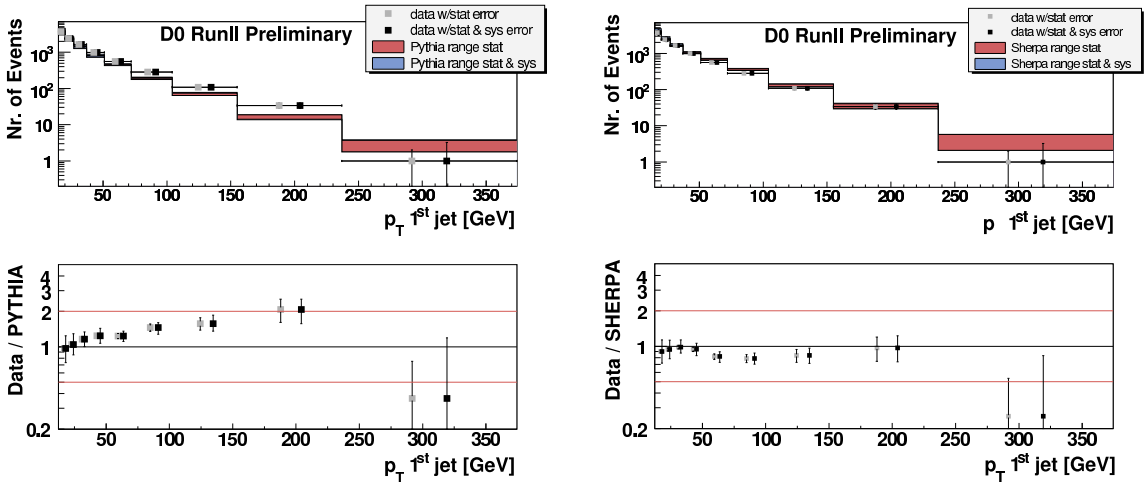

Figure 4.13: Transverse momentum distribution of the leading jet in $Z + \geq 1$ jet events compared to simulations obtained with Pythia *(left) and* Sherpa *(right) [182]. The lower plots show the ratio of the distributions (data/MC).*

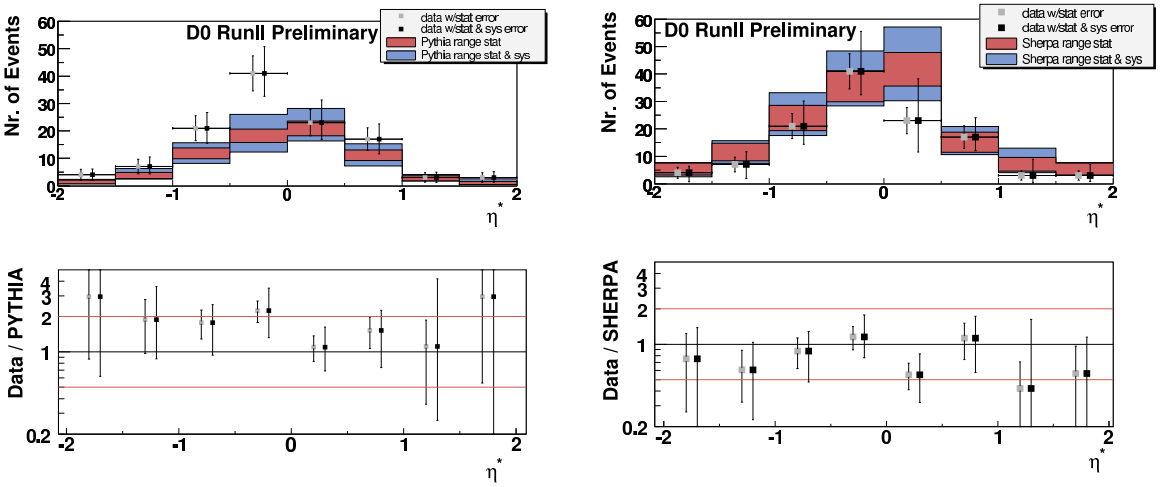

Figure 4.14: $\eta^* = \eta_3 - (\eta_1 + \eta_2)/2$ of third jet in $Z + \geq 3$ jet events compared to simulations obtained with Pythia *(left) and* Sherpa *(right) [182]. The lower plots show the ratio of the distributions (data/MC).*

is likely due the modelling of multi-parton interactions (underlying event) in Pythia, meaning that the additional jets at $\Delta\phi \approx \pi$ originate from an additional parton-parton interaction and not from the hard process. Since the minimal required jet transverse momentum is only $p_T = 15\,\text{GeV}$ the rate of jets produced in the underlying event is not negligible.

It is of particular interest to study the radiation of additional jets in Z + jets events with two leading jets which are well separated in rapidity as the rate for the production of a third enclosed jet is different for the QCD production process (colour-octet exchange) and the electroweak vector-boson-fusion (VBF, colour-singlet exchange) [183].

The electroweak production of Z + 2 jets via WW fusion has the same structure as Higgs boson production via VBF, which is expected to be one of the major Higgs discovery channels at the LHC [184, 185]. At the Tevatron the cross section for Z + 2 jet production via the VBF process is about two orders smaller than the dominating QCD process.

The DØ collaboration studied $Z+3$ jets events, with the two leading jets separated by $|\eta_1-\eta_2| > 2$ and the third jet radiated in between, i.e. $\eta_1 < \eta_3 < \eta_2$ or $\eta_2 < \eta_3 < \eta_1$. Fig. 4.14 shows the η distribution of the third jet relative to the two leading jets, defined as $\eta^* = \eta_3 - (\eta_1 + \eta_2)/2$. Due to the relatively small statistics for this class of events, the comparison between data and the Monte Carlo samples has only limited precision, but one can conclude that both Pythia and Sherpa can correctly model the topology of the third jet radiation. However, Pythia predicts a too low rate of three-jet events by a factor 1.7.

4.2 Heavy flavour jet production

The production of b quark jets in association with the weak bosons is of particular interest as these processes comprise the most important background to several new processes being searched for, e.g. single-top and associated Higgs production. b jets can be identified using characteristic properties of B mesons as semi-leptonic decays [52, 186], lifetime [51, 187], and mass [188, 189]. Also, the associated production of jets originating from c quarks is important, not only because those can be misidentified as b-tagged jets.

4.2.1 Predictions from perturbative QCD

Predictions for the production of a W or Z boson with a heavy-quark pair $Q\bar{Q}$ ($Q = c, b$) and possible additional light partons can be obtained using the leading-order event generators Alpgen [43] and Sherpa [60]. Both programs include a matching algorithm which allows to combine matrix-element calculations for different parton multiplicities and a parton shower simulation (cf. Section 2.3.5). The Alpgen code incorporates $WQ\bar{Q}$, $ZQ\bar{Q}$ (both with up to four additional light partons), and Wc (with up to five additional light partons). A case study on $Wb\bar{b} + n$ jets production was presented by the Alpgen authors in Ref. [89].

At Born level $Wb\bar{b}$ is only produced via gluon radiation off a quark and gluon splitting into a $b\bar{b}$-pair in the initial state, $q\bar{q}' \to Wb\bar{b}$. For $Zb\bar{b}$ production additional diagrams besides gluon radiation contribute: Z radiation off a b quark in the final state (included in $q\bar{q} \to Zb\bar{b}$) and gg scattering with a Z radiated off the b quark exchanged in the t-channel ($gg \to Zb\bar{b}$). The flavour excitation process $gs \to W^-c$ has no corresponding analogue involving a Z boson. These are all lowest order processes for the production of a W or Z boson with a heavy quark, if only light parton density functions are considered. Using heavy-quark PDFs, the production of a single heavy quark Q in association with a Z boson is allowed via $gQ \to ZQ$. Note, that this is only an alternative calculation scheme since this process is already included in the calculation of $gg \to ZQ\bar{Q}$ with one of the heavy quarks being emitted collinear to the

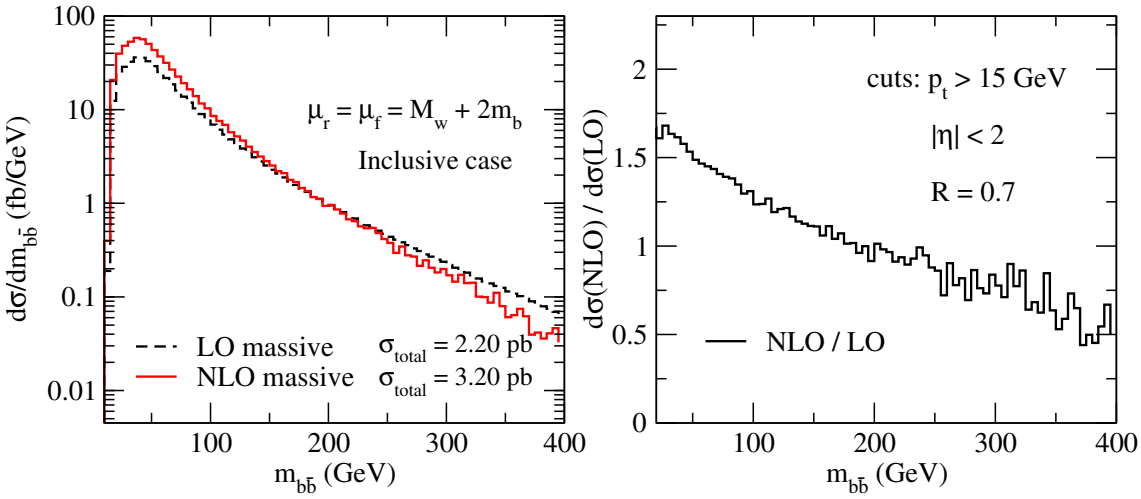

Figure 4.15: Left: The differential cross section $\mathrm{d}\sigma/\mathrm{d}m_{b\bar{b}}$ *as function of dijet mass for inclusive* $Wb\bar{b}$ *production at LO and NLO with massive* b *quarks. Right: The* K*-factor defined as the ratio of NLO and LO cross sections [66].*

beam and since the heavy flavour PDF is dynamically derived via $g \to Q\bar{Q}$ splittings in the DGLAP evolution equations [70–72]. Thus, one should be careful when combining different processes to avoid double-counting in the cross section. The calculation based on heavy-flavour PDFs has the advantage that the collinear logarithms in the initial state are absorbed in the PDF leading to a more convergent perturbative expansion [190, 191]. In addition NLO corrections are less complicated due to the less complex final states at LO [192, 193].

The next-to-leading order corrections to $Wb\bar{b}$ [64] and $Zb\bar{b}$ [65] production have been calculated in the approximation of massless b quarks and are implemented in the parton-level code MCFM [45, 63]. Recently, the NLO corrections to $Wb\bar{b}$ production were also derived for massive b quarks [66]. The calculation confirmed the expectation that the massive treatment is only of importance at low dijet mass $m_{b\bar{b}}$ and found that the difference between the massive and massless calculation is marginal for $m_{b\bar{b}} \gtrsim 50\,\mathrm{GeV}$.

Fig. 4.15 shows the LO and NLO predictions for the dijet mass distribution $\mathrm{d}\sigma/\mathrm{d}m_{b\bar{b}}$ for inclusive $Wb\bar{b}$ production using massive b quarks. At low dijet masses the ratio of NLO and LO cross sections (the K factor) is about 1.5, but decreases steadily with increasing $m_{b\bar{b}}$ so that it even falls below unity at $m_{b\bar{b}} \approx 175\,\mathrm{GeV}$. For a specific event selection the *effective* K-factor depends on the choice of the kinematic region for selected jets. Therefore care should be taken when using the displayed graphs to directly correct LO predictions (e.g. as obtained with event generators). Instead the MCFM program can be used to recalculate the NLO corrections to match specific kinematic selections.

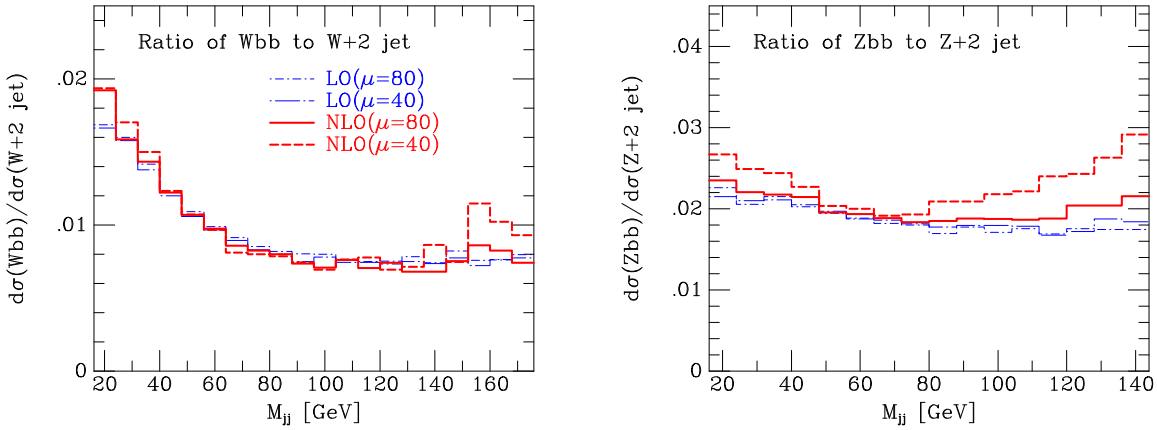

Figure 4.16: Ratio of W/Z + 2 b jets to W/Z + 2 jet events as function of the invariant dijet mass M_{jj} at LO and NLO for two different choices of the normalization and factorization scale μ [45]. The jets are required to have $p_T > 15\,\text{GeV}$ and $|y| < 2$.

The heavy-flavour content of jets produced with a W or Z boson was found not to depend strongly on the higher-order correction [45] which is demonstrated in Fig. 4.16. The ratio of the cross section of W events with two b jets over the cross section for W events with any two jets as a function of dijet mass m_{jj} is nearly constant at 0.8% for $m_{jj} \gtrsim 60\,\text{GeV}$. In this calculation jets where required to have $p_T > 15\,\text{GeV}$ and $|y| < 2$. The b quark content is increasing towards low m_{jj} as $b\bar{b}$ quark pairs are only produced via gluon splitting which involves a massless propagator term. The NLO and LO predictions for the $Wb\bar{b}/Wjj$ ratio are practically identical, except for low dijet masses, where the NLO correction increases the ratio. This prediction appears to be somewhat in contrast to the findings of the CDF and DØ collaborations, which obtain a better description of their data in the W + heavy-flavour-tagged jet sample when they reweight the LO prediction of the heavy-flavour content obtained with ALPGEN with a factor of about 1.5 [194, 195]. The predictions for the $Zb\bar{b}$ contribution to the Zjj cross section is considerably larger compared to $Wb\bar{b}$, in particular at large M_{jj} due to the additional diagrams contributing to $Zb\bar{b}$.

NLO corrections have also been calculated for Wc production [196], for the production of a single heavy-quark in association with a Z boson [192, 197] and for the production of a W boson and two associated jets with one b quark tag [193]. The latter two processes can be realized using heavy-flavour PDFs in the initial state. In addition, processes with two heavy-flavour quarks in the final state, e.g. $q\bar{q} \to Zb\bar{b}$, contribute to the one-tag rate, as one of the heavy quarks can be outside the selected range in p_T and y or both heavy quarks can collapse into a single jet. ZQ production is predicted to be dominated by the process $gQ \to ZQ$, with $q\bar{q} \to ZQ\bar{Q}$ contributing about a third at the Tevatron, while being much less significant at the LHC [192]. Ref. [193] showed that the process $bq \to Wbq'$ receives large NLO corrections thus that it is nearly as important as $Wb\bar{b}$ if only one b-tag is required. Note that $bq \to Wbq'$ is equivalent to the process $gq \to Wb\bar{b}q'$ which does not rely on b quarks in the initial

state and which is, for example, implemented in the ALPGEN event generator. At LO and using a renormalization and factorization scale of M_W the cross section calculated based on $gq \to Wb\bar{b}q'$ is about twice as large compared to the one obtained with $bq \to Wbq'$ [193]. Thus the higher order corrections to $gq \to Wb\bar{b}q'$ are expected to be of largely reduced size.

4.2.2 Measurements of b quark production in association with a Z or W boson

The measurement of the production of b quark jets in association with Z or W bosons is challenging as the production cross sections are relatively small and the background contributions are substantial. A crucial element of the analysis is a method to efficiently identify jets originating from a heavy-flavour quark, while effectively suppressing the rate of jets from light-flavour quarks or gluons which are misidentified as heavy-flavour jets.

Heavy flavour tagging

The b quarks produced in hard processes mostly hadronize into B mesons (and b-flavoured baryons), which have significant mean life times $\tau \approx 1.6 \times 10^{-12}\,\mathrm{s}$ and large semi-leptonic branching fractions of about 11% each for decays in final states with electrons or muons. Furthermore, B mesons can decay into D mesons which subsequently can decay semi-leptonically, thus increasing the effective rate of leptons from B decays. Both properties, lifetime and semi-leptonic decays, can be used to *tag* b quark jets. c quark jets which hadronize into D mesons can also be tagged albeit with a largely reduced tagging efficiency.

The identification of non-isolated muons within jets provides a *soft lepton tag* (SLT) for b jets. The b-tagging efficiency is limited by the semi-leptonic decay rates and by the muon identification efficiencies and thus typically reaches only values of about 11% [198]. This method was applied by both the CDF and DØ collaborations for the observation of top quark pair production at the Tevatron Run I [51, 52].

