Institut de Física
d'Altes Energies

Universitat Autònoma de Barcelona

Measurement of Inclusive Jet Cross Sections in $Z/\gamma^*(\rightarrow e^+e^-) + $ jets Production in $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV with the CDF Detector$^{(1)}$

Oriol Saltó Bauzà
Institut de Física d'Altes Energies
Universitat Autònoma de Barcelona
Departament de Física
Edifici Cn, Campus UAB
E-08193 Bellaterra (Barcelona)

April 2008

supervised by
Prof. Mario Martínez Pérez
ICREA / Institut de Física d'Altes Energies
and
Universitat Autònoma de Barcelona

---

[1] Ph.D. Dissertation

The studies presented in this thesis led to the following publication:

- Measurement of Inclusive Jet Cross Sections in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ Production in $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV, Phys. Rev. Lett. **100**, 102001 (2008).

Acknowledgments

The work reported in this Ph.D. thesis involved 3 years of my life, 2 continents, 3 notebooks, several terabytes of disk space, and interacting with many people.

Undoubtedly, the person that most influenced this thesis is my supervisor Mario Martínez. His broad knowledge of physics and QCD in particular brought endless inputs to this thesis, and his commitment to the well-done work was a constant guidance in the processes of the analysis. Monica is the person who day by day helped me in every step of the analysis. I would like to thank her for creating such a great and friendly working environment. Alon has also helped me in the final stages of this work providing useful ideas and cross checks. Thanks also to Regis for asking and answering so many questions, his aim for perfection surely had an impact in my work.

All this was possible thanks to Enrique Fernández who gave me the opportunity to work at IFAE. Special thanks go to Matteo Cavalli-Sforza who warmly welcomed me at IFAE, for sharing his neverending desire for learning and passion for physics. I would also like to express my gratitude to Ilya Korolkov, my Master Thesis supervisor, who had to deal with my learning curve in the research world and always supported and encouraged me.

I would like to thank all the CDF collaborators, specially the spokespeople and the QCD conveners, for their help and inputs to the analysis, and for the success of the CDF experiment.

Tota aquesta feina no hagués estat possible sense els companys i amics de l'IFAE. Primer de tot al Xavi i l'Olga, per ésser els meus amfitrions a l'IFAE, que amb el temps han esdevingut més que amics. Ells em van ensenyar tot el que calia saber per sobreviure fora de la nostra estimda terra, primer al CERN, i després als Estat Units. La seva facilitat per conèixer nova gent i fer amics allí on van han fet que sempre em sentís acompanyat. I també a la Carolina, amb qui després de conèixer-nos durant la carrera vam tornar a trobar-nos a l'IFAE, el seu esperit sempre alegre i la seva rialla encomanadissa han ajudat a passar el moments més durs. We had a lot of physics and non-physics discussions with Gianluca, we shared problems and solutions, and he has been a really good friend and office-mate. I overlapped some months with Stefano in Barcelona. We did not have much time to know each other but I wish him good luck with his thesis.

The friends that I met at Fermilab: Ernest, Óscar, Jónatan, Cristóbal, Kike, Miguel, Bárbara, Bruno, Arán, Ignacio, Roberto y alguno más que seguro que me olvido. Gracias a todos por crear un ambiente tan familiar y acogedor que me ha hecho sentir como en casa. Gracias sobretodo por los buenos momentos pasados fuera del trabajo, por las partidas de Go, las maratones cinematográficas, el volei, la piscina y un sinfín de cosas que nos permitian 'desconectar'.

També voldria donar les gràcies als meus amics del Cau, per no perdre el contacte durant tot aquest temps que he estat fora i per deixar que m'encarregués d'alguns aspectes informàtics des de la distància que m'han permès sentir-me part del Cau fins i tot des de tant lluny.

Gràcies de tot cor als meus pares i a les meves germanes, la Laia i la Núria, per haver-me donat suport en tot moment i animar-me sempre a fer el que més m'agradés. I finalment a la Vicky per haver-me estimat tot aquest temps i haver tingut infinita paciència amb mi, per les cartes, les llargues converses per teléfon i les inoblidables vacances passades en terres americanes.

Contents

1 Introduction **1**

2 Standard Model and QCD **3**

2.1 Standard Model . . . 3

2.2 QCD Physics . . . 5

2.3 Parton Model . . . 6

2.4 QCD Factorization . . . 8

2.5 Parton Distribution Functions . . . 10

2.5.1 PDF Parametrization . . . 11

2.6 Boson + jets pQCD Predictions . . . 13

2.6.1 Dependence on the Scale . . . 15

2.6.2 PDF Uncertainties . . . 15

2.7 QCD Phenomenology and Monte Carlo Simulation . . . 16

2.7.1 Parton Shower . . . 17

2.7.2 Hadronization . . . 18

2.7.2.1 String Model . . . 18

2.7.2.2 Cluster Model . . . 19

2.7.3 Underlying Event . . . 20

2.8 Monte Carlo Generator Programs . . . 20

2.8.1 PYTHIA Monte Carlo . . . 20

2.8.1.1 TUNE A Underlying Event Parametrization . . . 21

2.8.1.2 TUNE DW Underlying Event Parametrization . . . 21

2.8.2 HERWIG Monte Carlo 21
2.8.2.1 JIMMY Underlying Event Parametrization 21
2.8.3 ALPGEN Monte Carlo 22
2.8.4 Other Monte Carlo Generator Programs 22
2.9 Jet Algorithms 23
2.9.1 Cone Algorithms 24
2.9.2 Cluster Algorithms 26

3 CDF at Fermilab **27**
3.1 The Tevatron 27
3.2 The CDF Run II Detector 28
3.2.1 Tracking Systems 29
3.2.1.1 Silicon Detectors 29
3.2.1.2 Central Outer Tracker 31
3.2.2 Time-of-Flight System 33
3.2.3 Calorimeters 33
3.2.3.1 Central Calorimeters 33
3.2.3.2 Plug Calorimeters 35
3.2.3.3 Shower Profile Detectors 36
3.2.4 Muon Detectors 36
3.2.5 DAQ and Triggers 36
3.2.6 Luminosity Measurement 39
3.3 Reconstruction of Physics Objects 40
3.3.1 Electrons 41
3.3.2 Jets 41
3.3.2.1 Corrections of the Jets Transverse Momentum 43

4 Inclusive Jet Cross Sections in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ Production **47**
4.1 Definition of the Dataset 47
4.1.1 Electron Trigger Efficiency 47

4.2 Monte Carlo Samples 49
4.3 Electron Selection and Z/γ^* Reconstruction 50
    4.3.1 Event Trigger Efficiency 50
    4.3.2 Electron Energy Corrections 51
    4.3.3 Event Selection 52
    4.3.4 Electron ID Efficiency 54
    4.3.5 Inclusive $Z/\gamma^* \to e^+e^-$ Cross Section 59
4.4 Jet Reconstruction and Selection 60
    4.4.1 Dependence on Instantaneous Luminosity 62
4.5 Background Estimation 63
    4.5.1 QCD and $W(\to e + \nu) + \text{jets}$ Backgrounds 63
        4.5.1.1 Fake Rate 64
        4.5.1.2 QCD and $W(\to e + \nu) + \text{jets}$ Background Estimation 66
    4.5.2 Non-QCD Backgrounds 66
    4.5.3 Validation of the Background Estimation 67
4.6 Unfolding 69
4.7 Systematic Uncertainties 72
    4.7.1 Correlations between Systematic Uncertainties 76

5 Jet Shapes and Energy Flows in $Z/\gamma^*(\to e^+e^-) + \text{jets}$ **79**
5.1 Jet Shapes 79
    5.1.1 Systematic Uncertainties on the Jet Shapes 81
    5.1.2 Underlying Event in Monte Carlo Generators 82
5.2 Energy Flow in $Z/\gamma^*(\to e^+e^-) + \text{jets}$ Events 85

6 Results **89**
6.1 NLO pQCD Calculation 89
    6.1.1 Non-pQCD Contributions 89
    6.1.2 PDF Uncertainty on the Theoretical Predictions 90
    6.1.3 Dependence on Renormalization and Factorization Scale 91

6.1.4 Uncertainty on the non-pQCD Corrections 93
6.2 Comparison with theoretical predictions 93
6.3 Comparison to LO Matrix Elements + Parton Shower Predictions 105
6.3.1 Jet Multiplicity 105
6.3.2 Jet Topologies in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets Events 105

7 Summary and Conclusions **111**

A Data Quality Monitoring **113**
A.1 Online DQM System 113
A.1.1 DQMon Monitor 114
A.2 Offline DQM System 116
A.2.1 Partial Run Recovery 116

B Breakdown of the Unfolding Factors **117**

References **119**

List of Figures **125**

List of Tables **133**

Chapter 1

Introduction

Quantum Chromodynamics (QCD) [1] is the field theory that describes the strong interactions between quarks and gluons. The most prominent signature of QCD at hadron colliders is the production of collimated jets of hadrons. The measurement of the production of such jets in association with a vector boson, W or Z/γ^*, provides an stringent test of perturbative QCD (pQCD) calculations. Furthermore, some of new physics processes at hadron colliders, such as the production of Higgs bosons and supersymmetric particles, can be mimicked by the production of vector bosons in association with jets that constitute irreducible backgrounds to these searches. Therefore, the study and understanding of Z/γ^* +jets processes is a crucial part of the physics program of the Tevatron collider. At the Run I of the Tevatron, protons and antiprotons collided with an energy in the center-of-mass $\sqrt{s}$ of 1.8 TeV. $Z/\gamma^* +$ jets production was studied during Run I with the CDF detector using 106 pb^{-1} of data [2]. The measurements were compared to the predictions from leading order (LO) plus parton shower Monte Carlo generator programs. The comparisons were mainly dominated by the large scale uncertainties in the LO predictions. In Run II, The Tevatron and the CDF detector were upgraded and, among other important improvements, $\sqrt{s}$ was increased to 1.96 TeV. The good performance of the accelerator and the detector made available a larger amount of data, making possible precise differential measurements.

This Ph.D. thesis presents the measurement of inclusive jet cross sections in $Z/\gamma^* \rightarrow e^+e^-$ events using 1.7 fb^{-1} of data collected by the upgraded CDF detector during the Run II of the Tevatron. The Midpoint cone algorithm is used to search for jets in the events after identifying the presence of a Z/γ^* boson through the reconstruction of its decay products. The measurements are compared to next-to-LO (NLO) pQCD predictions for events with one and two jets in the final state. The perturbative predictions are corrected for the contributions of non-perturbative processes, like the underlying event and the fragmentation of the partons into jets of hadrons. These processes are not described by perturbation theory and must be estimated from phenomenological models. In this thesis, a number of measurements are performed to test different models of underlying event and hadronization implemented in LO plus parton shower Monte Carlo generator programs.

Chapter 2 is devoted to the description of the theory of strong interactions and jet phenomenology at hadron colliders. Chapter 3 contains the description of the Tevatron collider and the CDF detector. The analysis is described in detail in Chapter 4. Chapter 5 shows the measurement of those observables sensitive to non-perturbative effects compared to the predictions from several Monte Carlo programs. Chapter 6 discusses the final results and the comparison with theoretical expectations. Finally, Chapter 7 is devoted to the conclusions.

Chapter 2

Standard Model and QCD

2.1 Standard Model

The Standard Model is the most successful way to describe the elementary particles that constitute matter and their interactions at the subatomic level [3]. The framework of the Standard Model consists of a quantum field theory based in the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$ that describes the electromagnetic and weak interactions (electroweak sector based in the $SU(2)_L \times U(1)_Y$ group) and the strong interactions (QCD based in the $SU(3)_C$ group). The fourth force in Nature, gravity, is not included in the Standard Model and it is described by the Newton equations in classical physics refined by the Einstein description in General Relativity.

In the Standard Model, forces are mediated by gauge bosons (Table I). Bosons have an integer spin (1) and obey Bose-Einstein statistics. The electromagnetic force is mediated via the exchange of massless photons γ. The weak force is transmitted by exchange of three massive vector bosons: W^+, W^- and Z. Finally, the strong force is mediated by eight massless gluons.

Gauge bosons	Mediator interaction	Charge
γ	Electromagnetic	0
$W^\pm$	Weak	± 1
Z		0
8 gluons	Strong	0

Table I: Gauge bosons of the Standard Model, mediators of the fundamental interactions.