The decay of a long-lived b-flavoured hadron generates tracks attached to a secondary vertex which is well separated from the primary vertex of the hard interaction. The *secondary vertex tagger* (SVT) relies on high resolution tracking near the interaction point, which can be achieved with silicon tracking detectors. In CDF's top quark observation the SVT was the primary method to identify b jets due to a significantly higher efficiency as compared to the SLT. In Run II, both CDF and DØ employ SVT algorithms for b jet identification [187,194,199]. The details of the specific methods implemented by both experiments differ. To give a general idea of the method, reference to DØ's algorithm is given in the following, which is a bottom-up approach: Tracks are clustered into track-jets using a cone algorithm. For each track-jet (consisting of at least two tracks) any pair of tracks with a significant impact parameter relative to the primary vertex (distance of closest approach) is used as a seed for a secondary vertex. Additional tracks, which are consistent with originating from this vertex, are attached iteratively. A vertex transverse decay length L_{xy} is defined as the radial distance of the secondary vertex to the primary vertex in the plane perpendicular to the beam

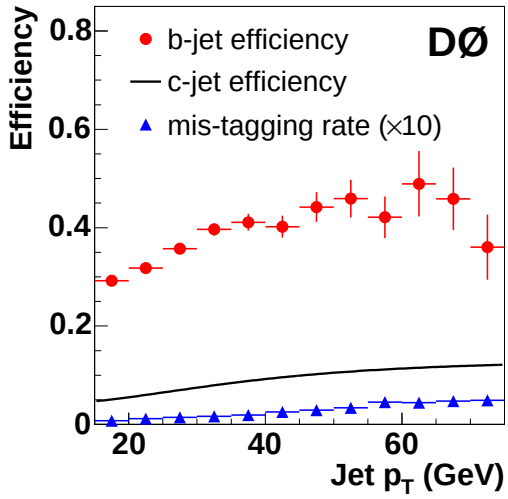

Figure 4.17: Tagging efficiencies for b and c jets and mistag rate for light flavour jets of DØ's secondary vertex tag algorithm [187].

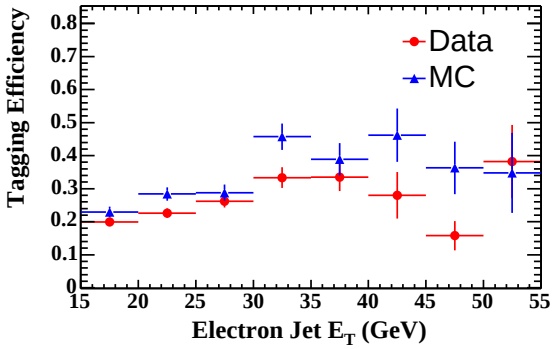

Figure 4.18: b-tagging efficiencies of CDF's secondary vertex tag algorithm measured in data and simulation [194]. The jet is identified as heavy-flavour jet with a soft electron tag.

pipe and is required to fulfil $L_{xy}/\sigma_{xy} > 7$ with σ_{xy} being the corresponding resolution. Examples for measured b-tagging efficiencies based on secondary vertex taggers are given in Figs. 4.17 and 4.18.

Both collaborations also employ jet lifetime probability taggers which estimate the combined probability of tracks within a jet to originate from the primary vertex [200, 201]. Small probabilities are an indication of b jets which contain tracks with large impact parameters. To further optimize b-tagging efficiencies and mistag rates, both DØ and CDF recently developed taggers based on neural networks which combine the outputs of the traditional taggers and additional kinematic variables [202, 203].

Jets originating from b or c quarks can be discriminated using the invariant mass M_s of charged tracks attached to a secondary vertex [188, 189]. Due to the large mass

of b hadrons, b jets have on average a larger M_s than c or light jets. The fraction of b, c and light jets in the secondary vertex tagged sample can be derived from fitting the observed M_s distribution with the sum of template shapes for the different flavours.

Zb production

During Run II both the DØ and CDF collaborations have measured the production of b jets in association with a Z boson. Inclusive $Z + b$ jet production which can contain an additional jet in the final state is expected to be a major background to associated Higgs production, ZH, with subsequent Higgs-boson decay into a $b\bar{b}$-pair. As about two thirds of the inclusive $Z + b$ jet production is expected to originate from the parton process $gb \to Zb$ (when using a five-flavour scheme with a b quark PDF) a measurement of the Zb production cross section would provide a constraint on the b quark PDF. Presently the b quark distribution function is evolved dynamically from the gluon PDF [94, 124, 190, 191] and found to be in agreement with the b contribution to the proton structure function $F_2^{b\bar{b}}$ measured in deep inelastic ep scattering up to momentum transfers $Q^2 = 650\,\mathrm{GeV}^2$ [204, 205]. The measurement of Zb production is sensitive to the b quark PDF at significantly higher $Q^2 \approx M_Z^2$. A precise knowledge of the b distribution function is essential when predicting the production rates for processes with b partons in the initial state, e.g. single top production [206] and the production of a supersymmetric Higgs h via $gb \to hb$ or $bb \to h$ [207–211].

The DØ collaboration published the first measurement of the ratio of inclusive cross sections $\sigma(p\bar{p} \to Z + b\,\mathrm{jet})/\sigma(p\bar{p} \to Z + \mathrm{jet})$ based on an integrated luminosity of $180\,\mathrm{pb}^{-1}$ [212]. Both $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ decay channels were utilized.

Jets were reconstructed using the Run II cone algorithm with $R = 0.5$ and were required to have $p_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$. The b jets were identified by secondary vertices separated from the primary vertex of the hard interaction. The primary vertex could be distinguished from additional vertices originating from minimum-bias events occurring in the same bunch crossing based on the p_T distribution of the associated tracks. Jets were considered taggable if a track jet consisting of at least two tracks was found within $\Delta R = 0.5$. The taggability for light jets was measured in data and determined to be 75%. Monte Carlo studies showed that the taggabilities for b and c jets are about 80%.

A secondary vertex tag algorithm was applied to each taggable jet. The b-tagging efficiency was determined from a data sample containing jets with muons (soft muon tag) and parameterized as function of jet p_T and η. The difference in b tagging efficiency between jets with and without muon was estimated using simulation. The average b-tagging efficiency was determined to be $\epsilon_b \approx 33\%$. The c-tagging efficiency was estimated assuming the ratio $\epsilon_c/\epsilon_b = 0.266$ as measured in simulation. The mistag rate for light-flavours was determined from data and averages $\epsilon_l \approx 0.28\%$.

Based on these tag efficiencies and taggabilities the number of events with a b, c, or light jet, N_b, N_c, and N_l, respectively, were derived from the number of events before and after b-tagging, corrected for the background contribution from multijet and Drell-Yan continuum production. As one constraint is missing to determine the three unknown quantities N_b, N_c, and N_l, the ratio of the number of events with c and b quarks were taken from the theoretical prediction $N_c = 1.69 N_b$ [192].

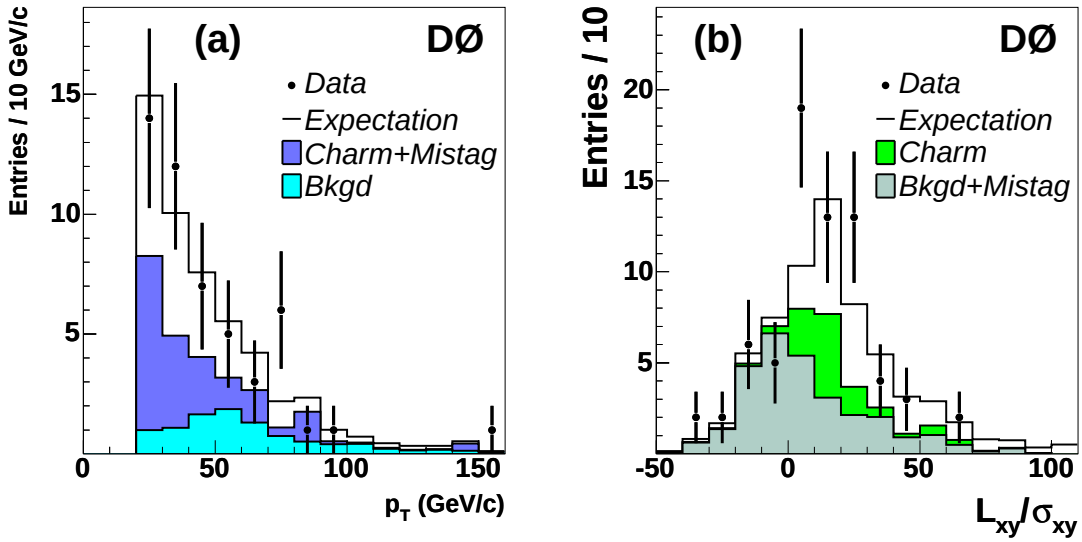

Figure 4.19: (a) The transverse momentum spectrum for b-tagged jets in Z + jets events. (b) Distribution in decay-length significance L_{xy}/σ_{xy} of secondary vertices in the transverse plane [212].

In Fig. 4.19 the shapes of the transverse momentum distribution of b-tagged jets and of the decay-length significance of the secondary vertices are compared to the sum of the expected Zb signal and the background contributions. The distribution of the decay-length significance shows a clear evidence for the heavy-flavour component.

The dominating systematic uncertainties for the measurement of the inclusive cross section ratio $\sigma(p\bar{p} \to Z + b\,\text{jet})/\sigma(p\bar{p} \to Z + \text{jet})$ were found to be uncertainties in the jet energy scale for b and light jets, in the background estimate, in the difference in tagging efficiency for $ZQ\bar{Q}$ events with the two heavy quarks being either combined in one jet or separated in two jets, and in the tagging efficiencies for b, c, and light jets. In addition an uncertainty on the theoretical prediction for $\sigma(Zc)/\sigma(Zb)$ [192] was considered.

The inclusive cross section ratio $\sigma(p\bar{p} \to Z+b\,\text{jet})/\sigma(p\bar{p} \to Z+\text{jet})$ was measured to be $0.021 \pm 0.004(\text{stat})^{+0.002}_{-0.003}(\text{syst})$. Here, $\sigma(p\bar{p} \to Z+\text{jet})$ includes the rate for any quark. The measurement is in good agreement with the NLO prediction of 0.018 ± 0.004 [192] obtained with the CTEQ6M parton distributions [94].

CDF's measurement of the $Z + b$ jet cross section was the first one not to rely on a theoretical prediction for the ratio $\sigma(Zc)/\sigma(Zb)$, but instead was able to discriminate b and c jets using the invariant mass distribution of charged particles originating from the secondary vertex [213]. This measurement is based on an integrated luminosity of $330\,\text{pb}^{-1}$ and includes both the electron and muon decay channels of the Z boson. The trigger and lepton selection were adopted from CDF's inclusive Z cross section measurement [119]. Jets were reconstructed using a cone algorithm with cone size $R = 0.7$ and were required to have $E_T > 20\,\text{GeV}$ and $|\eta| < 1.5$. b jets are identified using a secondary vertex algorithm [194] with an average efficiency of 33%. The b-tagging efficiency was derived from a sample of dijet events with a reconstructed b

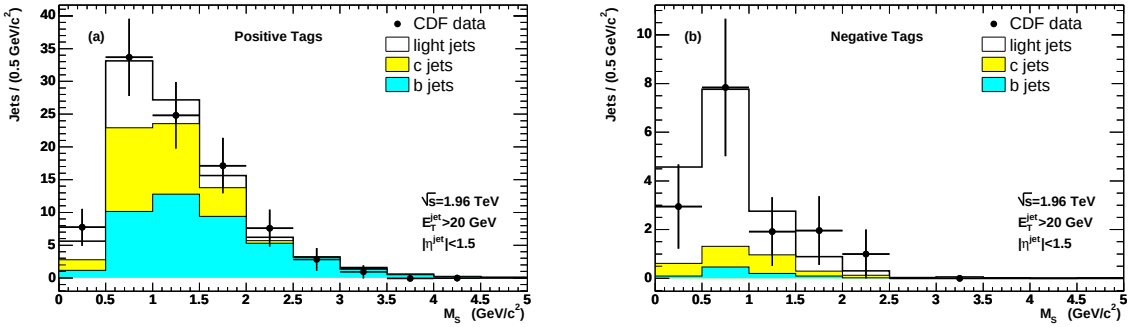

Figure 4.20: The invariant mass of all charged tracks attached to the secondary vertex, M_s, in Z + jets events [213]. The jets are required to have a positive (left) or negative tag (right) defined as a significant transverse displacement of the secondary vertex pointing into the direction (positive tag) or opposite to the direction (negative tag) of the jet hemisphere. The non-Drell-Yan background has been subtracted from the data. The data are compared with the sum of the light, c and b Monte Carlo templates, of which the normalization has been obtained by a fit to the observed M_s distribution.

or c decay. Non-negligible backgrounds were from ZZ and $t\bar{t}$ production, which were estimated using Monte Carlo simulations, and from multijet production with jets faking isolated electrons or muons. The lepton fake rates were estimated from data.

To discriminate between c and b quark jets in the tagged sample and to estimate the fraction of b jets, a fit to the invariant mass of all tracks attached to the secondary vertex M_s was performed. To better discriminate between the contributions from the different quark flavours, the M_s distribution for jets with both positive and negative tag, defined as a significant transverse displacement of the secondary vertex pointing into the direction or opposite to the direction of the jet hemisphere, were included in the fit. In Fig. 4.20 the M_s distributions for both positively and negatively tagged jets are compared to the sum of the scaled Monte Carlo templates for b, c and light quarks. The additional discrimination, in particular between light and c quarks, due to the inclusion of the M_s distribution for negative tags is evident. The template for b jets was validated with a dijet sample which was also used to determine the b-tagging efficiency. Using γ + jet data it was found that the simulation did not accurately describe the ratio of positive to negative tags for light jets, which was predicted a factor 1.3 too high. The light jet template was reweighted accordingly and the difference from unity was taken as a systematic uncertainty.

The main systematic uncertainty in this measurement of the Zb production cross section is the uncertainty of the simulated jet E_T and η distributions, which define the acceptance correction. The uncertainty was estimated by reweighting the Monte Carlo distributions within a range consistent with the data. Other dominant contributions to the systematic error include uncertainties on the track reconstruction efficiency, on tagging efficiencies and on the jet energy scale.

The measured cross section for Z + b jet production is evaluated for jet $E_T >$

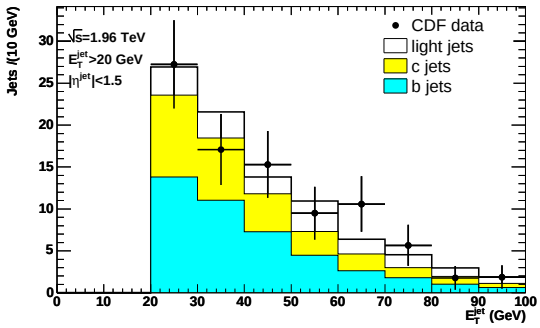

Figure 4.21: The transverse energy distribution for positively tagged jets in Z + jets events [213]. The non-Drell-Yan background has been subtracted from the data. The data are compared with the sum of the light, c and b quark Monte Carlo templates.

20 GeV, $|\eta| < 1.5$ and 66 GeV $< M_{ll} <$ 116 GeV and presented both as absolute and relative cross section. The $Z + b$ jet production cross section normalized to inclusive Z production is measured to be $\sigma(Z+b\,\text{jet})/\sigma(Z) = 0.0037 \pm 0.0011(\text{stat.}) \pm 0.0008(\text{syst.})$, which is in agreement with both the PYTHIA [122] prediction of 0.0035 and the NLO calculation of 0.0019 [192]. The rate of $Z + b$ jet compared to $Z +$ jet production was determined to be $\sigma(Z+b\,\text{jet})/\sigma(Z+\text{jet}) = 0.0236 \pm 0.0074(\text{stat.}) \pm 0.0053(\text{syst.})$, again in agreement with the PYTHIA estimate of 0.0218 and the NLO prediction of 0.0181. This CDF measurement agrees also well with the preceding DØ measurement described above. Fig. 4.21 demonstrates that the observed transverse energy distribution of tagged jets is well described by the sum of the MC templates of b, c and light jets, which were scaled using the result of the fit to the M_s distribution.

$Wb\bar{b}$ production

The DØ collaboration published a combined search for $Wb\bar{b}$ and WH production based on the W decay in electrons and an integrated luminosity of 174 pb^{-1} [214]. This analysis, which was optimized to maximize the sensitivity for WH, selected events with a reconstructed electron, significant missing transverse momentum and two jets with $p_T > 20$ GeV and $|\eta| < 2.5$. Both jets were required to be tagged with an algorithm based on jet lifetime probabilities and to have an angular separation $\Delta R_{bb} > 0.75$. The angular separation reduces the presence of b jets originating from gluon splitting which enhances the sensitivity for WH, but suppresses the selection efficiency for $Wb\bar{b}$. An upper cross section limit at 95% C.L. on $Wb\bar{b}$ production of 6.6 pb was set for $p_T^b > 20$ GeV and $\Delta R_{bb} > 0.75$, compared to the NLO prediction of 0.75 pb obtained with MCFM [45].

The CDF collaboration presented a preliminary measurement of the b jet cross section for $Wb\bar{b}$ production based on an integrated luminosity of 695 pb^{-1} and using the W decays in both electrons and muons [215]. In addition to a reconstructed electron or muon and missing transverse momentum, one or two jets with $E_T > 20$ GeV and $|\eta| \leq 2$ were required. At least one jet needed to have a secondary vertex tag and

the flavour composition of the tagged jets was determined using the invariant mass distribution M_s of the tracks attached to the secondary vertex similar to CDF's $Z + b$ jet analysis.

About 33% of the tagged jets are found to be true b jets, while c and light jets contribute about 47% and 20%, respectively. In contrast to the measurements of Zb production, the $W + b$ jet selection suffers from large background contributions, both from top and QCD multijet production. The top background, which has similarly large contributions from single and pair production, was estimated from simulation to be 30%. The multijet background was evaluated using data samples with inverted $\not\!\!E_{\mathrm{T}}$ and lepton isolation requirements, while its b jet contribution was determined from the M_s distribution at low E_T. The background from QCD production was found to account for 25% of the selected b-tagged jets.

The dominant sources of the systematic uncertainty were due to the Monte Carlo modelling of the templates for the M_s distribution and due to the uncertainties on the determination of the multijet background and on the b-tagging efficiency. The acceptance for the cross-section measurement was derived using simulated samples for $W + b\bar{b} + n$ additional partons (with $n = 0, 1, 2$) generated with ALPGEN [43].

The measurement does not derive a cross section for $Wb\bar{b}$ production as such, but presents the result as a b jet cross section in $Wb\bar{b}$ events, i.e. the measured cross section is proportional to the number of b jets in events with a leptonically decaying W boson and one or two jets. The cross section, which is corrected to the limited phase space corresponding to the kinematic acceptance and selection (i.e. $p_T(\ell^{\pm}) \geq 20\,\mathrm{GeV}$, $|\eta(\ell^{\pm})| \leq 1.1$, $p_T(\nu) \geq 25\,\mathrm{GeV}$, $E_T(\mathrm{jet}) > 20\,\mathrm{GeV}$, $|\eta(\mathrm{jet})| \leq 2$), was measured to be $0.90 \pm 0.20(\mathrm{stat.}) \pm 0.26(\mathrm{syst.})\,\mathrm{pb}$, which is in good agreement with the LO Alpgen prediction of $0.74 \pm 0.18\,\mathrm{pb}$.

In Run I, the CDF collaboration reported an anomaly observed in the heavy-flavour content of jets produced in association with a W boson, such that an excess of jets was observed which were tagged both by a secondary vertex tagging and a soft-lepton tagging algorithm [216]. It was suggested that this enhanced rate could be due to the production of the supersymmetric partner of the b quark, which could be relatively light [217]. In this scenario, the semi-leptonic decays of supersymmetric $\tilde{B}$ mesons would be the source of additional soft leptons.

Both, the DØ and CDF collaborations searched for this anomalous heavy-flavour production in association with a W boson based on data sets with integrated luminosities of about $150\,\mathrm{pb}^{-1}$ collected in Run II [198, 218]. Both analyses found good agreement with the standard model prediction of $Wb\bar{b}$ production and set upper limits on the cross section for new hypothetical processes.

4.3 Conclusions for new phenomena searches

The associated production of vector bosons with jets is the dominating background in many searches for new phenomena and the Higgs boson, which will be discussed in detail in the following chapter. Thus, precise measurements of W/Z + jet production as well as simulations of these processes are essential for the accurate prediction of their background contribution in searches.

Measurements of differential jet distributions in W and Z boson events are challenging as they require a detailed understanding of the jet energy calibration and energy resolution. They have sizable statistical uncertainties at large jet transverse momenta or high jet multiplicities and substantial systematic uncertainties arising from, e.g. the jet energy scale calibration or the estimation of the background. The measurements are in general well described by the prediction of NLO QCD calculations, where available, and by those event generators, which consistently combine matrix-element calculations up to high parton multiplicities with parton shower models, e.g. ALPGEN and SHERPA.

Of particular relevance is the associated production of W and Z bosons with heavy-flavour jets, which is the most important background in searches for the Higgs boson produced in associated WH and ZH production. These processes obtain sizable NLO corrections which modify both integral rates and shapes of kinematic distributions. Predictions at LO differ considerably and depend on the calculation scheme applied. Measurements of heavy-flavour jet production in association with vector bosons, are both limited by statistical and systematic uncertainties. Compared to flavour-inclusive jet production, additional systematic uncertainties are due to the heavy-flavour identification. They can be substantial in measurements which discriminate between b and c quark jets. As a consequence of both theoretical and experimental uncertainties, predictions for the production of $Wb\bar{b}$, $Zb\bar{b}$, and other W/Z + heavy flavour processes are less precise.

Chapter 5

Background from W/Z + jets production in searches for new phenomena and the Higgs boson

The production of W and Z bosons at hadron colliders is a copious source for both charged leptons and neutrinos. Many rare standard model or new phenomena processes result in event topologies which include charged leptons and missing transverse energy $\not\!\!E_{\mathrm{T}}$. Of particular interest are final states which include one or more hadron jets in addition to leptons and/or missing transverse momentum as they are generated in the production and decay of many new particles, e.g. the associated production of the Higgs boson with a W or Z boson and subsequent decay $H \to b\bar{b}$, the pair-production of scalar quarks and gluinos in supersymmetric models with decays into quarks, gluons and the lightest neutralino, or the pair-production of leptoquarks decaying into leptons and quarks etc.

In searches for these new processes, the production of the weak bosons with associated jets is a main background source. In many cases, it is even the dominating background contribution. Thus, it is essential both to correctly estimate the background contribution from W/Z + jets production and to understand how to effectively discriminate the signal of a new particle from this background. Additional background contributions from both multijet production and top-production are often important as well. On the other hand W + jet production is the dominating background when measuring single top and top pair production. The discussion of the latter is outside the scope of this work.

The previous chapters discussed in detail the phenomenology of vector boson production as well as measurements of inclusive and differential cross sections. In particular a focus on the production of hadron jets in association with the weak bosons and on the description and validation of perturbative QCD calculations and event generators was set. Based on this groundwork, this chapter discusses W/Z + jets production as a primary background source in searches. Using several searches performed by both the CDF and DØ collaborations as examples, this chapter reviews how this background can be estimated and effectively suppressed. First, new processes with charged leptons (and possible $\not\!\!E_{\mathrm{T}}$) in the final state are discussed, then those processes with missing transverse momentum (but without charged leptons). Both cases are further classified

dependent on the presence of associated heavy-flavour jets. The primary interest is not to discuss the phenomenological details of Higgs boson production or of particular models for new physics beyond the standard model, but instead to focus on the role of the Z/W + jet background. In most of the cases, reference to analyses performed by the DØ collaboration is given when discussing exemplary searches below. This choice is not based on any valuation, but is motivated by an attempt to avoid redundant discussions, as there is a large overlap between the analyses performed by both Tevatron collaborations.

5.1 Final states with charged leptons and jets in searches

The associated production of vector bosons with jets with subsequent leptonic decays, $Z \to \ell^+\ell^-$ and $W \to \ell\nu$ generates final states which include jets, charged leptons and possibly additional missing transverse energy. Similar topologies are predicted by the production and decay of excited quarks, leptoquarks and the Higgs boson (in WH and ZH production). For the latter the identification of b jets is essential for the signal discrimination.

5.1.1 Jets without heavy flavour identification

Several exemplary searches for excited quarks and leptoquarks are reviewed below.

Excited quarks: $q^* \to Zq$

The DØ collaboration searched for a heavy resonance decaying into a $Z \to e^+e^-$+ jet final state using an integrated luminosity of $370\,\mathrm{pb}^{-1}$ [219]. While the search was performed for a generic resonance, the non-observation of a signal was more specifically interpreted in terms of a model which includes an excited quark decaying into a quark and a Z boson, which would indicate quark substructure [220]. This model is implemented in the Pythia event generator [36], which was used to evaluate signal acceptances. The event selection required a reconstructed electron pair with an invariant mass consistent with the Z boson and at least one jet with $p_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$.

As a consequence of the high purity of the Z boson reconstruction, the background to the hypothetical heavy resonance is completely dominated by Z + jet production. The instrumental background due to QCD multijet production was estimated to account for only 0.6% of the rate and the contribution from W + jet production was found to be negligible. The Z + jet background was estimated using the parton shower event generator Pythia. Simulated samples of the $2 \to 1$ process $q\bar{q} \to Z/\gamma^*$, which includes corrections to the first branching in the parton shower (cf. Section 2.3.5), were found to agree with the data, but to be of insufficient statistics for Z jet invariant masses $M_{Zj1} > 300\,\mathrm{GeV}$, i.e. the region where a hypothetical signal would be expected. In addition, the higher-order $2 \to 2$ processes, $q\bar{q} \to Zg$ and $qg \to Zq$, were simulated using minimal thresholds on the Z-parton invariant mass M_{Zq} ranging from 100 GeV to 400 GeV. Fig. 5.1 shows the measured distribution of the invariant mass of the Z

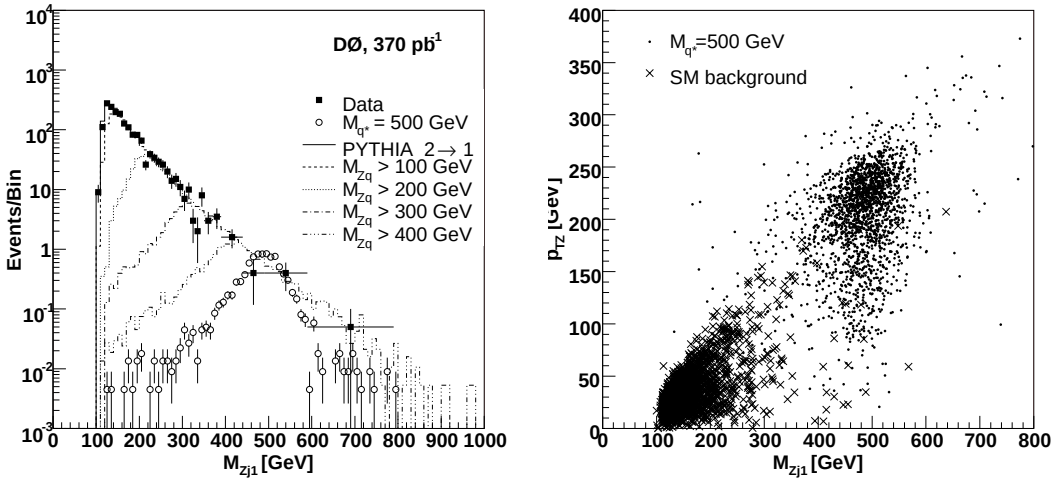

Figure 5.1: Search for heavy resonances decaying into Z + jet [219]: Left: the distribution of invariant mass of the reconstructed Z boson and the leading jet, M_{Zj1}, compared to the SM prediction and the hypothetical signal due to an excited quark q^* with mass $m_{q^*} = 500\,\text{GeV}$ and narrow width. The Z + jet background was simulated using PYTHIA*: The inclusive $2 \to 1$ process is shown as the solid line. The $2 \to 2$ process is shown for various M_{Zj} thresholds as discontinuous lines (only drawn up to where the distribution for a higher threshold on M_{Zj} takes over). Right: the $p_T(Z)$ vs. M_{Zj1} distribution for the standard model background and the hypothetical q^* signal.*

boson and the leading jet compared to the simulation which is in good agreement. The $2 \to 2$ samples corresponding to different thresholds on M_{Zq} were not combined but for each assumed heavy resonance mass a single sample was chosen which matches best the M_{Zj1} region to be investigated.

The two dimensional distribution given by M_{Zj1} and the reconstructed transverse momentum of the Z boson, p_{TZ}, was used to discriminate between the hypothetical signal of a resonance and the Z + jet background (see Fig. 5.1, right) as events from the resonance are clustered around its mass for M_{Zj1} and around half the mass for p_{TZ}. An elliptical cut around the central values of M_{Zj1} and p_{TZ} expected for a resonance with a certain assumed mass was chosen in a way to optimize the sensitivity. As no excess over background was observed, a 95% C.L. lower mass limit of 510 GeV for the excited quark mass M_{q^*} (with compositness scale $\Lambda = M_{q^*}$ which implies a narrow width of q^*) was set. The main contributions to the systematic uncertainty on the Z + jet background determination were found to be uncertainties in the jet energy scale and in the modelling of Z + jet production which was estimated using the difference between the predictions obtained with the PYTHIA and ALPGEN [43] event generators.

Leptoquark pair-production: $LQ\overline{LQ} \to \ell q \ell q$, $\to \ell q \nu q$

Leptoquarks, hypothetical coloured bosons which carry both lepton and quark quantum numbers and thus allow lepton-quark transitions, are predicted by numerous extensions

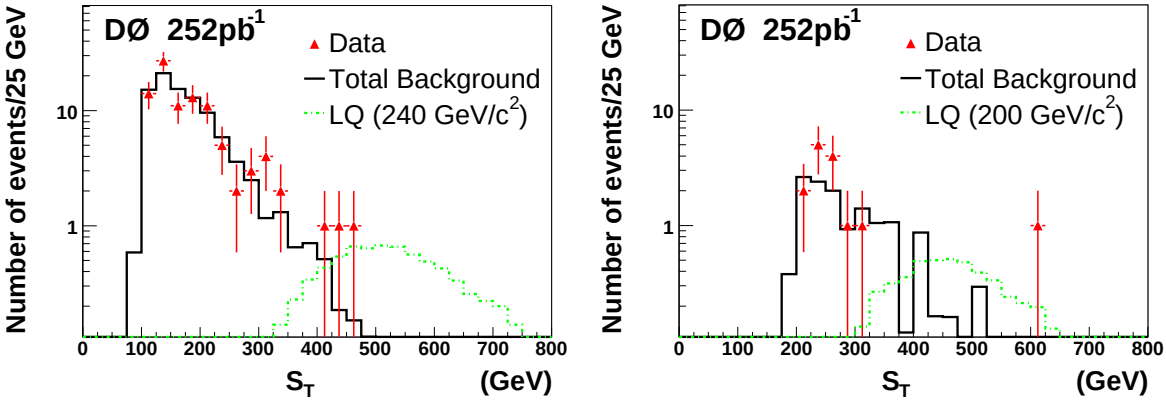

Figure 5.2: Search for first generation leptoquark pair production [226]: The distribution of the scalar sum of transverse energies, S_T, for $eejj$ events (left) and $e\nu jj$ events (right) after preselection and an additional cut on either the dielectron mass M_{ee} (l.) or transverse mass $M_T^{e\nu}$ (r.) to suppress the Z and W background. The data is compared with the SM prediction which is dominantly Z/γ^* + jet and QCD multijet production (l.) or W + jet production (r.), respectively. Also shown are the expected distributions for a 240 GeV LQ signal (l.) and for a 200 GeV LQ signal (r.).

of the standard model [221–225]. At hadron colliders, leptoquarks are predominantly produced in pairs via the strong coupling. They could, in principle, decay into any combination of a quark and a lepton, but leptoquarks with masses as low as $\mathcal{O}(100\,\mathrm{GeV})$ are only allowed to couple to one generation of quarks and leptons, since they otherwise would generate lepton number violation and sizable flavour-changing neutral currents. The branching fractions of the leptoquark decays into a charged lepton and quark or neutrino and quark are determined by the respective leptoquark-lepton-quark coupling. Thus, leptoquark pair-production can produce three characteristic final states: $\ell^+\ell^- qq$, $\ell^\pm \nu qq$, and $\nu\nu qq$. Main standard model backgrounds to leptoquark signals are therefore $Z/\gamma^*(\to \ell^+\ell^-)jj$, $W^\pm(\to \ell^\pm\nu)jj$, and $Z(\to \nu\bar{\nu})$, respectively. Here, some of the searches in the $\ell^+\ell^- jj$ and $\ell^\pm\nu jj$ final states are reported, whereas the $\nu\nu jj$ final state will be discussed in the following section.

During Tevatron Run II, the DØ and CDF collaborations searched for first and second generation leptoquark pair-production in both $\ell^+\ell^- jj$ and $\ell^\pm\nu jj$ final states [226–230]. Both experiments also reported preliminary results on the search for third generation leptoquark pair-production decaying into $\tau^+\tau^- jj$ [231, 232].

The DØ collaboration searched for first generation leptoquarks (LQ_1) in topologies arising from both $LQ_1\overline{LQ}_1 \to eqeq$ and $LQ_1\overline{LQ}_1 \to eq\nu q$ based on a data set corresponding to an integrated luminosity of $252\,\mathrm{pb}^{-1}$ [226]. The event selection required two high E_T electrons or one electron and significant $\not\!\!E_\mathrm{T}$ and at least two reconstructed jets with $E_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$. The background due to Z/γ^*, W, and $t\bar{t}$ production was estimated using ALPGEN (interfaced to PYTHIA) and PYTHIA, whilst the contribution of QCD multijet production was derived from data.

For the $eejj$ channel, most of the Z/γ^* + jet background could be suppressed using

a veto on events with a dielectron mass compatible with the Z boson mass, thus that subsequently $Z/\gamma^* +$ jets and multijet production were of an equal rate, whilst the $t\bar{t}$ contribution was an order of magnitude less. As the leptoquark decay products have large transverse energies due to the high mass of the leptoquark, the scalar sum of the transverse energies of the final state objects,

$$S_T = E_T(\ell_1) + E_T(\ell_2) + E_T(j_1) + E_T(j_2) \ , \tag{5.1}$$

was used as the final discriminant between signal and background, which is demonstrated in Fig. 5.2. Requiring $S_T > 450\,\text{GeV}$, which maximized the sensitivity of the search, the standard model background was reduced to 0.54 ± 0.11 events, dominated by both $Z/\gamma^* +$ jets and multijet production. The systematic uncertainty on the background contribution was dominated by uncertainties on the jet energy scale and on the modelling of the $Z/\gamma^* +$ jets background, which was estimated from the difference in the simulated samples obtained with ALPGEN and PYTHIA, respectively.