Besides bosons, there are elementary particles with half-integer spin ($\frac{1}{2}$) that follow Fermi-Dirac statistics: the fermions (Table II). There are two fundamentally different types of fermions: leptons and quarks. Both leptons and quarks can interact electroweakly, but only quarks feel the strong force. Fermions are classified in three generations or families. The lepton sector is formed by electrons e, muons μ, taus τ and their associated neutrinos ν_e, ν_μ and ν_τ. There are six types or *flavors* of quarks: up u, down d, strange s, charm c, bottom b and top t. Quarks have fractional electric charge: $Q = +2/3$ (up-type quarks) or $Q = -1/3$ (down-type quarks). They are not found isolated in Nature, only in bound states of quark-antiquark pairs (*mesons*) or three quarks or antiquarks (*baryons*).

	Generations			
	I	II	III	Charge
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$	$\pm 2/3$ $\mp 1/3$
Leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$	0 ± 1

Table II: Elementary fermions of the Standard Model and their electric charge.

The Standard Model formalism is written for massless particles and the Higgs mechanism of spontaneous symmetry breaking is proposed for generating non-zero boson and fermion masses [4]. The symmetry breaking requires the introduction of a new field that leads to the existence of a new massive boson, the Higgs boson, for which there is still no empirical evidence.

Despite its tremendous success in describing the existence and properties of particles and interactions, and that no experiment has been able to find any clear deviation from its predictions, the Standard Model contains 19 free parameters that need to be determined experimentally and it has some theoretical limitations that push the belief that the Standard Model is an “effective” theory of a much larger theory.

Several extensions of the Standard Model have been proposed. One of the most popular extensions is SuperSymmetry (SUSY) [5]. SUSY introduces another symmetry between bosons and fermions that doubles the spectrum of particles assigning one *sparticle* (SUSY particle) to every SM particle with a difference of $\frac{1}{2}$ in the spin. For every spin-$\frac{1}{2}$ fermion there is a spin-0 *sfermion* and for every spin-1 boson there is a spin-$\frac{1}{2}$ *sboson*. SUSY is broken at the TeV scale as we do not observed sparticles in Nature. There exist several SUSY models with different mechanisms to broke the symmetry and different parameters to be determined. The Minimal SuperSymmetric Standard Model (MSSM) [6] has more than hundred free parameters that need to be constrained by the measurements, and therefore is not very predictive. The Minimal SuperGravity (mSUGRA) [7] model includes gravity and unifies the masses of the Higgs sector, the sfermion sector, as well as the couplings, at GUT scale energies (10^{15} GeV). mSUGRA has only five free parameters which allow to predict all particles and couplings. Most of the extensions of the Standard Model predict the existence of new particles, usually with high masses. These particles could only be produced in hadron colliders, where the physics processes are dominated by a large background of QCD.

2.2 QCD Physics

Quantum Chromodynamics (QCD) is the gauge field theory that describes the strong interactions in the Standard Model. It was introduced by Gell-Mann and Fritzsch in 1972 [8] as a renormalizable non-abelian theory based on the $SU(3)$ group representation [9]. The corresponding Lagrangian, that contains quark q_k and gluon $\mathcal{A}^A_\alpha$ fields, is given by

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4} F^A_{\alpha\beta} F^{\alpha\beta}_A + \sum_j^{flavors} \bar{q}_j (i\gamma^\mu D_\mu - m_j)\, q_j \tag{2.1}$$

where $F^A_{\alpha\beta}$ is the field strength tensor for the spin-1 gluon field $\mathcal{A}^A_\alpha$,

$$F^A_{\alpha\beta} = \partial_\alpha \mathcal{A}^A_\beta - \partial_\beta \mathcal{A}^A_\alpha - g f^{ABC} \mathcal{A}^B_\alpha \mathcal{A}^C_\beta, \tag{2.2}$$

$D_\mu = \partial_\mu + ig\mathcal{A}^\alpha_\mu t^\alpha$ is the covariant derivative, t^α are the Gell-Mann matrices and f^{ABC} are the structure constants of $SU(3)$. Capital indexes A, B, C run over the 8 degrees of freedom of the gluon fields and index j runs over the quark flavors. The charge associated to the strong interaction is the *color*. The need of a color charge is introduced by the fact that quarks, as fermions, must satisfy the Pauli exclusion principle(2). The validity of the color quantum numbers and the choice of $SU(N_\mathrm{C} = 3)_{\mathrm{color}}$ is confirmed by many measurements that directly probe the value of N_C [10]. The physical vertices include the $gq\bar{q}$ vertex, analogous to the $\gamma f\bar{f}$ coupling in Quantum Electrodynamics, but also 3-gluon and 4-gluon vertices, described by the third term in the $F^A_{\alpha\beta}$ tensor, proper of the non-abelian nature of the theory. It is the non-abelian nature of the QCD that leads to two important characteristics of the strong interaction: asymptotic freedom and confinement.

---

[2] For instance, baryon N^{*++} with z-component spin $+3/2$ is made up of three up quarks ($u^\uparrow u^\uparrow u^\uparrow$) in an s-state. Its wave function is totally symmetric in space, in spin and in flavor, so that complete antisymmetry in color is required by Fermi statistics. This requirement is satisfied by $\epsilon_{abc} q^a q^b q^c$ where a, b and c are the $SU(3)_{\mathrm{color}}$ indexes.

The strong coupling constant $\alpha_s = \frac{g^2}{4\pi}$ changes with the scale of the interaction. At high energies (small distances) α_s decreases (Fig. 2.1). Then, strong interactions proceed via color fields of reduced strength and quarks and gluons behave as essentially free particles. This situation is called asymptotic freedom and is totally supported by the results from deep inelastic scattering (DIS) experiments (described in the next Section 2.3). It is important to notice that the asymptotic freedom is the basis of the perturbative approach to calculate QCD observables (pQCD). On the other hand, the coupling strength asymptotically diverges at low energies (large distances) making therefore impossible to produce isolated quarks. When in a $q\bar{q}$ pair, the quarks begin to separate from each other, the energy of the field between them increases. At some point, it is energetically favorable to create an additional $q\bar{q}$ pair. At the end, there are only colorless bound states (hadrons). This situation is called confinement and it is related to the process of jet formation (Section 2.7.2).

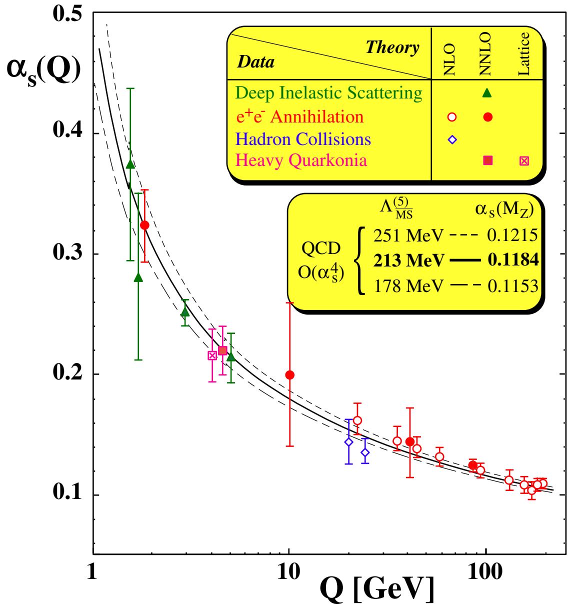


Figure 2.1: Value of the running strong coupling α_s as a function of the energy scale.

2.3 Parton Model

Deep inelastic scattering (DIS) experiments performed in the 1960s to understand the internal structure of the proton and the neutron led to the Quark Parton Model, predecessor of QCD. Electrons with energies up to 20 GeV were sent against a target of hydrogen:

$$e + P \rightarrow e + X \tag{2.3}$$

where P in the proton on the hydrogen nucleus and X is any hadronic final state (Fig. 2.2).

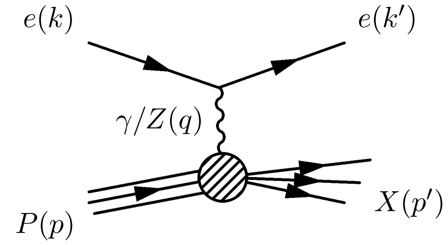


Figure 2.2: Schematic view of a deep inelastic scattering event: the photon (or a Z boson) interacts with a parton inside the proton.

At a given center of mass energy $\sqrt{s}$ the kinematics of DIS can be described by

$$Q^2 = -q^2 = (k - k')^2 \qquad\qquad x = \frac{Q^2}{2(p \cdot q)} \tag{2.4}$$

where k and k' are the four momentum of the incoming and outgoing electrons, p the momentum of the incoming proton and x is interpreted as the fraction of the proton momentum carried by the interacting quark. Figure 2.3 shows the $Q^2 - x$ range accessible by different fixed-target and collider experiments.

The results of the first DIS experiments were a larger number of large-angle deflected electrons than expected. A phenomenological explanation to these results was given considering the proton a compound of non-interacting point-like particles, the *partons*. Therefore, eP collisions can be regarded as 'hard' interactions between the electron and partons inside the proton. The Parton Model considered nucleons as bound states of three partons, each carrying a fraction x of the total nucleon momentum such that

$$\sum_{\text{partons}} x_p = 1 \tag{2.5}$$

In the Parton Model, the total cross section can be expressed in terms of electron-parton ep interaction cross section:

$$\sigma(eP \to eX) = f \otimes \sigma(ep \to ep) \tag{2.6}$$

where f is the parton density, also called parton distribution function (PDF). The term $f_i(x)dx$ gives the probability of finding a parton of type i in the proton carrying a fraction between x and $x + dx$ of the proton momentum. The Parton Model predicted that in the infinite-momentum frame of the proton, i.e. when $Q^2 \to \infty$ and the transverse momentum of the partons inside the proton are small, the parton densities were only a function of x. This behavior was called Bjorken scaling [11]. However, in QCD, the radiation of gluons from the quarks leads to a violation of the scaling predicted by the Parton Model. Then, the DIS cross section is written as:

$$\frac{d^2\sigma(e^{\pm}P)}{dx\,dQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left(y^2 xF_1(x,Q^2) + (1-y)F_2(x,Q^2) \mp y(1-y)xF_3(x,Q^2)\right) \tag{2.7}$$

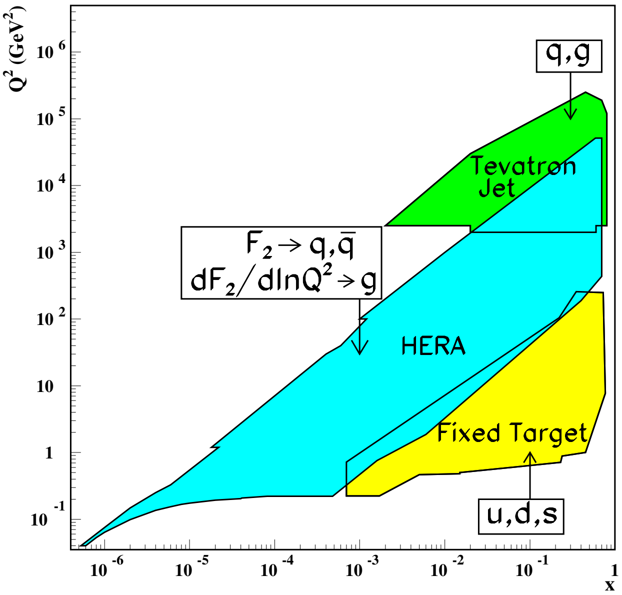


Figure 2.3: Kinematic domains in x and Q^2 probed by fixed-target and collider experiments shown together with the important constrains they make on the various parton distribution functions.

where y is the ratio $\frac{Q^2}{sx}$ and F_i are the proton structure functions defined as:

$$F_1 = \frac{1}{2}\sum_i e_i^2 f_i \qquad F_2 = \sum_i e_i^2 x f_i. \tag{2.8}$$

that explicitly depend on Q^2. As Q^2 increases, more gluons are radiated off the quarks, which split into $q\bar{q}$ pairs and more of the momentum of the proton is carried by the gluons (Fig. 2.4). This was observed in DIS experiments [12], were approximately half of the proton momentum was not carried by the charged particles that intervened in the electroweak process.

Additionally, experimental direct evidence of gluons were found TASSO [13]. Three-jet events in e^+e^- collisions were identified with $q\bar{q}$ production with a gluon emission (Fig. 2.5). All these evidences lead to the replacement of the Parton Model by the QCD theory.

2.4 QCD Factorization

A key ingredient of any cross section calculation in hadron interactions is the concept of QCD factorization. The QCD factorization theorem states that a cross section can be factorized

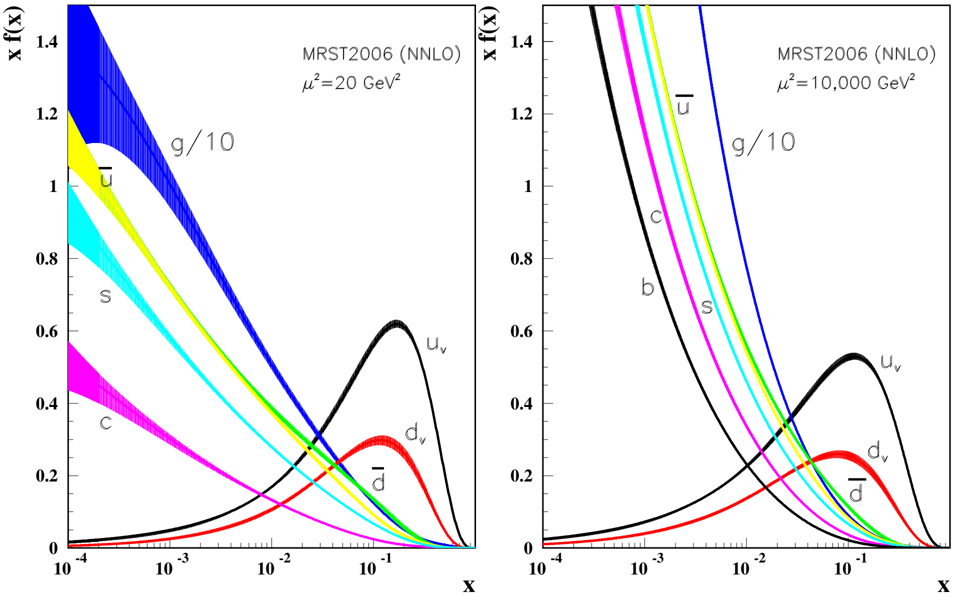


Figure 2.4: Example of proton PDFs measured at $Q^2 = 20\,\mathrm{GeV}^2$ and $Q^2 = 10,000\,\mathrm{GeV}^2$ in a DIS experiment for gluons and quarks. The contributions from the gluons and the quarks from the sea increase with Q^2. PDFs measured in DIS experiments are used in $p\overline{p}$ colliders.

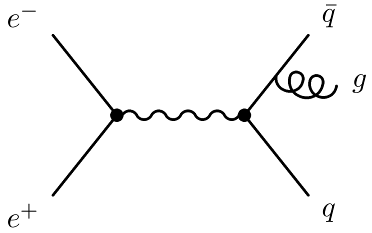


Figure 2.5: Three-jet event with a gluon emission.

into short- and long-distance effects delimited by a factorization scale μ_{F} (Fig. 2.6). According to this theorem the cross section of a two-hadron interaction with momenta P_1 and P_2, respectively, can be written as:

$$\sigma(P_1, P_2) = \sum_{i,j} \int dx_1 dx_2 \; f_i(x_1, \mu_{\mathrm{F}}^2) f_j(x_2, \mu_{\mathrm{F}}^2) \;\times\; \hat{\sigma}_{ij}\left(x_1, x_2, \alpha_s(\mu_{\mathrm{F}}^2, \mu_{\mathrm{R}}^2), Q^2/\mu_{\mathrm{F}}^2\right) \qquad (2.9)$$

where f_i are the PDFs, the sum runs over all parton types, and $\hat{\sigma}_{ij}$ is the parton cross section for incoming partons with momenta $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$. $\hat{\sigma}_{ij}$ is calculated at a given order on pQCD which introduces a dependence on a renormalization scale μ_{R}, that is usually chosen to be equal to μ_{F}.

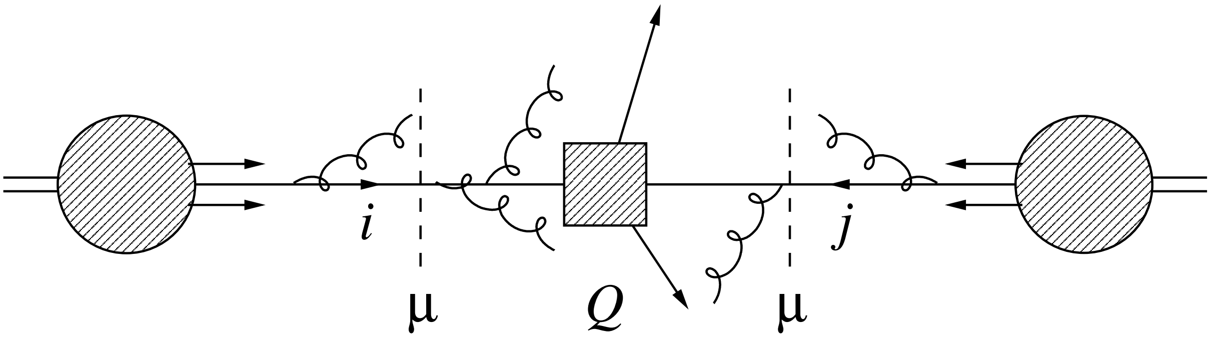
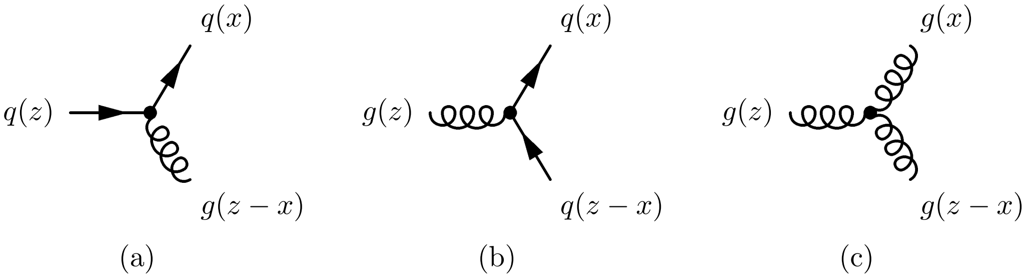


Figure 2.6: Schematic hadron-hadron interaction.

2.5 Parton Distribution Functions



(a) (b) (c)

Figure 2.7: Feynman diagrams of the lowest order processes of the splitting functions (Eq. 2.10): (a) gluon radiation, (b) quark pair production and (c) gluon splitting.

Perturbative QCD does not predict the form of the PDFs but can describe their evolution with the variation of the scale Q^2. The parton interactions at the lowest order in α_s are the three parton branching processes (Fig. 2.7): gluon radiation ($q \to qg$), gluon splitting

($g \rightarrow gg$) and quark pair production ($g \rightarrow q\bar{q}$). In the small-angle approximation, where the opening angle between the outgoing partons is small, and averaging over the polarizations and spins, these processes are described by the splitting functions [9]:

$$\begin{aligned} P_{gg}(z) &= 6\left[\frac{1-z}{z} = \frac{z}{1-z} + z(1-z)\right] && (2.10)\\ P_{qg}(z) &= \frac{1}{2}\left[z^2 + (1-z)^2\right] && (2.11)\\ P_{qq}(z) &= \frac{4}{3}\frac{1+z^2}{1-z} && (2.12)\end{aligned}$$

where z and x are defined in Fig. 2.7, and P_{ab} represents the probability that a parton of type a radiates a quark or a gluon and becomes a parton of type b carrying a fraction x/z of the momentum of parton a. The variation of the PDFs for quarks (q_i) and gluons (g) as a function of the scale Q^2 can be written in terms of these splitting functions:

$$\begin{aligned} \frac{dq_i(x,Q^2)}{d\log Q^2} &= \frac{\alpha_s}{2\pi}\int_x^1 \frac{dz}{z}\left[q_i(z,Q^2)P_{qq}\left(\frac{x}{z}\right) + g(z,Q^2)P_{qg}\left(\frac{x}{z}\right)\right] && (2.13)\\ \frac{dg(x,Q^2)}{d\log Q^2} &= \frac{\alpha_s}{2\pi}\int_x^1 \frac{dz}{z}\left[\sum_i q_i(z,Q^2)P_{qq}\left(\frac{z-x}{z}\right) + g(z,Q^2)P_{gg}\left(\frac{x}{z}\right)\right] && (2.14)\end{aligned}$$

These equations are called the DGLAP equations after Dokshitzer, Gribov, Lipatov, Altarelli and Parisi [14]. The first expression describes the change of the quark densities with Q^2 due to gluon radiation and gluon splitting, and the second expression describes the change of the gluon densities with Q^2 due to gluon radiation from quarks and gluons. The equations assume massless partons, hence only valid for light quarks (u, d and s), and work for scales down to $Q^2 \sim 1$ GeV.

2.5.1 PDF Parametrization

Experimental data is fitted to determine the parton densities at a given scale Q^2 and the DGLAP evolution equations predict the PDFs at higher scales. Figure 2.8 shows the structure function F_2 as a function of x and Q^2 as measured from DIS and fixed target experiments and the evolution predicted with the DGLAP equations.

The parameters from the fit are determined using a χ^2 minimization over data from different types of measurements: *ep* DIS experiments, fixed target neutrino experiments, as well as Drell-Yan production, W-asymmetry and jet inclusive cross sections in $p\overline{p}$ collisions. In this analysis, we used the parametrization provided by the Coordinated Theoretical-Experimental Project on QCD (CTEQ), but other fits, like the one performed by the Martin, Robers, Stirling and Thorne group (MRST), are also extensively used. The form of a typical parametrization of the quarks and gluon parton densities is:

$$xf_i(x,Q_0^2) = A_0 \cdot x^{A_1}(1-x)^{A_2}\left(1 + A_3 \cdot x^{A_4}\right) \qquad (2.15)$$

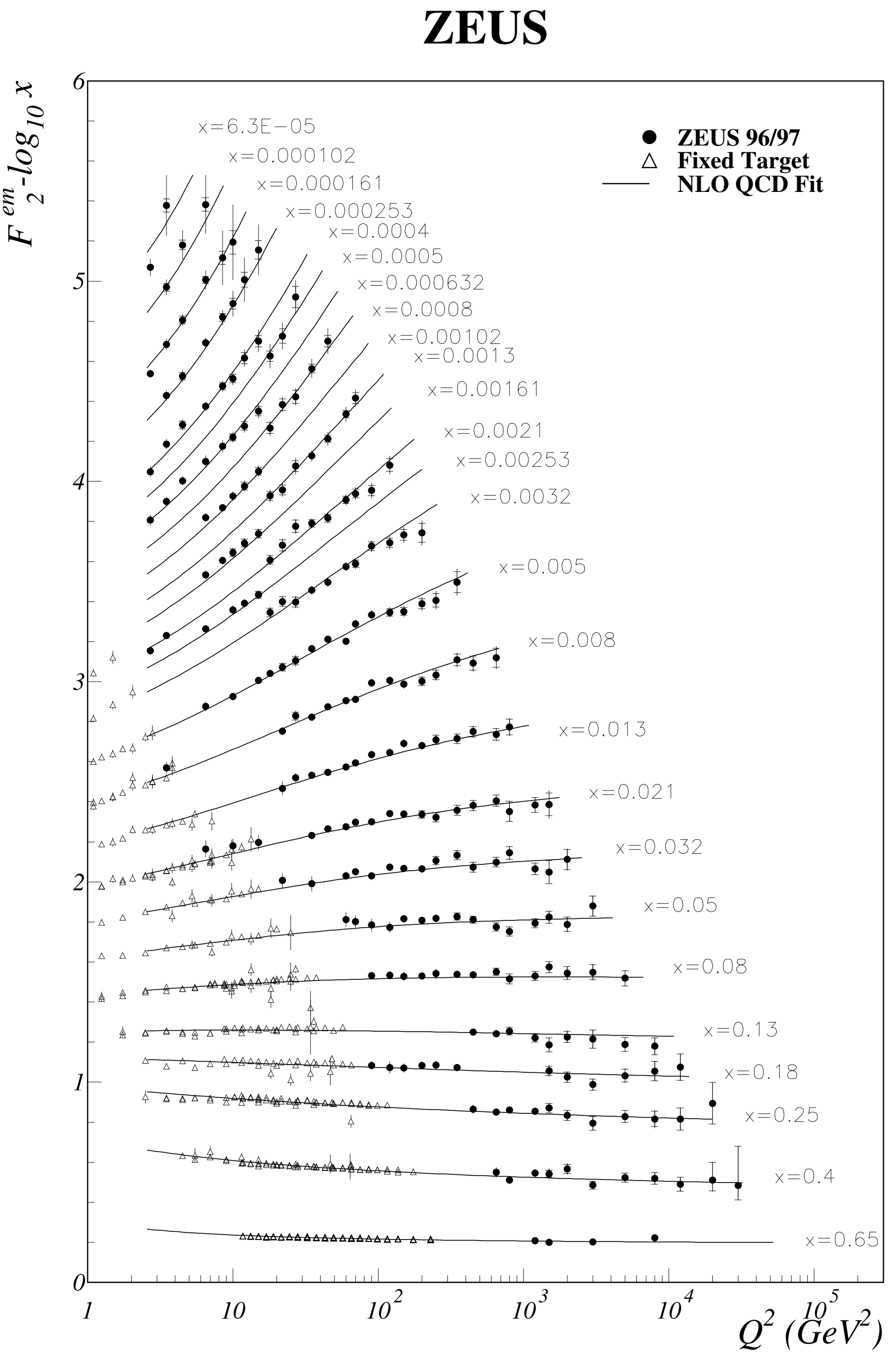


Figure 2.8: Structure function F_2 as a function of Q^2 and x measured in DIS (circles) and fixed target experiments (triangles). The curves are NLO pQCD fits.

where f_i is a particular parton density at Q_0^2 and A_i are the parameters to fit. Not all the parameters are free, as the functions must satisfy flavor and momentum sum rules.