For the $e\nu jj$ channel, the $W +$ jet background was largely suppressed by requiring an invariant transverse mass of the electron and the missing energy well above the W boson mass: $M_T^{e\nu} > 130\,\text{GeV}$. Similar to the $eejj$ channel the background was discriminated from the hypothetical leptoquark signal using the scalar sum of transverse energies $S_T = E_T(e) + \not\!\!E_T + E_T(j_1) + E_T(j_2)$, for which an optimal cut of $S_T > 330\,\text{GeV}$ was found (see Fig. 5.2). Finally, 3.6 ± 1.2 background events were left after the selection of which about two thirds were due to $W +$ jets production. Also in this case the systematic uncertainty was dominated by uncertainties in the jet energy calibration and in the modelling of the $W/Z +$ jet background.

Combining both channels and the searches for leptoquarks performed in Run I [233], lower limits on the mass of first generation scalar leptoquarks as function of the branching fraction $\beta = Br(LQ_1 \to l^{\pm}q)$ were derived. The 95% C.L. limits on the leptoquark mass were determined to be $M_{LQ_1} > 256\,\text{GeV}$ for $\beta = 1$ and $M_{LQ_1} > 234\,\text{GeV}$ for $\beta = 0.5$.

In the searches for second generation leptoquarks (LQ_2), the separation of the background due to $Z/W +$ jet production is more difficult, as the resolution in muon transverse momentum, which is obtained from reconstructed tracks, is worse than the momentum resolution for electrons which benefits from the precise energy measurement in the calorimeter. In particular, due to tails in the muon momentum resolution, background events from $Z/W +$ jet production can be less effectively removed using the invariant dimuon mass $M_{\mu\mu}$ or the transverse mass $M_T^{\mu\nu}$.

The DØ collaboration published a search for $LQ_2\overline{LQ_2}$ pair-production in the final state consisting of two muons with $p_T > 15\,\text{GeV}$ and at least two jets with $E_T > 25\,\text{GeV}$ and $|\eta| < 2.4$ [228]. To account for the decreasing muon momentum resolution at large p_T, corrections were applied to the muon momenta, which exploit the transverse momentum balance for both signal and background processes. The missing transverse momentum $\not\!\!E_T$ was estimated from the reconstructed momenta of the two muons and the jets, and the momentum of the muon most opposite to the $\not\!\!E_T$ direction was scaled such that the $\not\!\!E_T$ component parallel to the muon direction vanished. Whilst this correction degraded the resolution in invariant mass $M_{\mu\mu}$, the tails of the background distribution at high $M_{\mu\mu}$ which leak into the signal region were considerably suppressed. After requiring $M_{\mu\mu} > 105\,\text{GeV}$, six events were selected in a data sample of $300\,\text{pb}^{-1}$

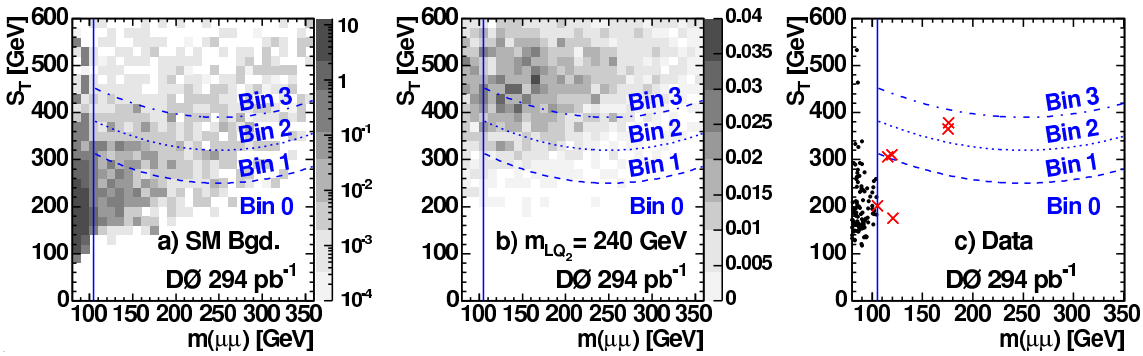

Figure 5.3: Search for second generation leptoquark production in the $\mu\mu jj$ final state [228]: Scalar sum of the transverse energies, S_T , as a function of the dimuon mass $m(\mu\mu)$: for the SM background (left), for a leptoquark signal with mass $m_{LQ_2} = 240\,\mathrm{GeV}/c^2$ and $LQ_2 \to \mu q$ branching fraction $\beta = 1$ (middle), and for data (right). The vertical line illustrates the Z boson veto and the curved lines show the boundaries between the signal bins used in the limit calculation.

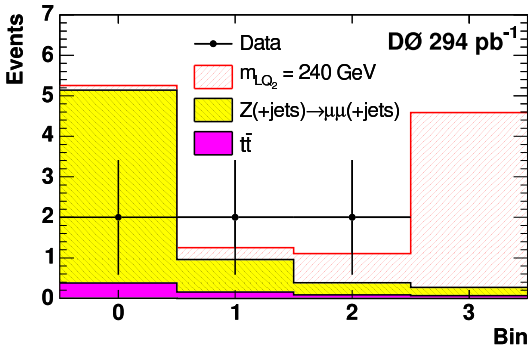

Figure 5.4: Distribution of events over the four bins in the $(m(\mu\mu), S_T)$ plane as defined in Fig. 5.3 compared to the SM prediction and the expected signal for a scalar leptoquark with mass $m_{LQ_2} = 240\,\mathrm{GeV}$ and $\beta = 1$ [228].

compared to 6.8 ± 2.0 events expected from standard model backgrounds, mainly consisting of Z/γ^* + jet production (6.1 ± 2.0 events).

The corrected dimuon mass $M_{\mu\mu}$ and the scalar sum of transverse energies S_T (cf. Eq. (5.1)) were used to further discriminate the background from a hypothetical signal. Fig. 5.3 shows the two-dimensional distribution in S_T vs. $M_{\mu\mu}$ for the standard model background, the leptoquark signal, and the observed events. As indicated in the figure, four bins were defined according to the signal-over-background ratio (S/B). The number of events in these four signal bins are given in Fig. 5.4, demonstrating that a hypothetical leptoquark signal with M_{LQ_2} could be well separated from the Z/γ^* + jets background. The main uncertainties on the predicted number of background events were determined to originate from the limited available statistics of the Monte Carlo samples and from uncertainties in the jet energy calibration and the Z/γ^* +

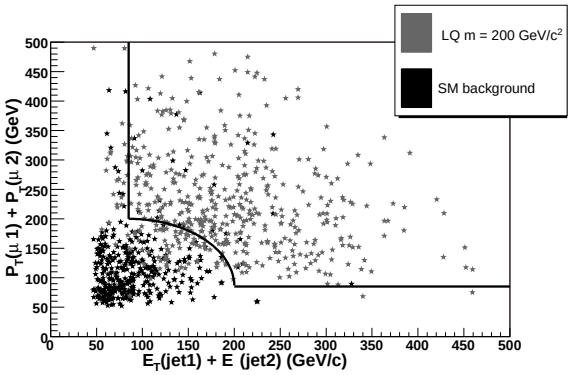

Figure 5.5: Search for second generation leptoquark production (CDF) [230]: $p_T(\mu_1) + p_T(\mu_2)$ vs. $E_T(jet_1) + E_T(jet_2)$ for simulated SM background and a hypothetical leptoquark signal with $M_{LQ_2} = 200\,\text{GeV}$. The line shows the main selection cut.

jet background prediction. The latter covers uncertainties on the normalization and on the shape of the jet p_T distribution. The event distribution in the four signal bins was used to derive upper cross-section limits on $LQ_2\overline{LQ_2}$ pair-production as function of the leptoquark mass M_{LQ_2}, which, in comparison with the NLO prediction for the $LQ_2\overline{LQ_2}$ cross section [234], were used to set a lower mass limit of $M_{LQ_2} > 251\,\text{GeV}$ for $\beta = 1$ (in combination with the Run I analysis [235]).

The CDF collaboration published a similar search for $LQ_2\overline{LQ_2}$ production based on an integrated luminosity of $200\,\text{pb}^{-1}$ [230]. Instead of including a requirement on S_T in the selection, they found good signal-to-background discrimination using the two-dimensional event distribution in $p_T(\mu_1) + p_T(\mu_2)$ vs. $E_T(jet_1) + E_T(jet_2)$, which is demonstrated in Fig. 5.5.

In the $\mu\nu jj$ channel, one cannot utilize the transverse momentum balance to correct for the limited muon p_T resolution in a way similar to the $\mu\mu jj$ channel, as the neutrino is undetected. Thus, relatively tight quality criteria on the track matched to the reconstructed muon need to be applied to avoid large tails in the p_T resolution, which as a consequence leads to reduced reconstruction efficiencies.

The DØ collaboration presented preliminary results on the search for second generation scalar leptoquarks in the $\mu\nu jj$ channel using the Run IIa data set of $1\,\text{fb}^{-1}$ [229]. The preselection required exactly one reconstructed energetic muon ($E_T > 20\,\text{GeV}$, $|\eta| < 2$), large missing transverse energy ($\not\!\!E_\text{T} > 30\,\text{GeV}$) and at least two jets with $E_T > 25\,\text{GeV}$ and $|\eta| < 2.5$. In addition, events with the $\not\!\!E_\text{T}$ direction opposite in azimuth to the muon were removed, as they were likely due to badly reconstructed muons resulting in an overestimated $\not\!\!E_\text{T}$. The background consisted of W + jets and Z + jets (with a non-identified μ), which were both estimated from simulation using Alpgen (interfaced to Pythia and incorporating the MLM matching prescription, cf. Section 2.3.5), $t\bar{t}$-production simulated with Pythia, and a small contribution of QCD multijet production, which was derived from data. The W + jets background was normalized to data at preselection level within a region of transverse mass $M_T(\mu\nu)$ dominated by W production. Fig. 5.6 demonstrates that the muon and leading jet

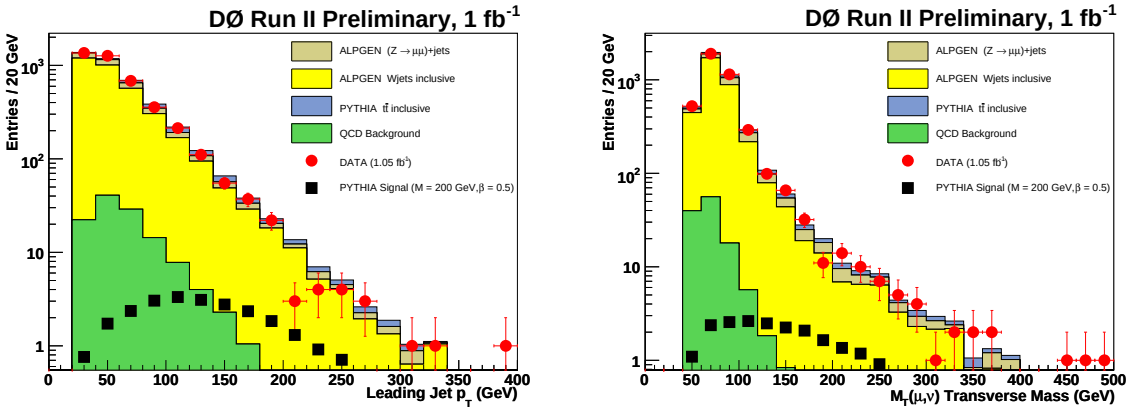

Figure 5.6: Search for second generation leptoquark production in the $\mu\nu jj$ final state [229]: Leading jet transverse momentum p_T (left) and transverse mass $M_T(\mu,\nu)$ (right) after preselection compared to the SM prediction and the expected leptoquark signal with $M_{LQ_2} = 200\,\text{GeV}$.

transverse momenta are well described by this background model.

The leptoquark signal was discriminated from the background using the muon-neutrino transverse mass $M_T(\mu\nu)$, the scalar transverse energy S_T, the transverse mass $M_T(\nu j_1)$ constructed from $\not\!\!E_T$ and the momentum of the leading jet, and the invariant mass of the muon jet combination closest to the assumed leptoquark mass. These selection requirements are motivated by the high leptoquark mass and consequently high transverse momenta of its decay products. For an assumed leptoquark mass $M_{LQ_2} = 200\,\text{GeV}$, six data events were selected with a background prediction of 6.4 ± 1.1 events, of which 50% were from W + jets production and the remainder from $t\bar{t}$, Z/γ^* + jets, and multijet production. The systematic error on the W + jet background prediction was found to be dominated by uncertainties in the jet energy scale and in the modelling of the jet transverse momentum shapes. The latter was evaluated by comparing the p_T distribution of the first and second leading jet observed in data with the predictions of ALPGEN and PYTHIA in a kinematic region dominated by W + jets production. From this analysis alone, a lower mass limit for scalar second generation leptoquarks of $M_{LQ_2} > 214\,\text{GeV}$ for $\beta = 0.5$ was derived.

5.1.2 Jets with heavy flavour identification

One of the most interesting and sought-after signals with leptons and jets in the final state is associated Higgs boson production, WH and ZH, with leptonic vector boson decays. At the Tevatron, ZH and in particular WH are the production modes with the highest discovery potential for the standard model Higgs boson with mass $M_H \lesssim 135\,\text{GeV}$, where it predominantly decays via $H \to b\bar{b}$. It is essential to require at least one b-tag in the final state to enhance the signal-to-background ratio in searches for the Higgs boson. Various algorithms which are employed to identify heavy-flavour jets were reviewed in Section 4.2.2.

Associated Higgs production: $W(\to \ell\nu)H(\to b\bar{b})$

For an assumed Higgs boson mass $M_H = 115\,\text{GeV}$ the standard model prediction for the WH production cross section at NNLO is $0.14\,\text{pb}^{-1}$ [236], compared to the cross section of the dominating $Wb\bar{b}$ background, $\sigma(Wb\bar{b}) = 4.0\,\text{pb}^{-1}$ (evaluated at NLO and including parton level cuts: $p_T(b,\bar{b}) > 20\,\text{GeV}$, $\Delta R_{b\bar{b}} > 0.75$ [45]). Evidently, the invariant mass of the two b jets, $M_{b\bar{b}}$, can be utilized to enhance the Higgs boson signal over the $Wb\bar{b}$ background, for which the differential cross section $\mathrm{d}\sigma/\mathrm{d}M_{b\bar{b}}$ is exponentially decreasing for $M_{b\bar{b}} \gtrsim 50\,\text{GeV}$ (cf. Section 4.2.1).

Both the DØ and CDF collaborations searched for associated WH production using leptonic W decay channels [214, 237]. Since then, several preliminary results using larger data samples and more sophisticated analysis techniques have been presented [238–242]. To give an example, the published DØ analysis [214] is summarized below.

The event selection required one high p_T electron, large missing transverse energy, and exactly two reconstructed jets with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.4$, which were both tagged as b jets using an impact parameter based algorithm. A selection with ≥ 3 jets was used as a control sample. To reduce the presence of b jets from gluon splitting, the separation between the two reconstructed jets was required to be $\Delta R > 0.75$, which suppressed part of the $Wb\bar{b}$ background. Before applying a cut on the dijet invariant mass, the signal-to-background ratio was found to be $S/B = 0.012$. The main background was determined to be $Wb\bar{b}$ jet production with a contribution of $\sim 40\%$, which was simulated using the ALPGEN event generator. Further background contributions were from $t\bar{t}$ and single-top production (simulated with PYTHIA , $\sim 30\%$) from W/Z production with mis-tagged light jets (simulated with ALPGEN, $\sim 20\%$), and from QCD multijet production ($\sim 20\%$), which was estimated from data using measured electron fake rates. The main systematic uncertainties on the predicted $Wb\bar{b}$ background were found to be due to uncertainties on the cross section, the b-tagging efficiency, and the jet energy calibration.

The dijet invariant mass distributions for the selected $W + 2$ jet events with ≥ 1 and 2 b-tags, respectively, are shown in Fig. 5.7 compared to the predicted background distribution and the expected signal from WH production. For an assumed Higgs boson mass $M_H = 115\,\text{GeV}$, a dijet mass window of $85 - 135\,\text{GeV}$ was selected, which increased S/B to 0.046. While the lower cross section limit on WH production derived in this analysis, based on an integrated luminosity of only $174\,\text{pb}^{-1}$, was a factor ~ 60 higher than the standard model prediction for $M_H = 115\,\text{GeV}$, the current preliminary combined DØ limit which is based on an integrated luminosity of $1\,\text{fb}^{-1}$ is about eight times the SM cross section for the same assumed M_H [243].

The $ZH \to \ell^+\ell^- b\bar{b}$ channel has a lower sensitivity for the search of the Higgs boson, as the production cross section and the branching ratio are each a factor of ~ 3 lower compared to $WH \to \ell\nu b\bar{b}$. Nevertheless this channel adds some sensitivity in the combination of the various Higgs searches. The DØ collaboration searched for the Higgs boson in $ZH \to \ell^+\ell^- b\bar{b}$ and found that the main background sources were $Zb\bar{b}$, Z + light jets (including c jets), and $t\bar{t}$ production, with approximately equal contributions [244].

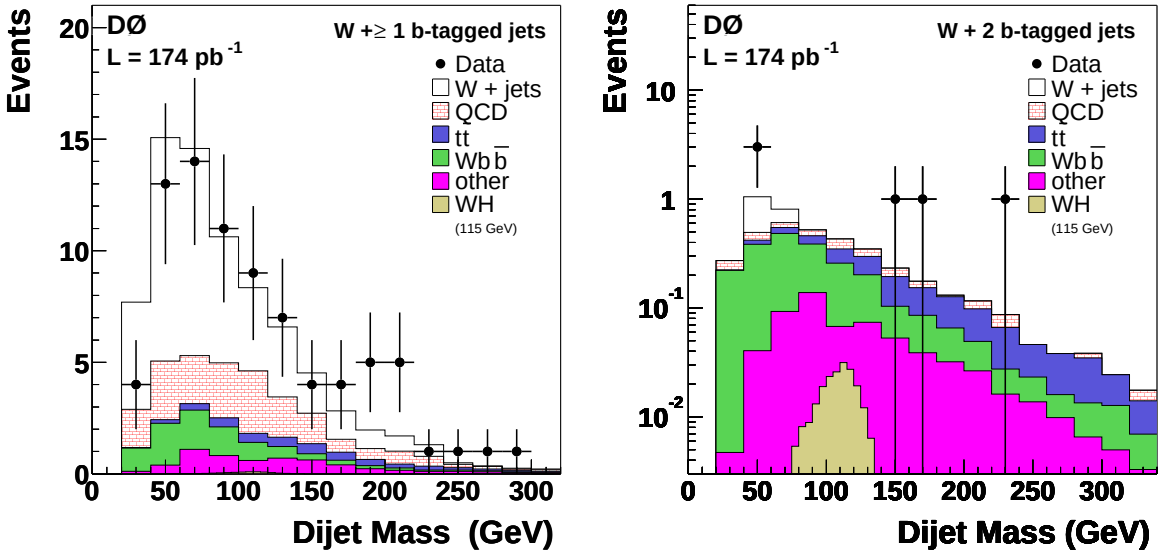

Figure 5.7: Search for $H \to b\bar{b}$ produced in association with W bosons [214]: Distribution of the dijet invariant mass for W + 2 jet events, when at least one jet is b tagged (left) and when both jets are tagged (right). The data is compared to the standard model expectation and for the double-tagged sample in addition to the expectation of WH production with $M_H = 115\,\text{GeV}$.

5.2 Final states with jets and missing transverse energy in searches

The production of jets in association with a Z boson which decays via $Z \to \nu\bar{\nu}$ leads to event topologies with missing transverse momentum $\not\!\!E_\text{T}$ and jets, including basic signatures such as monojets or acoplanar jets. These peculiar final states could also indicate new processes as well as the production of new particles, e.g. graviton emission in models with large extra dimensions, leptoquarks decaying into neutrinos and quarks, and the pair-production of scalar quarks and gluinos in supersymmetric models. Associated Higgs-production $ZH(\to \nu\bar{\nu}b\bar{b})$ has a signature of two acoplanar b jets, which need to be discriminated from the $Zb\bar{b}$ background.

5.2.1 Jets without heavy flavour identification

In event selections requiring only moderate minimal missing transverse energy, the instrumental background originating from QCD multijet production can be substantial, as the limited jet energy resolution leads to mismeasurements in the vectorial sum of deposited energies. The searches described below require high $\not\!\!E_\text{T}$ to ensure that this background is largely suppressed.

Large extra dimensions: monojet production

Despite the success of the standard model, one of its main unappealing features is the hierarchy problem, i.e. the observation that gravity is more than 30 orders of magnitude weaker than the other gauge forces. An elegant and natural solution to this problem would be given by the introduction of large extra dimensions (LED) in which only the gauge bosons of gravity, i.e. the gravitons, propagate [245–247]. In the presence of additional extra dimensions, the fundamental Planck scale, which characterizes the strength of gravity, could be vastly decreased from $\sim 10^{19}\,\text{GeV}$ to the TeV-scale, which is comparable to the scales of the other three gauge forces. The introduction of LEDs would lead to new phenomena observable at high energy colliders [248, 249]. The most prominent experimental signature of graviton production in hadron collisions would be the monojet topology, i.e. a final state consisting of a single high-p_T jet and consequently high $\not\!\!E_{\text{T}}$, as the graviton which would not be detected would be produced in association with a quark or gluon.

The DØ and CDF collaborations, both searched for the monojet signature of large extra dimensions in their Run I data samples [250, 251]. The CDF collaboration published a new search based on an integrated luminosity of $370\,\text{pb}^{-1}$ collected with a single-jet trigger during Run II [252]. The event selection required one reconstructed central ($|\eta| < 1.0$) jet with transverse energy $E_T > 150\,\text{GeV}$ and missing transverse energy $\not\!\!E_{\text{T}} > 120\,\text{GeV}$. To increase the acceptance for signal events with a second jet originating from initial or final state radiation, events containing a second jet with $E_T < 60\,\text{GeV}$ were accepted as well, but only in the absence of a third jet with $E_T > 20\,\text{GeV}$.

More than 90% of the standard model background was found to originate from $Z(\to \nu\bar{\nu})$ + jets and $W(\to \ell\nu)$ + jets production. For the latter, events with non-identified leptons could survive the signal selection, despite the fact that its contribution was suppressed by vetoing on the presence of isolated tracks and jets with high electromagnetic fractions. Instead of using the simulation to predict the number of Z and W boson background events, it was determined by measuring cross sections for $Z(\to \ell^+\ell^-)$ + jets and $W(\to \ell\nu)$ + jets production in the electron and muon decay channels, which are plotted in Fig. 5.8 (left) as a function of the leading jet E_T. The $Z(\to \nu\bar{\nu})$ + jets background was estimated with two methods, one based on the measured $Z(\to \ell^+\ell^-)$ + jets cross section and the Z branching fractions, the other based on the measured $W(\to \ell\nu)$ + jets cross section, the W and Z branching fractions, and a NLO calculation [45] of the ratio of the Z/W + jets cross sections. The latter method was found to be more precise due to the larger statistics of the W sample. The two independent methods gave consistent results and the combined estimate was 130 ± 14 background events from $Z(\to \nu\bar{\nu})$ + jets production. The background contribution from $W(\to \ell\nu)$ + jets was determined from its measured cross section and the percentage of events passing the signal selection, which was obtained from simulation, and found to be 113 ± 13. The uncertainties on the predictions for both background contributions are predominantly of statistical nature and fully correlated. The total expected background was determined to be 265 ± 30 events compared to 263 observed events.

As the kinematic distributions for the LED signal and the standard model back-

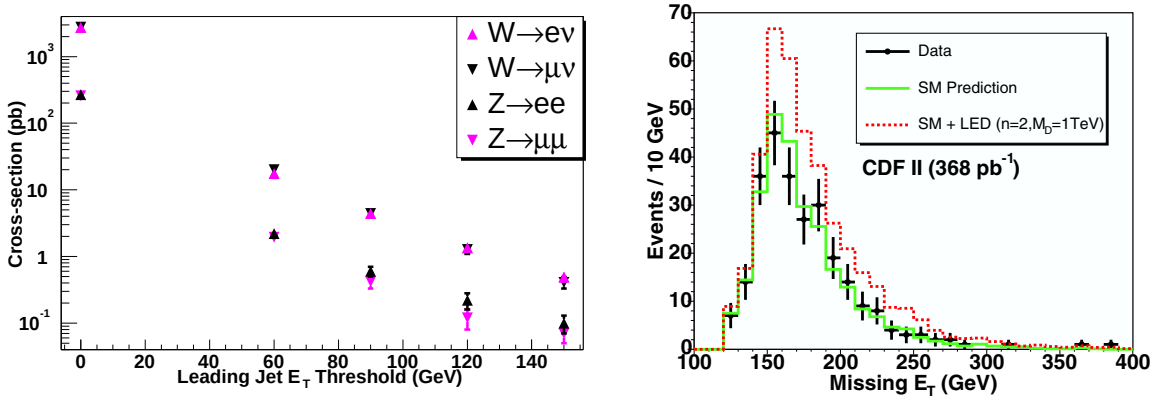

Figure 5.8: Search for large extra dimensions in the monojet signature [252]: Left: W and Z boson cross sections as a function of leading jet E_T. Right: the $\not\!\!E_{\mathrm{T}}$ distributions of the candidate data events compared to the distributions predicted by the standard model and when including an additional contribution from the LED signal.

ground are similar, which was explicitly demonstrated for the $\not\!\!E_{\mathrm{T}}$ distribution (see Fig. 5.8, right), only the integral rates for observed events, background, and expected signal were utilized to derive lower limits on the Planck scale M_D in the $(4+n)$-dimensional space for $n = 2-6$. For $n = 2$ a limit of $M_D > 1.18\,\mathrm{TeV}$ was obtained which corresponds to a compactification radius (i.e. the size of the extra dimensions) of 0.35 mm.

Leptoquark pair-production: $LQ\overline{LQ} \to \nu q \nu q$

The pair-production of leptoquarks with a vanishing branching ratio into charged leptons and quarks ($\beta = 0$) would generate final states consisting of two neutrinos and two quarks, i.e. an event topology of two acoplanar high-p_T jets and $\not\!\!E_{\mathrm{T}}$. Both collaborations, CDF and DØ, searched for scalar leptoquarks in events with this signature [253, 254]. The DØ analysis was based on $360\,\mathrm{pb}^{-1}$ of data collected using a jets + $\not\!\!E_{\mathrm{T}}$ trigger, which placed requirements on the vector sum of the jet transverse momenta, $\not\!\!H_{\mathrm{T}} = |\sum_{\mathrm{jets}} \vec{p}_T|$ [254]. The event selection required exactly two reconstructed central acoplanar jets with $p_T(\mathrm{jet}_1) > 60\,\mathrm{GeV}$, $p_T(\mathrm{jet}_2) > 50\,\mathrm{GeV}$, $|\eta| < 1.5$, and an azimuthal jet separation $\Delta\Phi(\mathrm{jet}_1, \mathrm{jet}_2) < 165°$. In addition, selected events also needed to have high missing transverse energy $\not\!\!E_{\mathrm{T}} > 80\,\mathrm{GeV}$ (and $\not\!\!H_{\mathrm{T}} > 40\,\mathrm{GeV}$), which suppressed the majority of the instrumental background due to QCD multijet production (see Fig. 5.9, left). A veto on isolated electrons, muons, and tracks rejected a large fraction of events originating from W/Z + jet processes. The correlations in azimuthal angle between the jet and $\not\!\!E_{\mathrm{T}}$ directions were utilized to further suppress both the instrumental and SM backgrounds. Fig. 5.9 (right pane) shows the sum of the jet-$\not\!\!E_{\mathrm{T}}$ angular separations, which provided additional discrimination between the LQ signal and the SM background.

After all cuts the SM background was entirely dominated by $Z(\to \nu\bar{\nu}) + 2$ jet and $W(\to \ell\nu)$ + jet with about equal contributions. These processes were simulated using

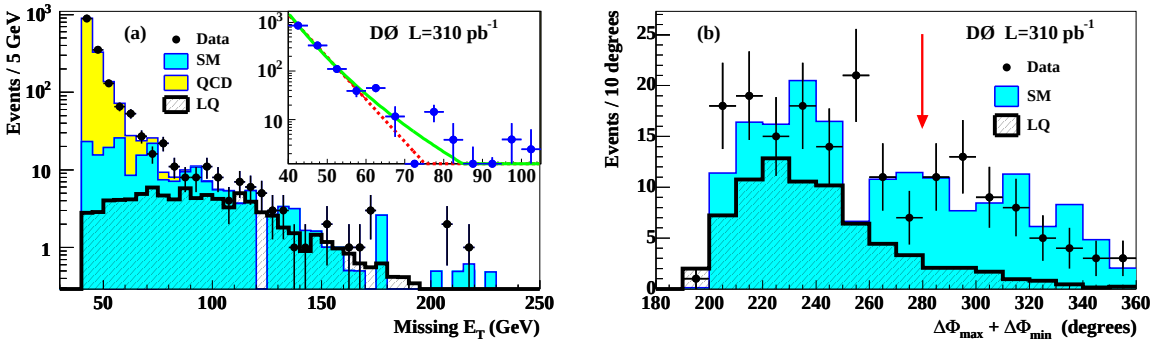

Figure 5.9: Search for scalar leptoquarks in the acoplanar jet topology [254]: Left: the E_T distribution after all selection cuts except the cut on $\not\!\!E_{\mathrm{T}}$ compared to the sum of instrumental background (QCD multijet production) and the SM background. Also shown is the expected signal for leptoquark pair production with $M_{LQ} = 140\,\mathrm{GeV}$. The inset shows how the instrumental background is estimated from both an exponential and power law fit to the observed $\not\!\!E_{\mathrm{T}}$ distribution (after subtraction of the SM background). Right: the distribution of $\Delta\Phi_{max} + \Delta\Phi_{min} = \Delta\Phi(\not\!\!E_{\mathrm{T}}, j_1) + \Delta\Phi(\not\!\!E_{\mathrm{T}}, j_2)$ in events with exactly two acoplanar jets compared to the SM background and the expected leptoquark signal. The selection cut is indicated by the arrow.

Alpgen interfaced with Pythia. The dominating uncertainties on the background prediction were found to be due to the limited statistics in the simulation and due to its normalization. The latter was inferred from a comparison of data and simulated $Z(\to \ell^+\ell^-) + 2$ jet events. The analysis set a lower mass limit of $M_{LQ} > 136\,\mathrm{GeV}$ for a single-generation scalar leptoquark, decaying exclusively in a neutrino and a quark.

Supersymmetry: $\tilde{q}, \tilde{g}$ production

Supersymmetric models predict the existence of scalar quarks (or squarks, $\tilde{q}$) and spin-1/2 gluinos ($\tilde{g}$) as super-partners of the standard model quarks and gluons [13]. R-parity is introduced as a new multiplicative quantum number to differentiate between standard model ($R = 1$) and supersymmetric ($R = -1$) particles. As a consequence of the assumption of R-parity conservation, supersymmetric particles are produced in pairs and the lightest supersymmetric particle (LSP) needs to be stable. In supersymmetric models inspired by supergravity, the lightest neutralino $\tilde{\chi}_1^0$, which is a mixture of the super-partners of the neutral gauge and Higgs bosons, is usually assumed to be the LSP. As it is only weakly interacting it escapes detection and thus gives, similar to the neutrino, a signature of missing transverse energy $\not\!\!E_{\mathrm{T}}$.

If sufficiently light, squarks and gluinos could be produced in pairs at the Tevatron. If $M(\tilde{q}) < M(\tilde{g})$, mostly pairs of squarks would be produced, which decay via $\tilde{q} \to q\tilde{\chi}_1^0$, resulting in an event signature of two acoplanar jets and $\not\!\!E_{\mathrm{T}}$. If $M(\tilde{g}) > M(\tilde{q})$, gluinos would decay according to $\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$ and their pair-production would give topologies with many jets and $\not\!\!E_{\mathrm{T}}$. In the case of $M(\tilde{g}) \approx M(\tilde{q})$ and $\tilde{q}\tilde{g}$-production the final state is expected to often consist of three jets and $\not\!\!E_{\mathrm{T}}$.

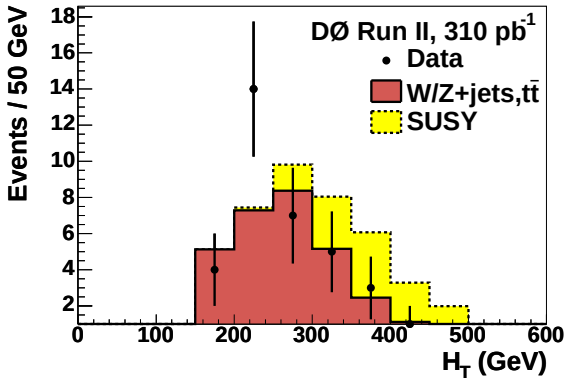

Figure 5.10: Search for squarks and gluinos in events with jets and missing transverse energy [255]: $H_T = \sum_{\text{jets}} E_T$ distribution for the "3-jet" selection which is optimized for $m_{\tilde{q}} \approx m_{\tilde{g}}$ and requires three high energetic jets and large $\not\!\!E_{\text{T}}$. Shown is the distribution after all selection cuts except the cut on H_T. After all cuts including $H_T > 350\,\text{GeV}$, 3.9 background events are expected, of which 2.3 are from W and Z production.

The DØ collaboration searched for the production of squarks and gluinos using three different event selections which were targeted at the scenarios described above [255]. For the three analyses, a similar data set, trigger, and preselection as in the leptoquark search (in the channel $LQ\overline{LQ} \to \nu q \nu q$) described above were used. A common preselection which required two central acoplanar jets and substantial $\not\!\!E_{\text{T}}$ was applied. The background from QCD multijet production was estimated from an exponential fit to the $\not\!\!E_{\text{T}}$ distribution and found to be negligible for the dijet and three-jet event selection after all selection cuts. The background from W/Z + jets events, which was simulated using ALPGEN interfaced with PYTHIA, was largely reduced by vetoing the presence of isolated electrons or muons.