2.6 Boson + jets pQCD Predictions

The production of W and Z bosons with hadronic jets in final state offers a good opportunity to test pQCD predictions. The presence of a boson with a large mass and high p_{T} jets provide a large enough hard scale Q to make perturbative calculations possible. The $Z/\gamma^* +$ jets production cross section can be decomposed into its multi-jet components:

$$\sigma_{Z/\gamma^*+\mathrm{jets}} = \sigma_{Z/\gamma^*+1j} + \sigma_{Z/\gamma^*+2j} + \sigma_{Z/\gamma^*+3j} + ... \tag{2.16}$$

where every jet multiplicity cross section can be perturbatively expanded as:

$$\begin{aligned} \sigma_{Z/\gamma^*+1j} &= a_1\alpha_s + a_2\alpha_s^2 + a_3\alpha_s^3 + a_4\alpha_s^4 + ... && (2.17) \\ \sigma_{Z/\gamma^*+2j} &= b_2\alpha_s^2 + b_3\alpha_s^3 + b_4\alpha_s^4 + ... && (2.18) \\ \sigma_{Z/\gamma^*+3j} &= c_3\alpha_s^3 + c_4\alpha_s^4 + ... && (2.19) \\ & ... && (2.20) \end{aligned}$$

The a_i, b_i and c_i coefficients in these expansions are in general functions of the jet definition, in particular the cone size used to cluster the partons into jets, and the transverse momentum, rapidity and separation cuts imposed on the jets. The contributions from the lowest order on α_s, i.e leading order (LO) contributions (a_1, b_2, c_3, ...), can be calculated from the matrix elements. Figure 2.9 shows the tree-level diagrams that contribute to the LO Z/γ^*+1 jet, and Fig. 2.10, some of the LO diagrams for Z/γ^*+2 jets. These LO contributions to the different jet multiplicities are used by some Monte Carlo generation programs to predict final states with high jet multiplicities. The techniques used to combine the cross sections of different jet multiplicities are described in Section 2.8.

Next-to-leading order (NLO) corrections (a_2, b_3, c_4...) are at present known only for $Z+1$ and $+2$ jets. The NLO contributions include virtual and real corrections to the tree-level processes:

$$\begin{aligned} \sigma_{Z+n}^{\mathrm{NLO}} &= \sigma_{Z+n}^{\mathrm{virtual}} + \sigma_{Z+n}^{\mathrm{real}} && (2.21) \\ &= \sigma_{Z+n}^{1-\mathrm{loop}} + \sigma_{Z+(n+1)}^{\mathrm{tree-level}} && (2.22) \end{aligned}$$

Virtual NLO contributions consist in one-loop corrections to the LO diagrams. Few examples of one-loop diagrams for $Z/\gamma^* + 1$ jet are shown in Fig. 2.11. Real corrections include the diagrams with the radiation of an extra parton in the final state. The soft, and collinear divergences that appear in the numerical calculation of the tree-level matrix elements of the real correction are evaluated using a subtraction method [15].

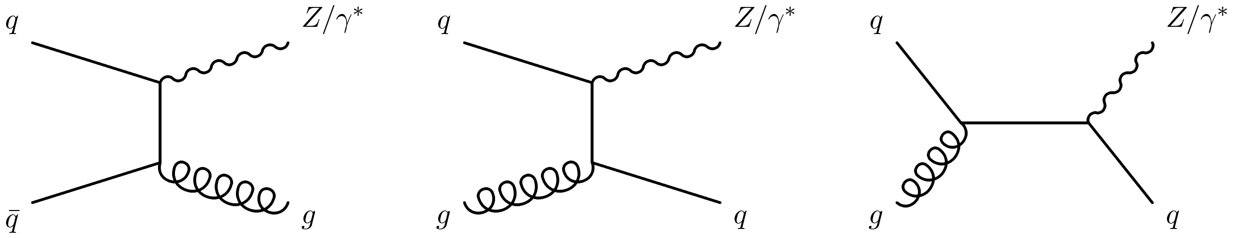

q
Z/γ*
q̄
g
q
Z/γ*
g
q
q
Z/γ*
g
q

Figure 2.9: Feynman diagrams for $Z/\gamma^* + 1$ jet production at LO in a hadron collider.

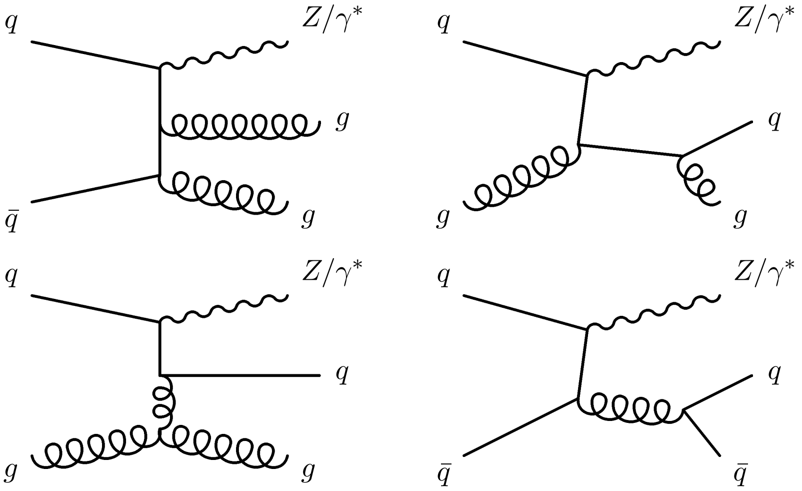

q
Z/γ*
g
q̄
g
q
Z/γ*
q
g
g
q
Z/γ*
q
g
g
q
Z/γ*
q
q̄
q̄

Figure 2.10: Some of the Feynman diagrams for $Z/\gamma^* + 2$ jets production at LO in a hadron collider.

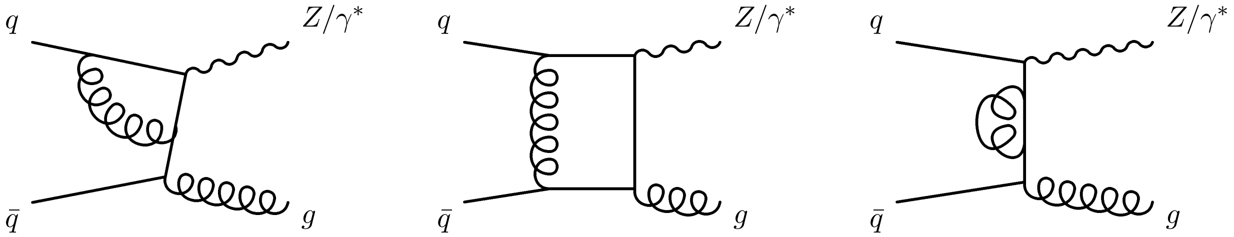

q
Z/γ*
q̄
g
q
Z/γ*
q̄
g
q
Z/γ*
q̄
g

Figure 2.11: Some of the Feynman diagrams contributing to virtual NLO corrections for $Z/\gamma^* + 1$ jet production in a hadron collider.

Next-to-next-to-leading order (NNLO) contributions include two-loop corrections to the LO diagrams, one-loop corrections to the NLO diagrams and the radiation of an extra parton. As the number of diagrams increases rapidly with the order of the calculation, predictions at the NNLO are not available yet for Z/γ^*+jets. In this analysis, we use the MCFM program [16] to obtain the NLO pQCD prediction for $Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet and ≥ 2 jets production cross sections, and the LO prediction for $Z/\gamma^*(\rightarrow e^+e^-)+\geq 3$ jets.

2.6.1 Dependence on the Scale

The independence of the Z/γ^* + jets cross section from the choice of scale μ is formally expressed as:

$$\mu^2 \frac{d\sigma}{d\mu^2} = 0 \tag{2.23}$$

Using the perturbative expansion of the cross section

$$\sigma = \sigma_0 + \alpha_s\sigma_1 + \alpha_s^2\sigma_2 + \alpha_s^3\sigma_3 + ... \tag{2.24}$$

the dependence on μ transforms into

$$\mu^2 \frac{d}{d\mu^2} \sum_{i=0}^{N} \alpha_s^i \sigma_i = \mathcal{O}(\alpha_s^{N+1}) \tag{2.25}$$

Then, as expected, the variation of the cross section with the scale gets smaller as more terms are included in the perturbative expansion. Figure 2.12 shows the scale dependence of the cross section prediction for $Z/\gamma^* \rightarrow e^+e^- + 1$ jet and 2 jets. The LO predictions rise sharply as the scale is decreased, while the predictions calculated at NLO produce a far flatter curve that exhibits a much less pronounced dependence on the scale choice. The dependence of the cross section on the scale at which the perturbation calculation is performed is usually estimated quoting the predicted cross section using half and twice the nominal scale value: $\mu/2$ and 2μ.

2.6.2 PDF Uncertainties

The Hessian formalism [17] is used to quantify the goodness of the fit of the PDFs to the experiment data. A matrix with the dimensions of the number of free parameters in the fit, 20 for the CTEQ parametrization, is diagonalized to obtain 20 orthogonal eigenvectors. This matrix determines the behavior of the minimization function χ^2 around $\chi^2(a_0)$ where a_0 is the vector of parameters that minimizes χ^2 and represents the best fit to the data. Points near a_0 are also considered acceptable fits if $\chi^2 - \chi^2(a_0) < T^2$, where T is a tolerance parameter and values $T^2 \sim 100$ are considered. For each eigenvector, displacements around a_0 in the

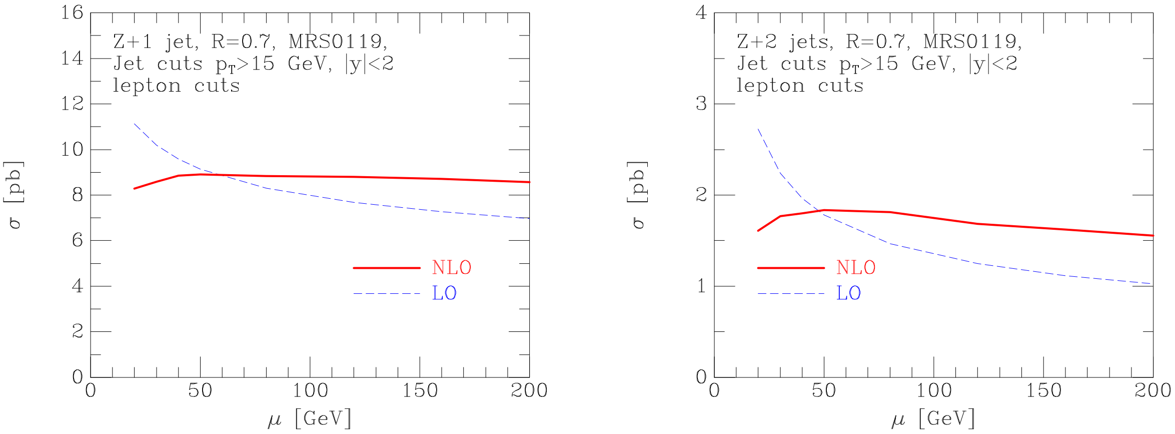


Figure 2.12: Scale dependence of the $Z/\gamma^*(\rightarrow e^+e^-) +$ jets cross section predictions with the factorization and renormalization scales set equal and given by μ.

directions along the vector are made, a_i^+ and a_i^- for the i^{th} eigenvector. This method results in a set of 40 PDFs: + and − for the 20 eigenvectors. Figure 2.13 shows the uncertainties on the up quark and gluon PDFs at $Q = 10$ GeV. While the up quark distribution is highly constrained for a wide range of x, the uncertainty on the gluon distribution increases rapidly for $x > 0.15$ up to values larger than 2 for the highest x values. The uncertainties also increase as Q^2 increases. For example, the uncertainty on the gluon PDF is the most important uncertainty affecting the theoretical prediction for the inclusive jet cross section at high values of the transverse momentum of the jets [18].