The analysis with the largest sensitivity for the scenario $M(\tilde{g}) \approx M(\tilde{q})$ required three jets with minimal E_T of 60, 40, and 30 GeV, respectively, and large $\not\!\!E_{\text{T}} \geq 100\,\text{GeV}$. A cut on the scalar sum of the jet transverse energies, $H_T = \sum_{\text{jets}} E_T \geq 350\,\text{GeV}$, provided additional discrimination between the expected signal and the SM background, which is illustrated in Fig. 5.10. The dominating contributions to the background were found to be from W + jet, $t\bar{t}$, and Z + jet production, with uncertainties mainly due to the precision of the jet energy calibration and the accuracy of the predictions for their cross sections. With an integrated luminosity of $310\,\text{pb}^{-1}$ the DØ collaboration set limits on the squark and gluino masses of $M_{\tilde{q}} > 325\,\text{GeV}$ and $M_{\tilde{g}} > 241\,\text{GeV}$ in their most conservative scenario.

5.2.2 Jets with heavy flavour identification

Event signatures with missing transverse energy and b or c quark jets provide promising discovery potentials for new signals as the standard model background from $Z(\to \nu\bar{\nu})$ + jet and QCD multijet production is largely reduced.

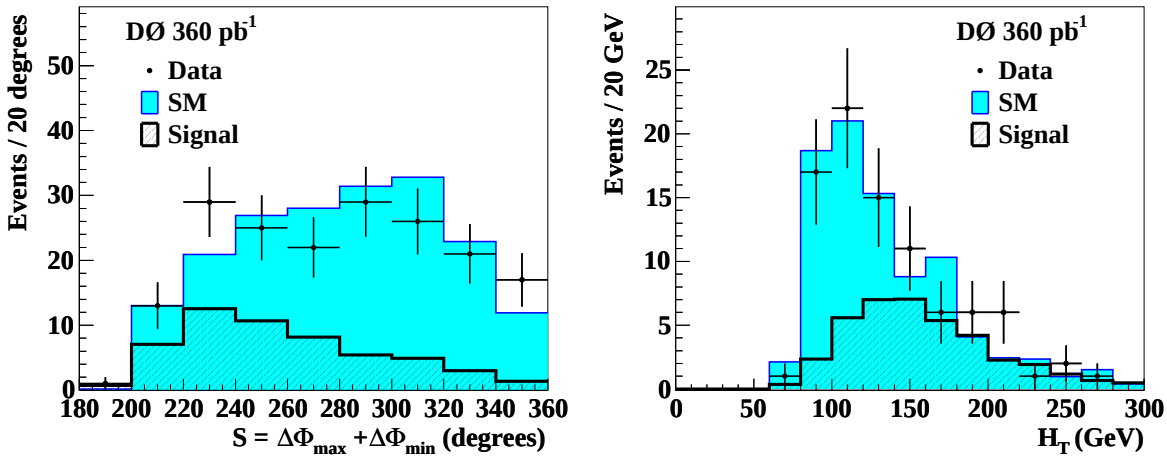

Figure 5.11: Search for the pair production of scalar top quarks in the acoplanar charm jet final state [260]: The distributions of $\Delta\Phi_{max} + \Delta\Phi_{min} = \Delta\Phi(\not\!\!E_T, j_1) + \Delta\Phi(\not\!\!E_T, j_2)$ (left) and of H_T (right) compared to the SM prediction and to the expected signal for $M_{\tilde{t}} = 140\,\text{GeV}$ and $M_{\tilde{\chi}_1^0} = 60\,\text{GeV}$.

Pair-production of scalar top quarks: $\tilde{t} \to c\chi_1^0$

Scalar top quarks (stop, $\tilde{t}$) could be substantially lighter than the other squarks, as the high mass of the top mass generates a large mixing between its chiral supersymmetric partners [256]. The mass eigenstates are therefore widely split, thus reducing the mass of the lighter scalar top quark considerably. If the favoured stop decay channels $\tilde{t} \to t\tilde{\chi}_1^0$ and $\tilde{t} \to b\tilde{\chi}_1^+$ are kinematically forbidden, the flavour-changing loop decay $\tilde{t} \to c\tilde{\chi}_1^0$ might be dominating [257]. Here, the neutralino $\tilde{\chi}_1^0$ and chargino $\tilde{\chi}_1^+$ denote the lightest mass eigenstates resulting from the mixing of the super-partners of the neutral or charged electroweak gauge and Higgs bosons. The pair production of scalar top quarks each decaying via $\tilde{t} \to c\chi_1^0$ lead to an event topology consisting of two acoplanar charm jets and $\not\!\!E_T$.

Both the CDF and DØ collaborations searched for the pair production of scalar top quarks in this final state using data collected in Run I [258,259]. The DØ collaboration recently published a search based on $360\,\text{pb}^{-1}$ of data collected during Run II [260], which is largely based on its search for pair production of leptoquarks decaying into a quark and neutrino [254] which was reported on above. The main differences between the analyses searching for stop or leptoquark production, respectively, are motivated by the large mass of the $\tilde{\chi}_1^0$ compared to the neutrino and the presence of a c quark in the stop decay. Consequently, the minimal transverse energies for the leading and second-leading jets were reduced to 40 GeV and 20 GeV, respectively, and the requirement on missing transverse energy was loosened to $\not\!\!E_T > 60\,\text{GeV}$. At least one of the jets was required to be tagged as a heavy-flavour jet using an impact-parameter based algorithm (cf. Section 4.2.2) with a loose threshold giving a relatively large tagging efficiency of $\sim 30\%$ for c jets at the cost of a substantial mistag probability of $\sim 4\%$ for light jets.

The instrumental background from QCD multijet production was further suppressed to a negligible level utilizing the correlations in azimuth between jet and $\not\!\!E_T$

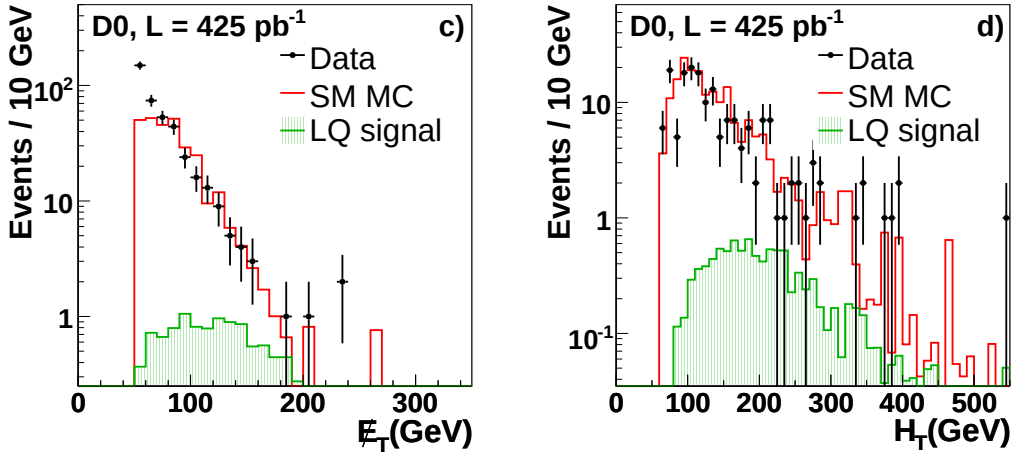

Figure 5.12: Search for third generation leptoquark pair-production decaying into $\nu b\bar{\nu}\bar{b}$ [261]: The $\not\!\!E_{\mathrm{T}}$ and scalar H_T distributions for the muon trigger sample before b-tagging. The leptoquark (LQ_3) signal is shown for an assumed mass $M_{LQ_3} = 200\,\mathrm{GeV}$.

directions and the asymmetry between $\not\!\!E_{\mathrm{T}}$ and the vector sum of the jet transverse momenta, $\not\!\!H_{\mathrm{T}}$. For the discrimination between the stop signal and the electroweak background from W/Z + jets production, the distributions in the scalar sum of the jet energies, H_T, and in the sum of the differences in azimuth between $\not\!\!E_{\mathrm{T}}$ and each jet, $\Delta\Phi(\not\!\!E_{\mathrm{T}}, j_1) + \Delta\Phi(\not\!\!E_{\mathrm{T}}, j_2)$, were used (see Fig. 5.11). The cuts on these two variables were optimized as a function of stop mass $M_{\tilde{t}}$. The dominating background processes after the full selection were $W(\rightarrow \ell\nu)$ + jets and $Z(\rightarrow \nu\bar{\nu})$ + jets with some small additional contributions from top and diboson production. The systematic uncertainty on the background estimate was determined to be similar as in case of the leptoquark search. An additional (but non-dominating) uncertainty was due to the heavy-flavour tagging. Exclusion regions in the plane given by the stop and neutralino masses were derived, which reached $M_{\tilde{t}} = 134\,\mathrm{GeV}$ for $M_{\tilde{\chi}_1^0} = 48\,\mathrm{GeV}$.

Leptoquark pair-production: $LQ_3\overline{LQ}_3 \rightarrow \nu b\bar{\nu}\bar{b}$

The DØ collaboration recently published a search for the pair-production of third-generation scalar leptoquarks decaying into a neutrino and a b quark using $425\,\mathrm{pb}^{-1}$ of data collected with a missing transverse energy and a single-muon trigger [261]. Similar searches were performed before by both CDF and DØ in Run I [262–264]. The new analysis combines two selections which start from data samples collected with the missing transverse energy and with the single-muon triggers, respectively. In both cases minimal requirements on the leading and second-leading jet p_T and on $\not\!\!E_{\mathrm{T}}$ were applied. Furthermore, two jets were required to be tagged as b jets with at least one of them by the jet lifetime probability algorithm based on the impact-parameters of the associated tracks. For the selection using the single-muon triggers, which is motivated by the semi-leptonic decays of B-mesons, one of the jets was required to be tagged with the soft-muon tagger.

Cuts on the final selection variables, $\not\!\!E_{\mathrm{T}}$ and scalar H_T, were optimized as function of the assumed leptoquark mass M_{LQ_3}. Fig. 5.12 shows the measured $\not\!\!E_{\mathrm{T}}$ and H_T distributions for the muon triggered sample compared to the SM background prediction and the expected leptoquark signal with $M_{LQ_3} = 200\,\mathrm{GeV}$. The main background sources were determined to be mostly top, $W/Z + b\bar{b}$, and $W/Z + c\bar{c}$ production. The uncertainty on their contribution was found to be dominated by uncertainties on the cross-section predictions, on the jet energy calibration, and on the b-tagging efficiency. Assuming that the leptoquarks decay exclusively in a neutrino and a b quark, a mass limit on third-generation leptoquarks of $M_{LQ_3} > 229\,\mathrm{GeV}$ was derived.

Associated Higgs production: $\boldsymbol{Z(\to \nu\bar{\nu})H(\to b\bar{b})}$

Assuming a relatively light standard model Higgs boson, $M_H \lesssim 135\,\mathrm{GeV}$, the $ZH \to \nu\bar{\nu}b\bar{b}$ channel is particularly promising due to the large $Z \to \nu\bar{\nu}$ and $H \to b\bar{b}$ branching fractions. Although the cross section for ZH production is a factor ~ 3 lower compared to WH production, the product of cross section times branching ratio for $ZH \to \nu\bar{\nu}b\bar{b}$ and $WH \to \ell\nu b\bar{b}$ are comparable. Therefore, ZH production leading to final states with significant $\not\!\!E_{\mathrm{T}}$ and two acoplanar b jets has a competitive sensitivity in the search for the SM Higgs boson. In addition, this final state has some sensitivity to WH production, as the charged lepton from the W boson decay might be undetected.

The DØ collaboration published a search for the Higgs boson in events with $\not\!\!E_{\mathrm{T}}$ and two acoplanar b jets using an integrated luminosity of $260\,\mathrm{pb}^{-1}$ [265]. Since then both CDF and DØ presented preliminary results based on nearly $1\,\mathrm{fb}^{-1}$ of data [266, 267].

The published DØ analysis required $\not\!\!E_{\mathrm{T}} > 50\,\mathrm{GeV}$, two or three reconstructed jets with $p_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$. Accepting events with an additional third jet increases the signal efficiency in the presence of initial or final state radiation. The main sources of SM background for this selection are Z/W + jets and top production. In addition, the instrumental background from multijet production has a large contribution. W and Z boson production with leptonic decays were suppressed with a veto on events containing an isolated electron or muon, whilst the $t\bar{t}$ background could be largely rejected by requiring the scalar sum of jet transverse momenta $H_T < 240\,\mathrm{GeV}$. The instrumental background was suppressed using several requirements on the acoplanarity of the leading jets, on the separation in azimuth between $\not\!\!E_{\mathrm{T}}$ and jets, on the vector sum of all track momenta, and on the asymmetry between $\not\!\!H_{\mathrm{T}}$ and $\not\!\!E_{\mathrm{T}}$. The latter distribution was also utilized to estimate the contribution of the instrumental background in the signal region. Before b-tagging the data was found to be well described by this background model, which is demonstrated in Fig. 5.13 (left).

To select b jets, the jet lifetime probability algorithm was applied. Two b-tag selections were used for the search: The single-tag sample required exactly one tight b-tag with an efficiency of $\sim 30\%$ and low mistag rate. For the double-tag sample looser b-tags with efficiencies of $\sim 40\%$ and $\sim 50\%$ were applied. The main backgrounds in the single-tag sample were estimated to be W/Z + light-flavour jets (including c quarks), multijet, and top quark production. In the double-tag sample $Zb\bar{b}$ and $Wb\bar{b}$ production were found to have larger contributions than the associated production of the vector bosons with light jets. Additional, main backgrounds were due to top quark and multijet production. As in the case of the search for $W(\to \ell\nu)H$ production

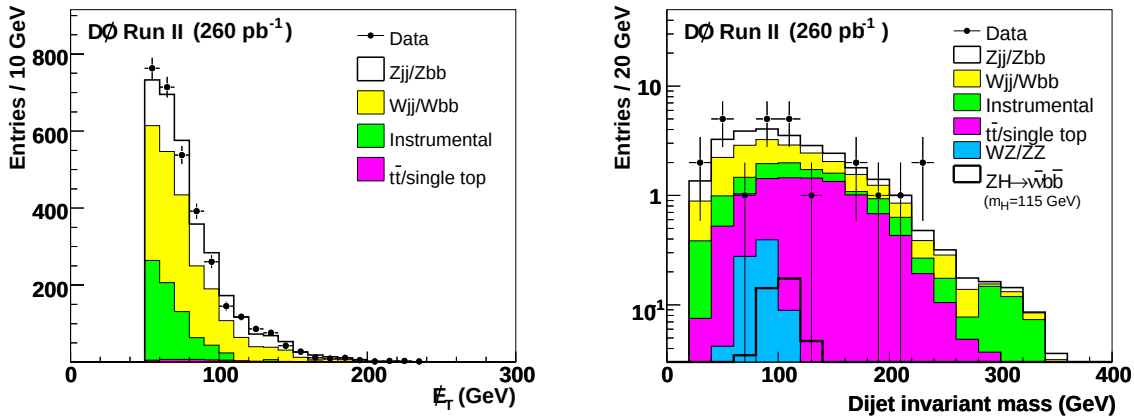

Figure 5.13: Search for the SM Higgs boson in the $ZH \to \nu\nu b\bar{b}$ channel [265]: Left: the $\not\!\!E_T$ distribution after event selection except for b-tagging compared to the SM background prediction. Right: the distribution of the invariant mass M_{jj} of the two tagged b jets after all selection cuts except the cut on M_{jj} , compared to the SM background prediction. The black histogram shows the expectation for $Z(\to \nu\bar{\nu})H$ with $M_H = 115\,\text{GeV}$.

summarized above, a window in the dijet mass distribution was utilized to search for an excess which could be attributed to a Higgs signal (see Fig. 5.13, right). For $M_H = 115\,\text{GeV}$ the cut on the dijet mass increased the signal-over-background ratio from $S/B = 0.0089$ to $S/B = 0.022$. Combining both the single-tag and double-tag selections, a sensitivity of about 40 times the ZH production cross section was achieved. The current preliminary DØ limit, which combines several analyses based on data sets with $1\,\text{fb}^{-1}$ is about eight times the SM cross section for M_H [243].

5.3 From Tevatron to LHC

The Large Hadron Collider (LHC) at CERN is expected to start operation with pp collisions at $\sqrt{s} = 14\,\text{TeV}$ in 2008. Many anticipated searches at the LHC will have substantial background contributions from W/Z + jet production.

One of the major goals will be the discovery of the Higgs boson. Once its existence is established, its properties, e.g. couplings and branching fractions, will be studied in-depth. In contrast to the situation at the Tevatron, the associated production of the Higgs boson with the W or Z boson has only limited discovery potential compared to gluon-gluon fusion and vector boson fusion processes [185, 268–270].

In the vector boson fusion (VBF) process, the Higgs boson is accompanied by one jet each in both the forward and backward regions of the detector, which originates from the initial quarks emitting the vector bosons. As no colour is exchanged between the quarks, central jet activity is suppressed. Although vector boson fusion accounts only for a moderate fraction of the total Higgs production cross section (e.g. about 20% for $M_H \lesssim 180\,\text{GeV}$) the topology of the jet production can be exploited to suppress

background processes. Nevertheless, the following production channels have substantial backgrounds from W/Z + jets processes after full event selection [184, 185]. The VBF mode $qqH(\rightarrow \tau\tau)$, with at least one tau decaying leptonically, is particularly sensitive for Higgs masses $M_H \approx 120\,\text{GeV}$. For $M_H \gtrsim 500\,\text{GeV}$, the most important discovery channels are $qqH(\rightarrow WW \rightarrow \ell\nu jj)$ and $qqH(\rightarrow ZZ \rightarrow \ell\ell jj)$. As the selection of these Higgs boson production modes is based on a central jet veto, the accurate simulation of the central jet activity in vector boson events with forward jets is essential (cf. Section 4.1.4).