The uncertainties in the pQCD cross section prediction due to the knowledge of the PDF are determined in the following way:

$$\delta\sigma^+ = \sqrt{\sum_{i=1}^{20}\left(\max\{\sigma(a_i^+) - \sigma(a_0),\ \sigma(a_i^-) - \sigma(a_0),\ 0\}\right)^2} \tag{2.26}$$

$$\delta\sigma^- = \sqrt{\sum_{i=1}^{20}\left(\min\{\sigma(a_i^+) - \sigma(a_0),\ \sigma(a_i^-) - \sigma(a_0),\ 0\}\right)^2} \tag{2.27}$$

where $\sigma(a)$ is the prediction of the cross section determined using the PDFs with the parameters in vector a.

2.7 QCD Phenomenology and Monte Carlo Simulation

As was mentioned in the previous section, available perturbative calculations are truncated at a given order in α_s. However, the parton shower (PS) approximation sums to all orders

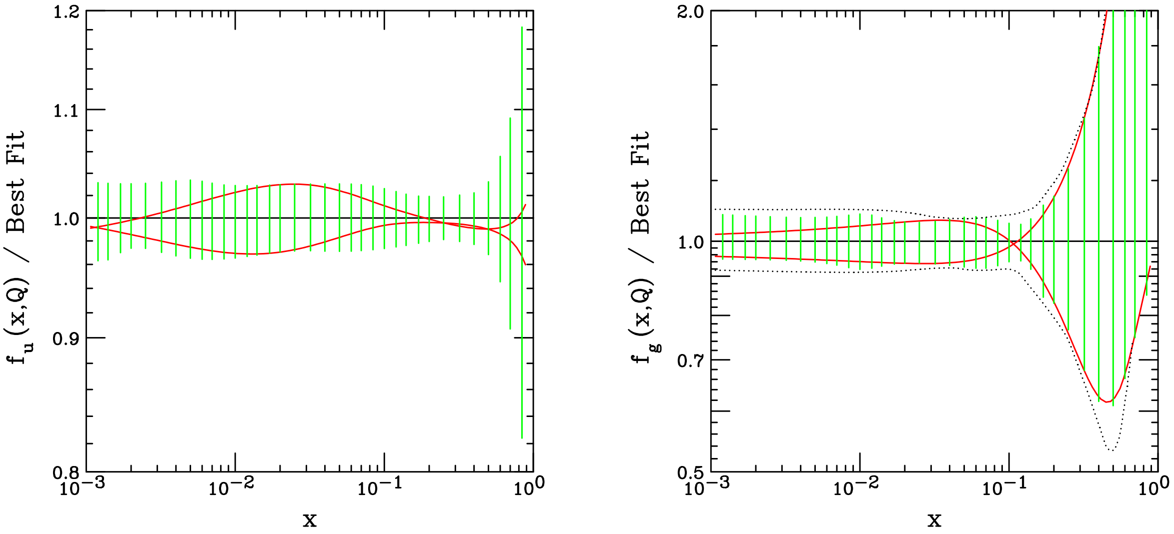


Figure 2.13: Uncertainties on up quark (left) and gluon (right) PDFs. The green bands represent the global uncertainty at $Q = 10$ GeV.

the leading soft and collinear enhanced contributions and their corresponding virtual corrections. This approximation is included in computed simulations called Monte Carlo generator programs (MC).

Monte Carlo programs also include models to reproduce non-perturbative effects like the hadronization process to produce colorless hadrons from the resulting partons, and the underlying event to generate soft interactions between the remnant partons from the colliding hadrons. Several Monte Carlo programs are described in detail in Section 2.8.

2.7.1 Parton Shower

Monte Carlo generators are usually limited to LO pQCD calculations but they include parton shower approximations to account for higher order contributions. By successive parton emission, the partons in the final state produce a cascade, where the splitting functions (Eq. 2.10) govern the radiation processes. For the computer implementation of the parton shower, the Monte Carlo programs use the Sudakov form factors $\Delta(t)$ [19]. They are derived from the splitting functions and represent the probability that a parton evolves from an initial scale t_o to a lower scale t without branching.

To simplify the implementation of the calculations in the Monte Carlo programs, the radiations are separated into initial-state and final-state showers, depending on whether they start off an incoming or outgoing parton of the hard scattering. In the final-state showers, the Monte Carlo branching algorithm operates in steps: when a branching $a \to b+c$ occurs at scale t_a, the fraction of momentum carried by the daughter partons x_b/x_a is determined using the appropriate splitting function P_{ab}, and the opening angle θ_a between b and c is given by $t_a = E_a^2(1-\cos\theta_a)$. In the next step, the scale at which partons b and c will branch, t_b and t_c, is determined using the Sudakov factors. The scale t is proportional to the virtual mass, thus t_b and t_c are kinematically constrained by $\sqrt{t_a} > \sqrt{t_b} + \sqrt{t_c}$, imposing that the subsequent branchings of the daughter partons will have smaller opening angles. The *angular ordering* is a coherence effect common to all gauge theories. Due to successive branching a parton cascade develops, and each outgoing line is the source of a new cascade until all outgoing lines have stopped branching. This stage depends on a cutoff scale Λ_{QCD} on the virtual mass of the partons.

In the initial-state showers, the algorithm applied is the same but operated backwards in time. Starting from an incoming parton at the hard interaction b, it finds the branching $a \to b + c$, where c can further branch in a final-state fashion. Therefore, in the construction of the initial-state shower, the fraction of momentum x is increased, and at the end it will match that described by the PDFs.

2.7.2 Hadronization

After the parton shower production, quarks and gluons recombine in color singlet states, the observed hadrons. The processes does not induce large transverse momentum transfers between the partons. This hypothesis, called *local parton-hadron duality*, is based on the observation that perturbation theory works well down to rather low scales $Q \sim 1$ GeV, and it states that hadrons are produced by partons which are close in phase space [20]. Therefore, the transition from a partonic to a hadronic jet does not wash out the original parton kinematics and flavor information, and the longitudinal and transverse momentum distributions are substantially unchanged. The final hadrons are collimated into a small angular region (jets) in the direction of the original parton.

The local parton-hadron duality does not describe the mechanism of hadron formation in any detail. For this, one must rely on models which have parameters that still need to be determined from experimental data. The hadron multiplicity, for instance, is tuned from the measurements of the spectra of particles in jets. The scale Λ_{QCD}, that decides when the parton shower stops and the hadronization starts, is also determined experimentally. The main hadronization models are the *string* and *cluster* models (Fig. 2.14) described below.

2.7.2.1 String Model

The string model [21] uses string dynamics to describe the color flux between a $q\bar{q}$ pair. The string produces a linear confinement potential. As q and $\bar{q}$ separate from each other, the

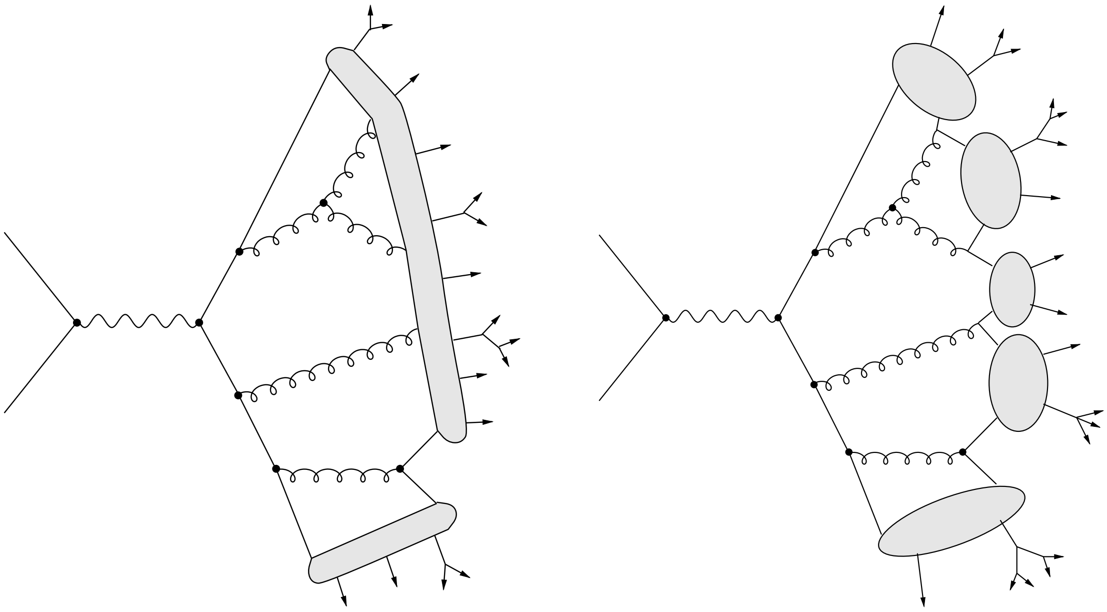


Figure 2.14: Parton shower with schematic string (left) and cluster (right) hadronization models.

potential energy in the string increases. When this energy exceeds the mass of a light $q\bar{q}$ pair, a new $q'\bar{q}'$ pair is created and the string breaks into two shorter strings creating two color-singlet states $q'\bar{q}$ and $q\bar{q}'$. If the relative momentum of the new $q\bar{q}$ pairs connected to the same string is large enough, the string may break again. Gluons produced in the shower act as 'kinks' in the string that add extra tension to it. The creation of heavy quarks is suppressed during the hadronization as the production of light quark pairs (u, d and s) is more energy favored. The heavy quarks present in the final state hadrons come from the perturbative parton shower through gluon splitting $g \to Q\bar{Q}$. The formation of baryons, 3-quark states, is achieved by considering them $quark - diquark$ states, where *diquarks* are simply treated as antiquarks. However, an accurate description of the baryon production is a problem in all hadronization models.

2.7.2.2 Cluster Model

The cluster model [22] is based on the color preconfinement property of the perturbative parton shower [23]. The process starts by splitting non-perturbatively all the gluons at the end of the parton shower into $q\bar{q}$ pairs $g \to q\bar{q}$, so that there are only quarks present. Quarks are then grouped in color singled clusters. The mass spectrum of these clusters peaks at low values of the order of the GeV/c^2 but has a broad tail at high masses. The clusters decay typically into two hadrons. Heavier hadrons are naturally suppressed by the mass spectra. The heaviest clusters can decay into smaller clusters that subsequently decay into hadrons.

The cluster model describes quite well the hadrons energy and transverse momentum in the data.

2.7.3 Underlying Event

In hadron-hadron colliders, the presence of hadron remnants that do not participate in the hard interaction leads to soft underlying event activity that contribute to the final state. The proper treatment of the underlying event involves taking into account their color and flavor connections with the hard interaction. These processes are non-perturbative and, as for the hadronization, Monte Carlo programs include models to describe them. As the dominant QCD cross sections decrease rapidly with $\hat{p}_{\mathrm{T}}$[3], the underlying event is modeled as 'minimum bias' processes with $\hat{p}_{\mathrm{T}} \gtrsim \hat{p}_{\mathrm{T,min}}$ and flat in rapidity. $\hat{p}_{\mathrm{T,min}}$ has a value around $1.5-2.5$ GeV/c and, as the others parameters in the models, it has been tuned to reproduced correctly the experimental results using observables specially sensitive to the underlying event modeling like, for example, the jet shapes and event profiles discussed in Chapter 5.

2.8 Monte Carlo Generator Programs

The following sections are devoted to describe briefly some of the most used Monte Carlo generator programs. Some of them are employed in different steps of the analysis.

2.8.1 PYTHIA Monte Carlo

PYTHIA is a Monte Carlo event program that uses LO matrix element to generate hard interactions between partons [24]. It is optimized for 2 → 1 and 2 → 2 calculations and the current version includes around 300 hard processes. It includes initial and final state parton shower evolution that add extra activity in the final state. The contributions from the soft underlying event are described by a combination of analytical results and various models. They allow several options for the probability of additional interactions, the level of detail in the generation of scatterings and how the correlations in flavor and color between the different scatterings are modeled. The hadronization in PYTHIA is performed using string fragmentation in the Lund model [21].

---

[3] $\hat{p}_{\mathrm{T}}$ is the transverse momentum of the hard scattering process in the center-of-mass frame, defined as $\hat{p}_{\mathrm{T}} = (\hat{t} \cdot \hat{u})/\hat{s}$, where $\hat{s}$, $\hat{t}$ and $\hat{u}$ are the Mandelstam variables.

2.8.1.1 Tune A Underlying Event Parametrization

There is a special parametrization of the underlying event in Pythia called Tune A. The Tune A settings mainly affect the initial-state showers from the incoming hadrons, where the scale Q^2 is increased and the lower cutoff is decreased allowing more radiation. There is also an increase of the probability that the multiple parton interaction produces two gluons with color connections to the nearest neighbors. These tunings are the result of specific measurements performed in CDF during the Run I of the Tevatron [25]. As a result of being tuned using data, the normalization of the hadron level distributions generated with Pythia-Tune A reproduce approximately the overall cross section measured in the data.

2.8.1.2 Tune DW Underlying Event Parametrization

Tune DW is very similar to Tune A describing the underlying event [26]. But, in addition, it also describes accurately the distribution of the transverse momentum of the Z/γ^* boson at very low values $p_{\mathrm{T}}^{Z} < 5$ GeV/c. This is achieved by increasing the width of the Gaussian distribution of the intrinsic $k_\perp$ of the partons inside the hadrons and its upper cutoff. These extra tunings have a significant effect on the shape and normalization of the p_{T} distribution of the jets.

2.8.2 Herwig Monte Carlo

Herwig is a general-purpose Monte Carlo event generator [27], which includes the simulation of hard hadron-hadron scattering and soft hadron-hadron collisions. It uses the parton-shower approach for initial- and final-state QCD radiation, including color coherence effects and azimuthal correlations both within and between jets. The model for the underlying event is based on a minimum bias $p\overline{p}$ event generator. Herwig also includes an interface to employ the multiple interaction model Jimmy to generate the underlying event activity (see next section). Finally, Herwig uses a cluster model to reproduce the hadronization of the partons.

2.8.2.1 Jimmy Underlying Event Parametrization

Jimmy provides secondary $parton - parton$ interactions to simulate the underlying event activity [28]. It is based on an eikonal model which assumes that individual interactions are uncorrelated. The multi-parton interaction rate is calculated using the LO matrix elements for 2 → 2 parton scattering, PDFs and the distribution of partons inside the hadrons as a function of the impact parameter $A(b)$. The impact parameter distribution is derived from the structure functions of the hadrons ignoring pQCD effects. The probability of the secondary interactions happening is determined by a p_T threshold $p_{\mathrm{T},2}$. Higher $p_{\mathrm{T},2}$ thresholds lead to lower amounts of underlying event (see Section 5.1.2).

2.8.3 ALPGEN Monte Carlo

ALPGEN is an event generator of multi-parton hard processes in hadronic collisions [29]. It uses the ALPHA algorithm [30] to compute tree level scattering amplitudes for large parton multiplicities in the final state. The advantage of ALPHA is that its complexity increases slower than the number of Feynman diagrams when increasing the particles in the final state. The algorithm determines the matrix elements from a Legendre transform of the effective action.

To combine different samples into an inclusive multi-parton sample, every event is weighted according to the matrix element calculation. The inclusion of parton showers to produce multi-jet samples leads to a double-counting problem: if the parton shower evolution produces an extra jet in a n-parton event, it will overlap with the contribution from $n+1$-parton configurations. The key is to clearly separate the phase space that is covered for the matrix elements calculation, and the space covered by the shower evolution. To properly match the matrix elements and the parton shower, ALPGEN uses the MLM matching algorithm [31]. The matching procedure starts by 'unweighting' the generated events in a exclusive sample of n partons in the final state and the color flow is randomly determined proportionally to the amplitudes calculated with ALPHA. Then, the parton shower is developed in that event using a PS Monte Carlo program like PYTHIA or HERWIG. The parton configuration after the showering is processed with a cone jet algorithm (see Section 2.9 for a description of jet algorithms) with a radius R_jet. Then, the original n partons are matched to the jets if $\Delta R(\mathrm{jet}, \mathrm{parton}) < R_\mathrm{jet}$. If all the partons are matched to a jet and there are no extra jets, i.e. $N_\mathrm{jets} = n$, the event is accepted. Otherwise it is rejected avoiding hard emissions in the parton shower that would lead to additional jets. Finally, the events with different jet multiplicities $n = 0, 1, 2...k$ are combined in an inclusive sample. Notice that with this recipe there would not be events with more than k jets. Therefore, the events in the sample with highest parton multiplicity are accepted if $N_\mathrm{jets} \geq k$.

2.8.4 Other Monte Carlo Generator Programs

Besides PYTHIA, HERWIG and ALPGEN that has been used in this analysis, there are other generators with interesting characteristics. SHERPA and MADGRAPH are multi-purpose parton level event generators, both aiming for a simple implementation of new models, interactions, particle properties, underlying event, etc.

Sherpa has been interfaced with Pythia for the parton shower evolution [32] using a different matching technique. The CKKW matching algorithm [33] operates in the following way: given a configuration with n partons in the final state, it generates a parton branching history using the $k_\perp$ algorithm [34] (see Fig. 2.15). The matrix elements are reweighted using the value of α_s in every vertex of the branching and the Sudakov factor from every line between the vertices. The initial conditions of the shower are set to have a smooth transition between the reweighted matrix elements and the parton shower, where the hard emissions in the shower evolution with enough transverse momentum to produce a separate jet (according to the $k_\perp$ algorithm) are vetoed.

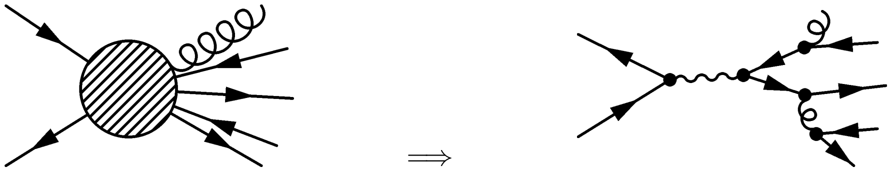


Figure 2.15: Branching history construction using the $k_\perp$ algorithm.

An interesting feature of Madgraph is that it implements both MLM and CKKW algorithms to add parton shower evolution to the hard scatterings. In the MLM implementation, a k_T clustering jet algorithm can be used instead of the original cone jet algorithm used in the original MLM implementation in Alpgen, making more straightforward direct comparisons with CKKW.

2.9 Jet Algorithms

The role of the jet algorithms is to associate clusters of particles into jets such that the kinematic properties of the jet can be related to the properties of the originating parton from the hard scattering. As will be seen, it is an important property of a good jet algorithm that it shows no infrared or collinear sensitivity. Jet algorithms start from a list of four-vectors that can be either calorimeter towers for detector measurements, stable particles in the final state from a Monte Carlo generated event, or partons in a fixed order pQCD prediction or a parton shower. The list of vectors fed to the jet algorithms are generally called “particles”. The method used to combine the momenta of the particles forming a jet to obtain the kinematical properties of the jet is called the recombination scheme. There exist two main classes of jet algorithms: cone-type algorithms and cluster-type algorithms, which are described in the next sections.

2.9.1 Cone Algorithms

The iterative cone-based algorithms are the most common in hadron colliders. They start with cones in the $\eta - \phi$ space$^{(4)}$ centered around seed particles of energy above a given threshold. Particles within a distance $R^{(5)}$ around the jet axis are clustered into a protojet. Then, the position of the E_{T}-weighted$^{(6)}$ *centroid* of the jet is determined using the Snowmass recombination scheme [35]:

$$\eta_{\text{centroid}} = \frac{\sum_i E_i \times \eta_i}{\sum_i E_i} \tag{2.28}$$

$$\phi_{\text{centroid}} = \frac{\sum_i E_i \times \phi_i}{\sum_i E_i} \tag{2.29}$$

With this new cone axis, the particles are clustered again. The iterative process is repeated until stable cones are found, i.e. the particles in the cone is unchanged from the previous iteration. The cone algorithms do not prevent particles from belonging to more than one jet, i.e. overlapping cones. As a result, a procedure must be specified to decide how to split or merge the jets that overlap. If the fraction of transverse energy shared by the cone with smaller transverse energy is larger than a given value $f_{\mathrm{m/s}}$ the two jets are merged into a single jet, otherwise, two jets are formed and the common towers are assigned to the closer jet. Notice that the merging and splitting of jets leads to larger cones with an effective radius different than R.

The jet cone algorithm described above corresponds to the JetClu CDF Run I algorithm. This algorithm turned out to be infrared unsafe when applied to parton level calculations. Figure 2.16 illustrates the concept of infrared sensitivity in cone algorithms: if two partons are within a distance $r \lesssim R$ they will be clustered in one or separate jets depending on the presence of soft radiation between them. Infrared sensitivity appears where there can be three nearby partons, as in NLO real corrections to the $Z/\gamma^* + 2$ jets cross section. The algorithm is also sensitive to collinear radiation. This behavior is illustrated in Fig. 2.17. If the energy of a calorimeter tower is split between two contiguous towers, a seed is no longer found and therefore no jet is formed. In order to address these theoretical difficulties, an additional step is included once all the stable cones are found and before the merging/splitting procedure. The midpoint between each pair of cones separated by less than $2R$ is added to the list of jets. The algorithm is rerun again until stability is achieved. The Midpoint algorithm [36] used in this analysis is named after this step. Additionally, the Midpoint algorithm uses the

---

[4] η is the pseudorapidity $\eta = -\ln \tan \frac{\theta}{2}$

[5] ΔR distance in the $\eta - \phi$ space.

[6] E_{T} is the transverse energy $E_{\mathrm{T}} = E \sin\theta$.

E-scheme for the recombination of the particles that consists in adding the four-momenta:

$$\mathbf{p}_{\text{jet}} = \sum_{\text{particles}} \mathbf{p}_{\text{particle}} = \left(\sum_{\text{particles}} \vec{p}_{\text{particle}}, \sum_{\text{particles}} E_{\text{particle}} \right) \tag{2.30}$$

which makes the jets massive, contrary to the snowmass scheme where jets are treated as massless. Cone jet algorithms used in NLO pQCD predictions require an extra parameter R_{sep} [37]. When used in theoretical calculations, the cone algorithms use an enlarged cone size $R^{'} = R \times R_{\text{sep}}$ to emulate the effects of merging and splitting cones applied to experimental data. R_{sep} is typically set to 1.3. The dependence on the choice of the value of R_{sep} introduces an uncertainty in the jet cross section around 2%.

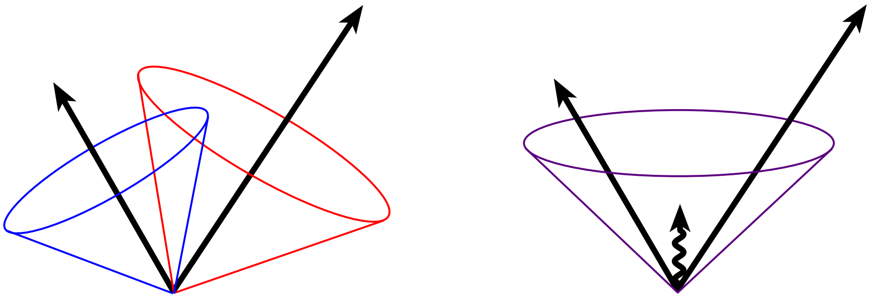


Figure 2.16: Illustration of infrared sensitivity in jet clustering. The presence of soft radiation between the two jets may cause the merging of the jets that would not occur in the absence of the soft radiation.