Evidence for Supersymmetry could be already obtained soon after the start of the LHC. If squarks and gluinos are kinematically accessible, they should be pair-produced with large cross sections, which dominate the inclusive production rate for any supersymmetric particle. Squarks and gluinos would decay in cascades to quarks and the lightest neutralino, $\tilde{\chi}_1^0$, possibly via $\tilde{\chi}_2^0$ or $\tilde{\chi}_1^\pm$, which themselves decay to $\tilde{\chi}_1^0$ and, with some probability, leptons. If R-parity is conserved, the weakly interacting $\tilde{\chi}_1^0$ is stable and escapes detection. Therefore, inclusive signatures of Supersymmetry at the LHC will be topologies with missing transverse momentum and several jets and possibly additional leptons [268, 269, 271–274]. Major background contributions are expected from $Z(\rightarrow \nu\bar{\nu})$ + jets and $W(\rightarrow \ell\nu)$ + jets production. Early studies on the sensitivity reach at the LHC utilized parton shower event generators for the simulation of the weak boson background, which considerably underestimated their contribution.

It can be expected that many other searches in final states including leptons, $\not\!\!E_\text{T}$, and jets will be performed at the LHC, including searches for leptoquark production [275, 276]. All these anticipated searches at the LHC would benefit from a more precise understanding of the associated production of vector bosons with jets. Future measurements of differential production cross sections for these processes at the Tevatron are expected to have reduced statistical and systematic uncertainties, due to the increase in integrated luminosity and potential improvements in the jet energy calibration. Subsequently, these measurements could be utilized to further tune and validate event generators. Ultimately, similar measurements at the LHC will follow, which will reach unprecedented kinematic regions, but as they require a precise understanding of the jet energy calibration, corrected differential cross sections are unlikely to be obtained soon after the start of the LHC. Alternatively, the background from Z/W + jets production can be estimated from uncorrected data, similar to the method applied in CDF's monojet search (cf. Section 5.2), e.g. an uncorrected measurement of $Z \rightarrow \mu\mu$ + jets production can be used to determine the background contribution from $Z \rightarrow \nu\bar{\nu}$ and $W \rightarrow \tau\nu$ in the $\not\!\!E_\text{T}$ + jets signal selection [269].

Chapter 6

Summary and Conclusions

The production of the weak bosons, Z and W, is a dominant background contribution in many searches for new phenomena and the Higgs boson at the Tevatron and LHC colliders. In particular, their associated production with jets generates final state signatures with leptons, missing transverse momentum, and one or more jets, which resemble those of a vast collection of signals for new phenomena. Thus, an accurate description of vector boson production, based both on precision data and phenomenological modelling, is required.

The inclusive production of vector bosons is well described by QCD. Perturbative corrections on integral cross sections and differential distributions were calculated up to NNLO and reach high precision. At the Tevatron cross sections for inclusive Z and W boson production, including differential shapes as function of rapidity and transverse momentum, were measured in detail. These measurements can be used to test QCD models and to tune event generators.

The production of vector bosons in association with jets was measured at the Tevatron. Jet multiplicities, jet transverse momentum distributions, and angular correlations were analysed. Traditional parton shower event generators, e.g. Pythia and Herwig, are well established to describe the general features of inclusive W and Z boson production, but underestimate both the jet multiplicity and the hardness of the jet energy distributions. By contrast, event generators, which consistently combine matrix-element calculations up to large parton multiplicities with parton shower models, e.g. Alpgen and Sherpa, provide a good model for W/Z + jets production. Future measurements, which are expected to reach higher precision and will study additional kinematic distributions, will help to further improve and validate these event generators. The production of heavy-flavour jets with a W or Z boson is an important background contribution in searches for the Higgs boson in associated WH and ZH production. Cross sections for Zb and $Wb\bar{b}$ production were measured, but the measurements are still limited by substantial statistical and systematic uncertainties.

Finally, the relevance of W/Z + jets production as main background in searches for new phenomena was discussed in detail. Searches in final states with leptons and jets and with missing transverse momentum and jets, for which the associated production of vector bosons is a dominating background, were reviewed. In most cases the electroweak background was estimated using simulation, but it was also attempted to constrain it primarily with data itself. Methods, to discriminate new phenomena

signals from the background were presented. For models which predict heavy new particles, the high energy of the decay particles could be exploited to partly suppress the electroweak background. The invariant dijet mass distribution of b-tagged jets was utilized to remove large parts of $Wb\bar{b}$ and $Zb\bar{b}$ backgrounds in searches for the Higgs boson. Systematic uncertainties due to the modelling of the W/Z + jet background were generally found to be significant.

While no indication for new phenomena at the Tevatron have yet been found, the Higgs boson should be discovered no later than at the LHC and prospects for the discovery of physics beyond the standard model are enormous. For many of these searches an in-depth and precise understanding of the associated production of vector bosons with jets will be essential.

References

[1] S. L. Glashow, Nucl. Phys. **22**, 579 (1961).

[2] A. Salam and J. C. Ward, Phys. Lett. **13**, 168 (1964).

[3] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[4] H. D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973).

[5] D. J. Gross and F. Wilczek, Phys. Rev. Lett. **30**, 1343 (1973).

[6] H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Lett. **B47**, 365 (1973).

[7] P. W. Higgs, Phys. Lett. **12**, 132 (1964).

[8] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).

[9] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

[10] J. F. Gunion, H. E. Haber, G. L. Kane and S. Dawson, *The Higgs Hunter's Guide* (Addison Wesley, Reading, MA, USA, 1990).

[11] LEP Collaborations ALEPH, DELPHI, L3, OPAL, J. Alcaraz *et al.*, hep-ex/0612034.

[12] LEP Working Group for Higgs boson searches, R. Barate *et al.*, Phys. Lett. **B565**, 61 (2003), [hep-ex/0306033].

[13] S. P. Martin, in G. L. Kane, *Perspectives on supersymmetry* (World Scientific, Singapore, 1990) (1997), [hep-ph/9709356].

[14] S. D. Drell and T.-M. Yan, Ann. Phys. **66**, 578 (1971).

[15] J. D. Bjorken, Phys. Rev. **179**, 1547 (1969).

[16] R. P. Feynman, Phys. Rev. Lett. **23**, 1415 (1969).

[17] J. C. Collins, D. E. Soper and G. Sterman, Phys. Lett. **B134**, 263 (1984).

[18] J. C. Collins, D. E. Soper and G. Sterman, Nucl. Phys. **B261**, 104 (1985).

[19] R. K. Ellis, W. J. Stirling and B. R. Webber, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8**, 1 (1996).

[20] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[21] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[22] UA2, R. Ansari *et al.*, Phys. Lett. **B186**, 452 (1987).

[23] UA2, J. Alitti *et al.*, Z. Phys. **C49**, 17 (1991).

[24] CDF, T. Dorigo, hep-ex/9806022.

[25] D0 Collaboration, D0 Note 5205-CONF.

[26] W. J. Stirling and M. R. Whalley, J. Phys. **G19**, D1 (1993).

[27] CDF, F. Abe *et al.*, Phys. Rev. Lett. **67**, 2937 (1991).

[28] CDF, F. Abe *et al.*, Phys. Rev. Lett. **66**, 2951 (1991).

[29] D0, B. Abbott *et al.*, Phys. Rev. Lett. **80**, 5498 (1998), [hep-ex/9803003].

[30] D0, B. Abbott *et al.*, Phys. Rev. Lett. **84**, 2792 (2000), [hep-ex/9909020].

[31] D0, B. Abbott *et al.*, Phys. Rev. **D61**, 032004 (2000), [hep-ex/9907009].

[32] D0, B. Abbott *et al.*, Phys. Lett. **B513**, 292 (2001), [hep-ex/0010026].

[33] D0, V. M. Abazov *et al.*, Phys. Lett. **B517**, 299 (2001), [hep-ex/0107012].

[34] D0 Collaboration, D0 Note, July 2006.

[35] A. S. Ito *et al.*, Phys. Rev. **D23**, 604 (1981).

[36] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001), [hep-ph/0010017].

[37] R. Hamberg, W. L. van Neerven and T. Matsuura, Nucl. Phys. **B359**, 343 (1991).

[38] W. L. van Neerven and E. B. Zijlstra, Nucl. Phys. **B382**, 11 (1992).

[39] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, Phys. Rev. **D69**, 094008 (2004), [hep-ph/0312266].

[40] J. C. Collins, D. E. Soper and G. Sterman, Nucl. Phys. **B250**, 199 (1985).

[41] C. T. H. Davies and W. J. Stirling, Nucl. Phys. **B244**, 337 (1984).

[42] G. A. Ladinsky and C. P. Yuan, Phys. Rev. **D50**, 4239 (1994), [hep-ph/9311341].

[43] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, JHEP **07**, 001 (2003), [hep-ph/0206293].

[44] F. Caravaglios, M. L. Mangano, M. Moretti and R. Pittau, Nucl. Phys. **B539**, 215 (1999), [hep-ph/9807570].

[45] J. Campbell and R. K. Ellis, Phys. Rev. **D65**, 113007 (2002), [hep-ph/0202176].

[46] CDF, F. Abe *et al.*, Phys. Rev. Lett. **77**, 448 (1996), [hep-ex/9603003].

[47] CDF, F. Abe *et al.*, Phys. Rev. Lett. **79**, 4760 (1997), [hep-ex/9709016].

[48] F. A. Berends, H. Kuijf, B. Tausk and W. T. Giele, Nucl. Phys. **B357**, 32 (1991).

[49] S. D. Ellis, R. Kleiss and W. J. Stirling, Phys. Lett. **B154**, 435 (1985).

[50] F. A. Berends, W. T. Giele, H. Kuijf, R. Kleiss and W. J. Stirling, Phys. Lett. **B224**, 237 (1989).

[51] CDF, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995), [hep-ex/9503002].

[52] D0, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995), [hep-ex/9503003].

[53] A. Pukhov *et al.*, hep-ph/9908288.

[54] CompHEP, E. Boos *et al.*, Nucl. Instrum. Meth. **A534**, 250 (2004), [hep-ph/0403113].

[55] S. Tsuno, T. Kaneko, Y. Kurihara, S. Odaka and K. Kato, Comput. Phys. Commun. **175**, 665 (2006), [hep-ph/0602213].

[56] T. Stelzer and W. F. Long, Comput. Phys. Commun. **81**, 357 (1994), [hep-ph/9401258].

[57] F. Maltoni and T. Stelzer, JHEP **02**, 027 (2003), [hep-ph/0208156].

[58] F. Krauss, R. Kuhn and G. Soff, JHEP **02**, 044 (2002), [hep-ph/0109036].

[59] T. Gleisberg *et al.*, Czech. J. Phys. **55**, B529 (2005), [hep-ph/0409122].

[60] T. Gleisberg *et al.*, JHEP **02**, 056 (2004), [hep-ph/0311263].

[61] G. Marchesini and B. R. Webber, Nucl. Phys. **B310**, 461 (1988).

[62] G. Corcella *et al.*, JHEP **01**, 010 (2001), [hep-ph/0011363].

[63] J. Campbell, R. K. Ellis and D. L. Rainwater, Phys. Rev. **D68**, 094021 (2003), [hep-ph/0308195].

[64] R. K. Ellis and S. Veseli, Phys. Rev. **D60**, 011501 (1999), [hep-ph/9810489].

[65] J. M. Campbell and R. K. Ellis, Phys. Rev. **D62**, 114012 (2000), [hep-ph/0006304].

[66] F. Febres Cordero, L. Reina and D. Wackeroth, Phys. Rev. **D74**, 034007 (2006), [hep-ph/0606102].

[67] C. Balazs and C. P. Yuan, Phys. Rev. **D56**, 5558 (1997), [hep-ph/9704258].

[68] F. Landry, R. Brock, P. M. Nadolsky and C. P. Yuan, Phys. Rev. **D67**, 073016 (2003), [hep-ph/0212159].

[69] T. Sjostrand, S. Mrenna and P. Skands, JHEP **05**, 026 (2006), [hep-ph/0603175].

[70] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).

[71] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[72] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).

[73] V. V. Sudakov, Sov. Phys. JETP **3**, 65 (1956).

[74] H. Baer and M. H. Reno, Phys. Rev. **D44**, 3375 (1991).

[75] M. H. Seymour, Comp. Phys. Commun. **90**, 95 (1995), [hep-ph/9410414].

[76] J. Andre and T. Sjostrand, Phys. Rev. **D57**, 5767 (1998), [hep-ph/9708390].

[77] M. Bengtsson and T. Sjostrand, Phys. Lett. **B185**, 435 (1987).

[78] G. Miu and T. Sjostrand, Phys. Lett. **B449**, 313 (1999), [hep-ph/9812455].

[79] G. Corcella and M. H. Seymour, Nucl. Phys. **B565**, 227 (2000), [hep-ph/9908388].

[80] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, JHEP **11**, 063 (2001), [hep-ph/0109231].

[81] F. Krauss, JHEP **08**, 015 (2002), [hep-ph/0205283].

[82] S. Hoche *et al.*, hep-ph/0602031.

[83] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).

[84] A. Schalicke and F. Krauss, JHEP **07**, 018 (2005), [hep-ph/0503281].

[85] L. Lonnblad, Comput. Phys. Commun. **71**, 15 (1992).

[86] G. Gustafson, Phys. Lett. **B175**, 453 (1986).

[87] G. Gustafson and U. Pettersson, Nucl. Phys. **B306**, 746 (1988).

[88] S. Mrenna and P. Richardson, JHEP **05**, 040 (2004), [hep-ph/0312274].

[89] M. L. Mangano, M. Moretti and R. Pittau, Nucl. Phys. **B632**, 343 (2002), [hep-ph/0108069].

[90] S. Frixione and B. R. Webber, JHEP **06**, 029 (2002), [hep-ph/0204244].

[91] F. Krauss, A. Schalicke, S. Schumann and G. Soff, Phys. Rev. **D70**, 114009 (2004), [hep-ph/0409106].

[92] F. Krauss, A. Schalicke, S. Schumann and G. Soff, Phys. Rev. **D72**, 054017 (2005), [hep-ph/0503280].

[93] OPAL, R. Akers *et al.*, Z. Phys. **C63**, 197 (1994).

[94] J. Pumplin *et al.*, JHEP **07**, 012 (2002), [hep-ph/0201195].

[95] CDF, R. Field, Acta Phys. Polon. **B36**, 167 (2005).

[96] CDF, A. A. Affolder *et al.*, Phys. Rev. **D65**, 092002 (2002).

[97] S. Alekhin, Eur. Phys. J. **C10**, 395 (1999), [hep-ph/9611213].

[98] M. Botje, Eur. Phys. J. **C14**, 285 (2000), [hep-ph/9912439].

[99] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C28**, 455 (2003), [hep-ph/0211080].

[100] J. Pumplin *et al.*, Phys. Rev. **D65**, 014013 (2002), [hep-ph/0101032].

[101] D. Stump *et al.*, JHEP **10**, 046 (2003), [hep-ph/0303013].

[102] W. K. Tung *et al.*, JHEP **02**, 053 (2007), [hep-ph/0611254].

[103] Particle Data Group, W. M. Yao *et al.*, J. Phys. **G33**, 1 (2006).

[104] CTEQ, H. L. Lai *et al.*, Eur. Phys. J. **C12**, 375 (2000), [hep-ph/9903282].

[105] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Phys. Lett. **B531**, 216 (2002), [hep-ph/0201127].

[106] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Phys. Lett. **B604**, 61 (2004), [hep-ph/0410230].

[107] UA1, C. Albajar *et al.*, Z. Phys. **C44**, 15 (1989).

[108] UA1, C. Albajar *et al.*, Phys. Lett. **B253**, 503 (1991).

[109] UA2, J. Alitti *et al.*, Phys. Lett. **B276**, 365 (1992).

[110] CDF, F. Abe *et al.*, Phys. Rev. Lett. **69**, 28 (1992).

[111] CDF, F. Abe *et al.*, Phys. Rev. **D52**, 2624 (1995).

[112] CDF, F. Abe *et al.*, Phys. Rev. Lett. **76**, 3070 (1996), [hep-ex/9509010].

[113] CDF, F. Abe *et al.*, Phys. Rev. **D59**, 052002 (1999).

[114] D0, S. Abachi *et al.*, Phys. Rev. Lett. **75**, 1456 (1995), [hep-ex/9505013].

[115] D0, B. Abbott *et al.*, Phys. Rev. **D60**, 052003 (1999), [hep-ex/9901040].

[116] D0, B. Abbott *et al.*, Phys. Rev. **D61**, 072001 (2000), [hep-ex/9906025].

[117] D0, B. Abbott *et al.*, Phys. Rev. Lett. **84**, 5710 (2000), [hep-ex/9912065].

[118] CDF, D. Acosta *et al.*, Phys. Rev. Lett. **94**, 091803 (2005), [hep-ex/0406078].

[119] CDF, A. Abulencia *et al.*, hep-ex/0508029.

[120] CDF, D. Acosta *et al.*, Phys. Rev. **D71**, 032001 (2005), [hep-ex/0412071].