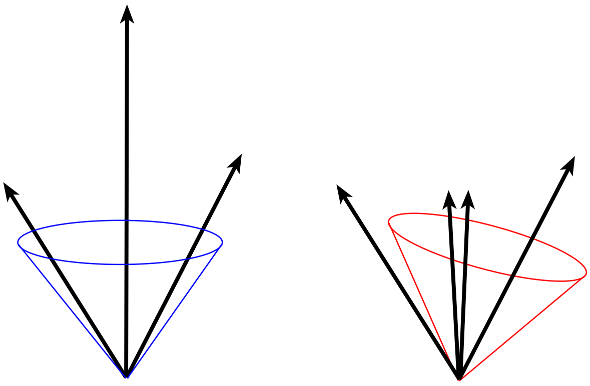


Figure 2.17: Illustration of collinear sensitivity in jet reconstruction. The configuration in the left fails to produce a seed because its energy is split among several detector towers.

2.9.2 Cluster Algorithms

A second class of algorithms, inspired by the QCD evolution of the parton showers described in Section 2.7.1, clusters particles according to their relative momentum rather than their spacial separation. Cluster-type algorithms are characterized by a scale y_{cut} that stops the iterative procedure in the case of exclusive k_{T} mode as it is run in e^+e^- collisions, or a $p_{\mathrm{T,\,min}}^{\mathrm{jet}}$ parameter in the case of inclusive k_{T} mode, as it is used in $p\overline{p}$ collisions. The latter is described below. The quantities

$$k_{\mathrm{T},i} = p_{\mathrm{T},i}^2 \quad ; \quad k_{\mathrm{T},(i,j)} = min(p_{\mathrm{T},i}^2, p_{\mathrm{T},j}^2) \cdot \Delta R_{i,j}^2 / D^2, \tag{2.31}$$

are computed for each protojet and pair of protojets, respectively, where $p_{\mathrm{T},i}$ denotes the transverse momentum of the i^{th} protojet, $\Delta R_{i,j}$ is the distance in $y - \phi$ space$^{(7)}$ between each pair of protojets, and D is a parameter that approximately controls the size of the jet by limiting, in each iteration, the clustering of protojets according to their spacial separation. All $k_{\mathrm{T},i}$ and $k_{\mathrm{T},(i,j)}$ values are then collected into a single sorted list. In this list, if the smallest quantity is of the type $k_{\mathrm{T},i}$, the corresponding protojet is promoted to be a jet and removed from the list. Otherwise, if the smallest quantity is of the type $k_{\mathrm{T},(i,j)}$, the protojets are combined into a single protojet by summing up their four-vector components. The procedure is iterated over protojets until the list is empty. In this way, every particle in the event is assigned to an unique jet avoiding overlap situations and the need for a merging/splitting criteria. Furthermore, the k_{T} algorithm is, by definition, infrared and collinear safe to all orders in pQCD and does not require the use of an extra parameter when it is applied to theoretical predictions.

---

[7] y is the rapidity $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$.

Chapter 3

CDF at Fermilab

The data used in this analysis was collected with the Collider Detector at Fermilab (CDF) during Tevatron Run II. In this chapter, the accelerator chain and the main CDF detector subsystems are explained.

3.1 The Tevatron

The Tevatron is currently the most powerful hadron collider operational in the world. It is located at Fermilab (Batavia, IL, USA). There, protons and antiprotons are accelerated to interact with a center-of-mass energy of 1.96 TeV. The CDF and DØ experiments are located at the two interaction points of the Tevatron.

The acceleration chain starts with hydrogen negative ions that are accelerated to 750 keV by a Cockcroft-Walton electrostatic accelerator. The hydrogen ions are then inserted in a linear accelerator (Linac), a 150 meter long chain of radio-frequency accelerator cavities where they reach an energy of 400 keV. Electrons are stripped from the ions leaving solely the protons. The bare protons are then injected into the Booster, a 75-meter-radius synchrotron, where they are split in several bunches and accelerated to 8 GeV. Before entering the Tevatron ring, the proton bunches are transfered to the Main Injector, a 3-km synchrotron that merges proton bunches into high-density bunches and accelerates them to 150 GeV. Once in the Tevatron, the 36 proton bunches, separated 396 ns, arrive to the final energy of 980 GeV.

The production of antiprotons is significantly more complicated and its production rate is the major limiting factors in the available instantaneous luminosity in the Tevatron. The antiprotons are produced when 120 GeV protons hit a nickel target. The production efficiency is $2{\cdot}10^{-5}$ antiprotons per incoming proton. Pulsed magnets and lithium lens separate the antiprotons from the other particles and focus them into a beam that is stored in the

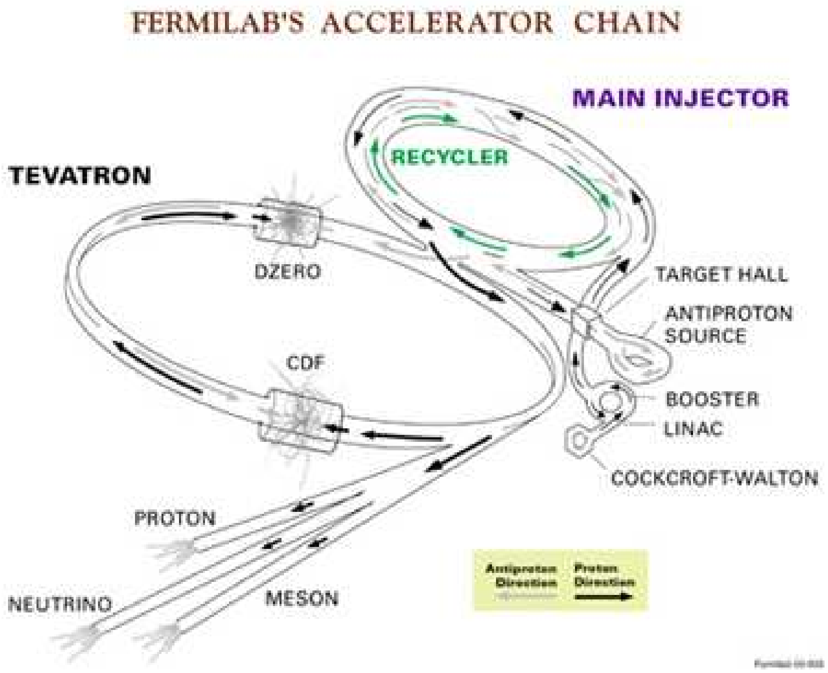

FERMILAB'S ACCELERATOR CHAIN
MAIN INJECTOR
RECYCLER
TEVATRON
DZERO
CDF
TARGET HALL
ANTIPROTON SOURCE
BOOSTER
LINAC
COCKCROFT-WALTON
PROTON
NEUTRINO
MESON
Antiproton Direction
Proton Direction

Figure 3.1: Chain of accelerators at Fermilab (not to scale) with the Tevatron and the CDF II detector.

Accumulator ring. When enough antiprotons are accumulated ($\sim 10^{12}$), they are transfered to the Main Injector where their energy is increased up to 150 GeV and then transported to the Tevatron, where, with the protons, they are accelerated to 980 GeV.

As winter of 2008, Tevatron instantaneous luminosity have reached record values of more than $3{\cdot}10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (Fig. 3.2). Since 2001, the Tevatron has delivered 3.8 fb^{-1} of data, and CDF has recorded 3.1 fb^{-1} of data to tape (Fig. 3.3).

3.2 The CDF Run II Detector

The CDF Run II detector [38], shown schematically in Fig. 3.4, is a general purpose experiment in operation since 2001. It is a cylindric-shape detector with azimuthal and forward-backward symmetry(8).

The next sections are devoted to the description of the subsystems more relevant to the analysis, and following the path of a particle coming from the interaction point.

---

[8] CDF uses a cylindrical coordinate system. The origin is set at the geometrical center of the detector. The z axis is taken along the beam direction and the x axis ($\phi = 0$) in the accelerator plane pointing away from the center of the ring.

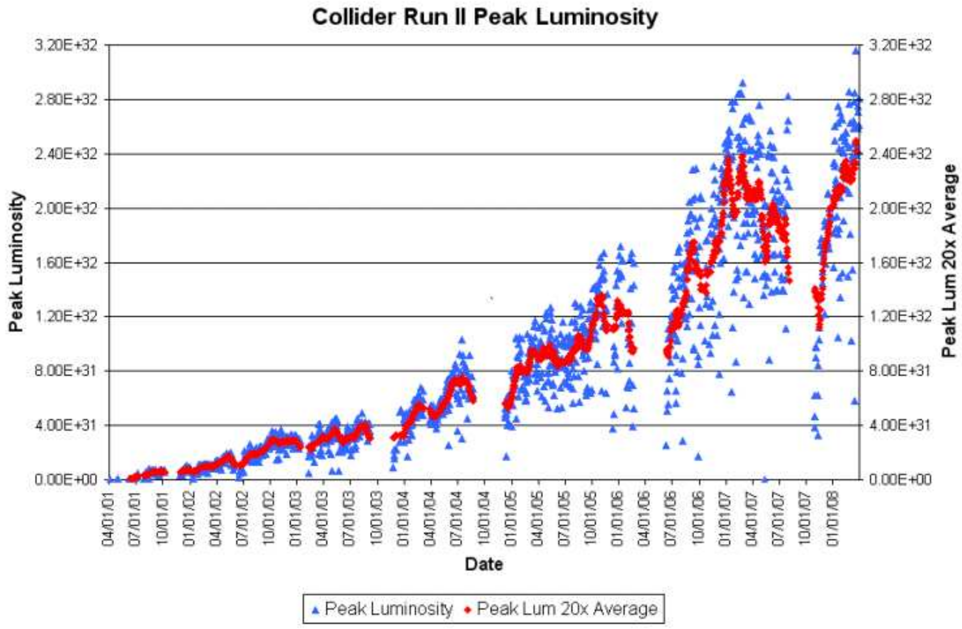


Figure 3.2: Tevatron Run II peak instantaneous luminosities since 2001 for every store (blue triangles) and averaged every 20 stores (red diamonds).

3.2.1 Tracking Systems

The tracking systems of the CDF detector (shown in Fig. 3.5) are used to measure the trajectory of the charged particles produced in the collisions. They are immersed in a 1.4 Tesla magnetic field generated by a superconducting solenoid of 1.5 m of radius. The tracking is performed by the silicon detectors and the Central Outer Tracker.

3.2.1.1 Silicon Detectors

The silicon detectors [39] are the subdetectors closest to the beam. They provide 3-dimensional $(r-\phi-z)$ measurements of the trajectory of charged particles. They consist of several layers of silicon microstrip detectors that cover a pseudorapidity region up to $|\eta| < 2.8$. The closest layer to the beam ($r = 1.35$ cm) is Layer 00 (L00), a silicon strip that carries out $r-\phi$ position measurements. It is designed to improve the resolution in the determination of the impact parameter$^{(9)}$ of the track. The next five layers, that expand from $r = 2.4$ cm to 10.6 cm, constitute the Silicon Vertex detector (SVXII). They consist of silicon sensors

---

[9] The impact parameter d_0 is defined in the transverse plane as the shortest distance of a particle trajectory from the primary vertex to the point where the particle decays.

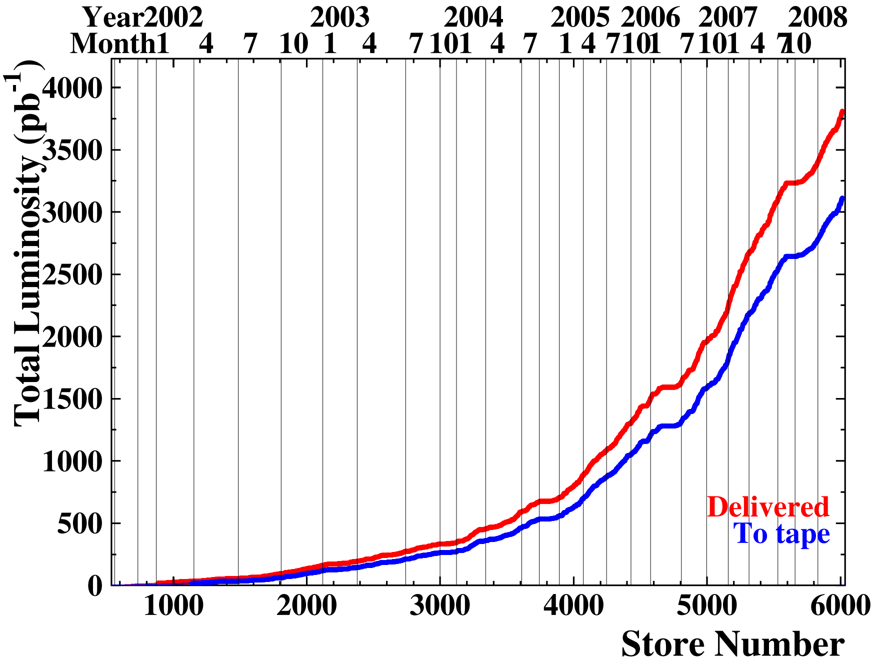


Figure 3.3: Tevatron Collider Run II delivered integrated luminosity (red) and integrated luminosity recorded by CDF (blue).