[121] D0, V. M. Abazov *et al.*, Nucl. Instrum. Meth. **A565**, 463 (2006), [physics/0507191].

[122] T. Sjostrand, L. Lonnblad and S. Mrenna, hep-ph/0108264.

[123] R. Brun, R. Hagelberg, M. Hansroul and J. C. Lassalle, (1978), CERN-DD-78-2-REV.

[124] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C35**, 325 (2004), [hep-ph/0308087].

[125] D0 Collaboration, D0 Note 4403-CONF.

[126] D0 Collaboration, D0 Note 4573-CONF.

[127] D0 Collaboration, D0 Note 4750-CONF.

[128] D0, V. M. Abazov *et al.*, Phys. Rev. **D71**, 072004 (2005), [hep-ex/0412020].

[129] CDF, D. Acosta *et al.*, Phys. Rev. **D71**, 052002 (2005), [hep-ex/0411059].

[130] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **95**, 151805 (2005), [hep-ex/0504032].

[131] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **95**, 091801 (2005), [hep-ex/0505018].

[132] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **95**, 161602 (2005), [hep-ex/0506063].

[133] CDF, A. Abulencia *et al.*, Phys. Rev. Lett. **95**, 252001 (2005), [hep-ex/0507104].

[134] CDF, A. Abulencia *et al.*, Phys. Rev. Lett. **96**, 211801 (2006), [hep-ex/0602045].

[135] CDF, A. A. Affolder *et al.*, Phys. Rev. Lett. **87**, 131802 (2001), [hep-ex/0106047].

[136] D0, B. Abbott *et al.*, Phys. Rev. Lett. **82**, 4769 (1999), [hep-ex/9812010].

[137] D0, S. Abachi *et al.*, Phys. Lett. **B385**, 471 (1996).

[138] CDF, F. Abe *et al.*, Phys. Rev. Lett. **79**, 2192 (1997).

[139] D0, B. Abbott *et al.*, Phys. Rev. Lett. **86**, 1156 (2001), [hep-ex/0008065].

[140] D0 Collaboration, D0 Note 4757-CONF.

[141] CDF Collaboration, CDF Note 8694.

[142] D0 Collaboration, D0 Note 5195.

[143] CDF, F. Abe *et al.*, Phys. Rev. Lett. **81**, 5754 (1998), [hep-ex/9809001].

[144] CDF, D. Acosta *et al.*, Phys. Rev. **D71**, 051104 (2005), [hep-ex/0501023].

[145] D0 Collaboration, D0 Note 4855-CONF.

[146] Q.-H. Cao and C. P. Yuan, Phys. Rev. Lett. **93**, 042001 (2004), [hep-ph/0401026].

[147] CDF, A. A. Affolder *et al.*, Phys. Rev. **D63**, 011101 (2001), [hep-ex/0006025].

[148] CDF Collaboration, CDF Note (30.11.2006).

[149] D0, V. M. Abazov *et al.*, Phys. Rev. **D76**, 012003 (2007), [hep-ex/0702025].

[150] R. M. Thurman-Keup, A. V. Kotwal, M. Tecchio and A. Byon-Wagner, Rev. Mod. Phys. **73**, 267 (2001).

[151] CDF, A. A. Affolder *et al.*, Phys. Rev. Lett. **84**, 845 (2000), [hep-ex/0001021].

[152] R. K. Ellis and S. Veseli, Nucl. Phys. **B511**, 649 (1998), [hep-ph/9706526].

[153] C. T. H. Davies, B. R. Webber and W. J. Stirling, Nucl. Phys. **B256**, 413 (1985).

[154] D0 Collaboration, D0 Note 5187-CONF.

[155] E. Barberio and Z. Was, Comput. Phys. Commun. **79**, 291 (1994).

[156] S. Berge, P. Nadolsky, F. Olness and C. P. Yuan, Phys. Rev. **D72**, 033015 (2005), [hep-ph/0410375].

[157] W. T. Giele and S. Keller, Phys. Rev. **D57**, 4433 (1998), [hep-ph/9704419].

[158] CDF, F. Abe *et al.*, Phys. Rev. **D45**, 1448 (1992).

[159] G. C. Blazey *et al.*, hep-ex/0005012.

[160] D0, B. Abbott *et al.*, Nucl. Instrum. Meth. **A424**, 352 (1999), [hep-ex/9805009].

[161] A. Bhatti *et al.*, Nucl. Instrum. Meth. **A566**, 375 (2006), [hep-ex/0510047].

[162] D0, V. M. Abazov *et al.*, Phys. Rev. **D74**, 092005 (2006), [hep-ex/0609053].

[163] A. Messina, hep-ex/0610053.

[164] CDF Collaboration, B. Cooper, A. Messina and D. Waters, CDF Note 8381.

[165] CDF, A. A. Affolder *et al.*, Phys. Rev. **D63**, 072003 (2001).

[166] H. L. Lai *et al.*, Phys. Rev. **D51**, 4763 (1995), [hep-ph/9410404].

[167] D0, S. Abachi *et al.*, Phys. Rev. Lett. **75**, 3226 (1995).

[168] W. T. Giele, E. W. N. Glover and D. A. Kosower, Nucl. Phys. **B403**, 633 (1993), [hep-ph/9302225].

[169] CDF, F. Abe *et al.*, Phys. Rev. Lett. **81**, 1367 (1998).

[170] A. D. Martin, W. J. Stirling and R. G. Roberts, Phys. Lett. **B354**, 155 (1995), [hep-ph/9502336].

[171] H. L. Lai *et al.*, Phys. Rev. **D55**, 1280 (1997), [hep-ph/9606399].

[172] CDF, F. Abe *et al.*, Phys. Rev. Lett. **73**, 2296 (1994).

[173] CDF, F. Abe *et al.*, Phys. Rev. Lett. **62**, 3020 (1989).

[174] CDF, F. Abe *et al.*, Phys. Rev. Lett. **71**, 679 (1993).

[175] D0, B. Abbott *et al.*, Phys. Lett. **B464**, 145 (1999), [hep-ex/9908017].

[176] T. Sjostrand, Comput. Phys. Commun. **82**, 74 (1994).

[177] V. A. Khoze and W. J. Stirling, Z. Phys. **C76**, 59 (1997), [hep-ph/9612351].

[178] D0, V. Abazov, hep-ex/0608052.

[179] CDF Collaboration, CDF Note 8827.

[180] R. Field, FERMILAB-CONF-06-408-E. Proceedings of TeV4LHC Workshop - 4th Meeting, FNAL, Batavia, IL, 2006.

[181] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands, hep-ph/0308153.

[182] D0 Collaboration, D0 Note 5066-CONF.

[183] D. L. Rainwater, R. Szalapski and D. Zeppenfeld, Phys. Rev. **D54**, 6680 (1996), [hep-ph/9605444].

[184] S. Asai *et al.*, Eur. Phys. J. **C32S2**, 19 (2004), [hep-ph/0402254].

[185] S. Abdullin *et al.*, Eur. Phys. J. **C39S2**, 41 (2005).

[186] CDF, F. Abe *et al.*, Phys. Rev. **D50**, 2966 (1994).

[187] D0, V. M. Abazov *et al.*, Phys. Lett. **B626**, 35 (2005), [hep-ex/0504058].

[188] D. J. Jackson, Nucl. Instrum. Meth. **A388**, 247 (1997).

[189] CDF, A. Abulencia *et al.*, Phys. Rev. **D74**, 032008 (2006).

[190] M. A. G. Aivazis, J. C. Collins, F. I. Olness and W.-K. Tung, Phys. Rev. **D50**, 3102 (1994), [hep-ph/9312319].

[191] J. C. Collins, Phys. Rev. **D58**, 094002 (1998), [hep-ph/9806259].

[192] J. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, Phys. Rev. **D69**, 074021 (2004), [hep-ph/0312024].

[193] J. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, Phys. Rev. **D75**, 054015 (2007), [hep-ph/0611348].

[194] CDF, D. Acosta *et al.*, Phys. Rev. **D71**, 052003 (2005), [hep-ex/0410041].

[195] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **98**, 181802 (2007), [hep-ex/0612052].

[196] W. T. Giele, S. Keller and E. Laenen, Phys. Lett. **B372**, 141 (1996), [hep-ph/9511449].

[197] F. Maltoni, T. McElmurry and S. Willenbrock, Phys. Rev. **D72**, 074024 (2005), [hep-ph/0505014].

[198] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **94**, 152002 (2005), [hep-ex/0411084].

[199] D0, V. M. Abazov *et al.*, Phys. Rev. **D74**, 112004 (2006), [hep-ex/0611002].

[200] S. Greder, (2004), FERMILAB-THESIS-2004-28.

[201] CDF, A. Abulencia *et al.*, Phys. Rev. **D74**, 072006 (2006), [hep-ex/0607035].

[202] T. Scanlon, (2006), FERMILAB-THESIS-2006-43.

[203] CDF Collaboration, CDF Note 8272.

[204] H1, A. Aktas *et al.*, Eur. Phys. J. **C40**, 349 (2005), [hep-ex/0411046].

[205] H1, A. Aktas *et al.*, Eur. Phys. J. **C45**, 23 (2006), [hep-ex/0507081].

[206] T. Stelzer, Z. Sullivan and S. Willenbrock, Phys. Rev. **D56**, 5919 (1997), [hep-ph/9705398].

[207] C. S. Huang and S.-H. Zhu, Phys. Rev. **D60**, 075012 (1999), [hep-ph/9812201].

[208] J. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, Phys. Rev. **D67**, 095002 (2003), [hep-ph/0204093].

[209] D. Dicus, T. Stelzer, Z. Sullivan and S. Willenbrock, Phys. Rev. **D59**, 094016 (1999), [hep-ph/9811492].

[210] C. Balazs, H.-J. He and C. P. Yuan, Phys. Rev. **D60**, 114001 (1999), [hep-ph/9812263].

[211] F. Maltoni, Z. Sullivan and S. Willenbrock, Phys. Rev. **D67**, 093005 (2003), [hep-ph/0301033].

[212] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **94**, 161801 (2005), [hep-ex/0410078].

[213] CDF, A. Abulencia *et al.*, Phys. Rev. **D74**, 032008 (2006), [hep-ex/0605099].

[214] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **94**, 091802 (2005), [hep-ex/0410062].

[215] CDF Collaboration, CDF Note 8410.

[216] CDF, D. Acosta *et al.*, Phys. Rev. **D65**, 052007 (2002), [hep-ex/0109012].

[217] CDF, D. Acosta *et al.*, Phys. Rev. **D69**, 072004 (2004), [hep-ex/0311051].

[218] CDF, A. Abulencia *et al.*, Phys. Rev. **D73**, 051101 (2006), [hep-ex/0512065].

[219] D0, V. M. Abazov *et al.*, Phys. Rev. **D74**, 011104 (2006), [hep-ex/0606018].

[220] U. Baur, M. Spira and P. M. Zerwas, Phys. Rev. **D42**, 815 (1990).

[221] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32**, 438 (1974).

[222] J. C. Pati and A. Salam, Phys. Rev. **D10**, 275 (1974).

[223] E. Eichten, I. Hinchliffe, K. D. Lane and C. Quigg, Phys. Rev. **D34**, 1547 (1986).

[224] W. Buchmuller and D. Wyler, Phys. Lett. **B177**, 377 (1986).

[225] E. Eichten, K. D. Lane and M. E. Peskin, Phys. Rev. Lett. **50**, 811 (1983).

[226] D0, V. M. Abazov *et al.*, Phys. Rev. **D71**, 071104 (2005), [hep-ex/0412029].

[227] CDF, D. Acosta *et al.*, Phys. Rev. **D72**, 051107 (2005), [hep-ex/0506074].

[228] D0, V. M. Abazov *et al.*, Phys. Lett. **B636**, 183 (2006), [hep-ex/0601047].

[229] D0 Collaboration, D0 Note 5370-CONF.

[230] CDF, A. Abulencia *et al.*, Phys. Rev. **D73**, 051102 (2006), [hep-ex/0512055].

[231] D0 Collaboration, D0 Note 5447-CONF.

[232] CDF Collaboration, CDF Note 8309.

[233] D0, V. M. Abazov *et al.*, Phys. Rev. **D64**, 092004 (2001), [hep-ex/0105072].

[234] M. Kramer, T. Plehn, M. Spira and P. M. Zerwas, Phys. Rev. Lett. **79**, 341 (1997), [hep-ph/9704322].

[235] D0, B. Abbott *et al.*, Phys. Rev. Lett. **84**, 2088 (2000), [hep-ex/9910040].

[236] O. Brein, A. Djouadi and R. Harlander, Phys. Lett. **B579**, 149 (2004), [hep-ph/0307206].

[237] CDF, A. Abulencia *et al.*, Phys. Rev. Lett. **96**, 081803 (2006), [hep-ex/0512051].

[238] D0 Collaboration, D0 Note 5054-CONF.

[239] D0 Collaboration, D0 Note 5357-CONF.

[240] D0 Collaboration, D0 Note 5365-CONF.

[241] CDF Collaboration, CDF Note 8240.

[242] CDF Collaboration, CDF Note 8390.

[243] D0 Collaboration, D0 Note 5380-CONF.

[244] D0, V. M. Abazov *et al.*, arXiv:0704.2000 [hep-ex].

[245] N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali, Phys. Lett. **B429**, 263 (1998), [hep-ph/9803315].

[246] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali, Phys. Lett. **B436**, 257 (1998), [hep-ph/9804398].

[247] N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali, Phys. Rev. **D59**, 086004 (1999), [hep-ph/9807344].

[248] G. F. Giudice, R. Rattazzi and J. D. Wells, Nucl. Phys. **B544**, 3 (1999), [hep-ph/9811291].

[249] E. A. Mirabelli, M. Perelstein and M. E. Peskin, Phys. Rev. Lett. **82**, 2236 (1999), [hep-ph/9811337].

[250] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **90**, 251802 (2003), [hep-ex/0302014].

[251] CDF, D. Acosta, Phys. Rev. Lett. **92**, 121802 (2004), [hep-ex/0309051].

[252] CDF, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 171802 (2006), [hep-ex/0605101].

[253] CDF, D. Acosta *et al.*, Phys. Rev. **D71**, 112001 (2005), [hep-ex/0410076].

[254] D0, V. M. Abazov *et al.*, Phys. Lett. **B640**, 230 (2006), [hep-ex/0607009].

[255] D0, V. M. Abazov *et al.*, Phys. Lett. **B638**, 119 (2006), [hep-ex/0604029].

[256] J. R. Ellis and S. Rudaz, Phys. Lett. **B128**, 248 (1983).

[257] K.-I. Hikasa and M. Kobayashi, Phys. Rev. **D36**, 724 (1987).

[258] CDF, A. A. Affolder *et al.*, Phys. Rev. Lett. **84**, 5704 (2000), [hep-ex/9910049].

[259] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **93**, 011801 (2004), [hep-ex/0404028].

[260] D0, V. M. Abazov *et al.*, Phys. Lett. **B645**, 119 (2007), [hep-ex/0611003].

[261] D0, V. M. Abazov *et al.*, arXiv:0705.0812 [hep-ex].

[262] CDF, F. Abe *et al.*, Phys. Rev. Lett. **78**, 2906 (1997).

[263] CDF, A. A. Affolder *et al.*, Phys. Rev. Lett. **85**, 2056 (2000), [hep-ex/0004003].

[264] D0, B. Abbott *et al.*, Phys. Rev. Lett. **81**, 38 (1998), [hep-ex/9803009].

[265] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **97**, 161803 (2006), [hep-ex/0607022].

[266] CDF Collaboration, CDF Note 8442.

[267] D0 Collaboration, D0 Note 5353-CONF.

[268] ATLAS Collaboration, CERN/LHCC/99-15.

[269] CMS Collaboration, CERN/LHCC/2006/021.

[270] V. Buscher and K. Jakobs, Int. J. Mod. Phys. **A20**, 2523 (2005), [hep-ph/0504099].

[271] H. Baer, C.-h. Chen, F. Paige and X. Tata, Phys. Rev. **D52**, 2746 (1995), [hep-ph/9503271].

[272] H. Baer, C.-h. Chen, F. Paige and X. Tata, Phys. Rev. **D53**, 6241 (1996), [hep-ph/9512383].

[273] D. R. Tovey, Eur. Phys. J. direct **C4**, N4 (2002).

[274] CMS, S. Abdullin *et al.*, J. Phys. **G28**, 469 (2002), [hep-ph/9806366].

[275] V. A. Mitsou, N. C. Benekos, I. Panagoulias and T. D. Papadopoulou, Czech. J. Phys. **55**, B659 (2005), [hep-ph/0411189].

[276] CMS, S. Abdullin, F. Charles and F. Luckel, CMS Note 1999/027.

Acknowledgements

This work would not have been possible to undertake without the support and advice of many colleagues. I would like to thank Dorothee Schaile, Otmar Biebel, and Johannes Elmsheuser for helpful comments on the manuscript. I want to express my gratitude to Dorothee for making it possible for me to work on this subject over the past years and for her continuous support. I thank Otmar for many discussions and a lot of advice. I enjoyed working with my DØ colleagues at Munich, many thanks to: Britta Tiller, Philippe Calfayan, Johannes Elmsheuser, Daniela Görisch, Tim Christiansen, Meta Binder, Raimund Ströhmer, Frank Fiedler, Petra Haefner, Philipp Schieferdecker, and Alexander Grohsjean. Finally, I am grateful to Sabine and Lorenz for their support, patience, and forbearance.