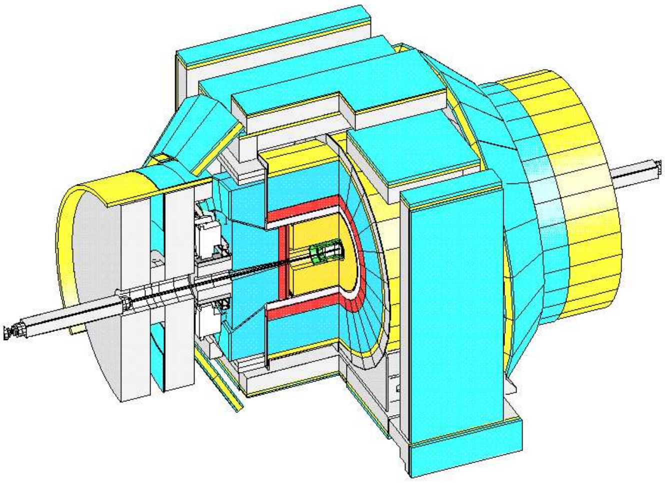


Figure 3.4: Isometric view of the CDF Run II detector with its main subsystems (innermost to outermost): the Silicon Vertex Detector (green), the Central Outer Tracker (orange), the superconducting solenoid (white), the electromagnetic calorimeter (red), the hadronic calorimeter (blue) and the muon chambers (yellow and light blue).

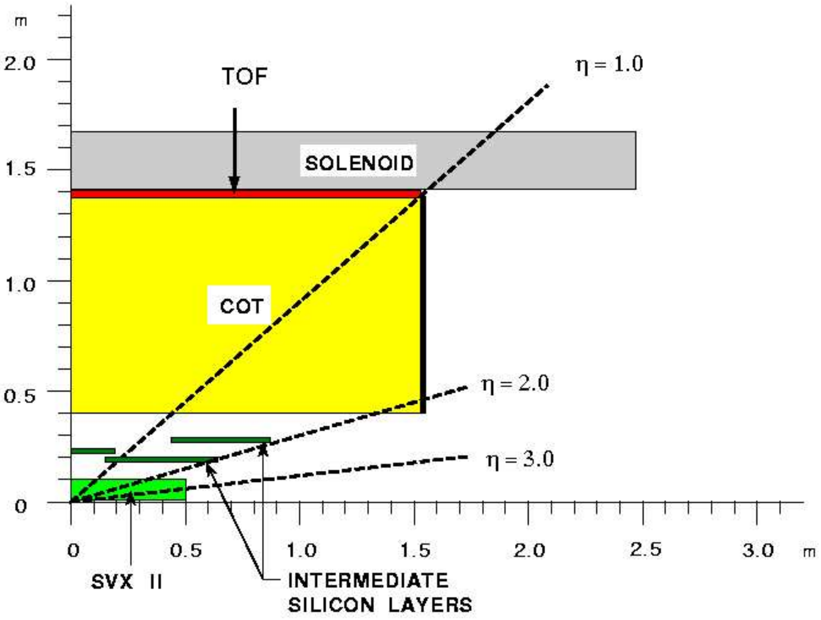


Figure 3.5: Longitudinal view of the CDF Run II detector, showing the coverage in pseudorapidity of the tracking systems.

with a combination of both 90° and small-angle stereo layers, providing up to 5 $r - \phi - z$ track position measurements. The last two layers, with an outer radius of $r = 29$ cm, are the Intermediate Silicon Layers (ISL). The ISL increases the efficiency of identifying long-lived particles. The detector achieves resolutions of about 9 μm and 40 μm on position and impact parameter measurements, respectively, and 70 μm in the determination of the z_0(10) position of a track.

3.2.1.2 Central Outer Tracker

The silicon detectors are surrounded by the Central Outer Tracker (COT). The COT [40] is a cylindrical, open cell, multiwire drift chamber. The cylinder is 310 cm long, and radially expands from $r = 40$ cm to 137 cm. It provides a coverage of $|\eta| < 1$ (Fig. 3.5).

The sense wire of the drift chamber are radially grouped in 8 "superlayers" (SL). The SL are divided in ϕ into "supercells", each supercell containing 12 sense wires and 13 potential wires (40 μm diameter gold coated tungsten wires), and delimited by 6.3 μm thick gold coated mylar field sheets. Figure 3.6 shows the layout of a supercell. The size of the supercells

---

[10] z_0 is the intersection of the track with the beamline.

is similar in all the superlayers, therefore the number of supercells in a given SL scales approximately with the radius. “Axial” and “stereo” SL alternate. In axial SL, the wires run parallel to the z axis. In stereo SL, the wires strung at a small angle (2°) with respect to the z axis. The combination of the axial and stereo measurements provide z and $r-\phi$ position information. Single hit resolutions in the COT of 140 μm translate into transverse momentum resolutions of $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}} = (0.15\%) \times p_{\mathrm{T}}[(GeV/c)^{-1}]$. If silicon tracking information is added, the resolution improves to $(0.07\%) \times p_{\mathrm{T}}[(GeV/c)^{-1}]$.

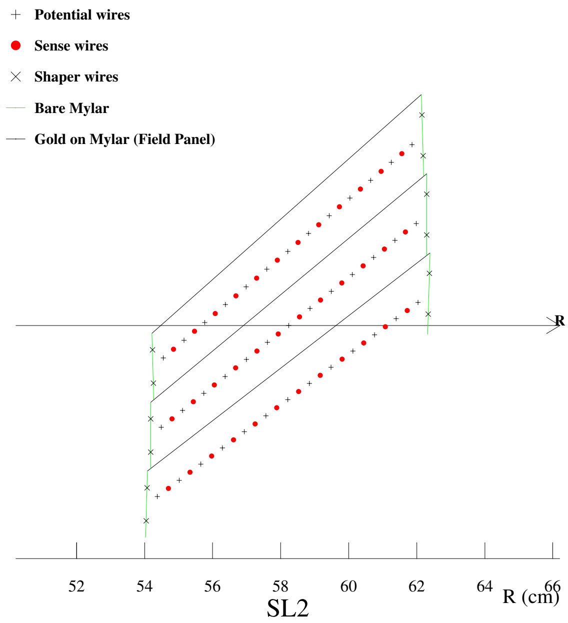


Figure 3.6: Layout of three COT supercells.

The COT chamber is filled with a gas mixture of Argon-Ethane (50:50) that provides a constant electron drift velocity across the cells. As the COT is immersed in a magnetic field, the electrons drift at a Lorentz angle(11) of 35°. Supercells are tilted by 35° with respect to the radial direction to compensate this effect. Small quantities of Oxygen and Isopropyle are added to the gas mixture to reduce aging effects on the wires.

---

[11] Charged particles in an electric field $\vec{E}$ are accelerated along the field direction. If a magnetic field $\vec{B}$ is also present, the moving particles are also affected by the Lorentz force perpendicular to $\vec{B}$ and the direction of the particle. This effect make the particle to drift at a given angle (Lorentz angle) with respect to $\vec{E}$.

3.2.2 Time-of-Flight System

Between the COT and the superconducting solenoid there is a time-of-flight detector (TOF) [41]. It is formed by 216 scintillating bars of almost 3 m of longitude located at $r \sim 140\,cm$ with one photo-multiplier tube attached to each end. Each bar covers 1.7° in ϕ and $-1 < \eta < 1$ in pseudorapidity (as shown in Fig. 3.5). The TOF allows the measurement of the time-of-arrival of the particle with respect to the collision time with a resolution of $\sim$ 100 ps. This information, combined with the momentum measurement in the tracking systems makes possible to distinguish between kaons, protons and pions with a good separation power for $p_{\rm T} < 1.5$ GeV/c.

3.2.3 Calorimeters

The calorimetry system of CDF is located after the solenoid coil. The calorimeters are non-compensated, sampling calorimeters that use scintillating plastics as active material. They are segmented in projective towers pointing to the nominal interaction point at the center of the detector. In total, CDF consists in 1536 calorimeter towers. The light produced by the particles of a shower that cross a scintillating plate is collected by wavelength shifting (WLS) fibers that transport it to photomultiplier tubes (PMT) located in the outermost part of the calorimeters. Every projective tower is read by one or two PMTs.

The subdetectors that constitute the calorimeter of CDF are classified in two main groups: the central calorimeters, that approximately cover the region $|\eta| < 1.1$, and the plug calorimeters, that cover $1.1 < |\eta| < 3.6$. The central calorimeters consist of two separate halves that meet at $\eta = 0$. Due to this configuration, two "gap" regions exist around $\eta = 0$ and $|\eta| = 1.1$. Figure 3.7 shows the spacial disposition of the calorimeters. Table III lists the main characteristics of each calorimeter.

3.2.3.1 Central Calorimeters

The innermost calorimeter is the central electromagnetic calorimeter (CEM). The CEM [42] is designed to contain the electromagnetic showers created by particles coming from the interaction point. It is segmented azimuthally in 15° (0.26 radians) wedges and 0.1 in pseudorapidity. Every tower in the central calorimeters is read by two WLS fibers and two PMTs. It uses lead as absorber and has a depth of 18 radiation lengths (X_0)(12).

---

(12) The radiation length X_0 describes the characteristic amount of matter transversed. It is defined as the average path length required for a relativistic charged particle to lose all but $\frac{1}{e}$ of its energy by bremsstrahlung as it passes through matter, $\frac{1}{X_0} = -\frac{(dE/dx)_{bremss}}{E}$. It is also equivalent to $\frac{7}{9}$ of the mean free path for e^+e^- production of high-energy photons.

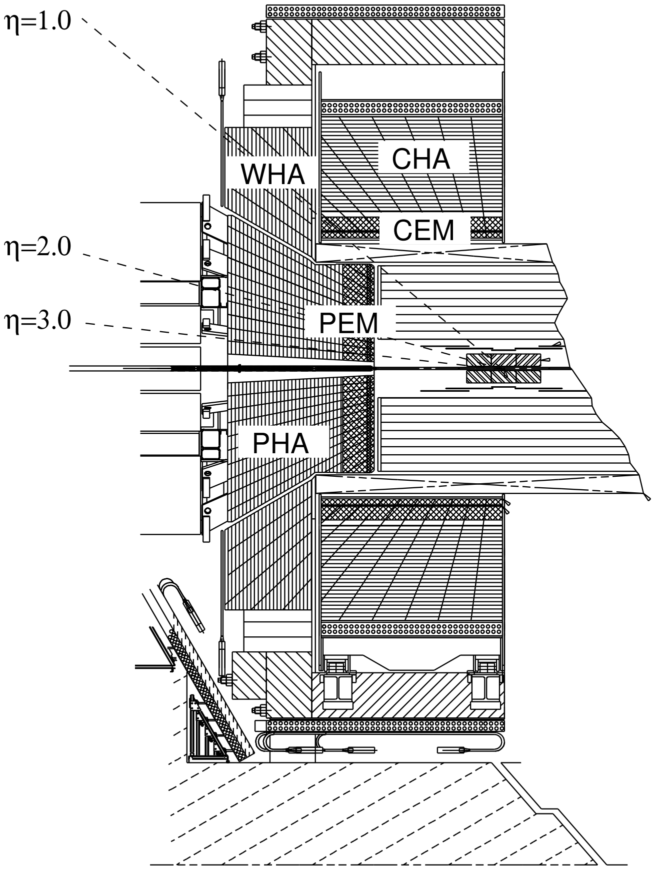

η=1.0
η=2.0
η=3.0
WHA
CHA
CEM
PEM
PHA

Figure 3.7: Elevation view of the CDF Run II detector calorimeters.

The CEM is surrounded by the central hadronic calorimeter (CHA). The CHA [43] has the same segmentation of the CEM and every CHA tower matches a CEM tower. It uses steel as absorber material and has a depth of 4.7 interaction lengths $(\lambda_I)^{(13)}$. To extend the coverage of the CHA, the Endwall Hadronic calorimeter (WHA) is located behind the CEM in the region $0.9 < |\eta| < 1.3$. The WHA [43] has a similar construction of the CHA, with the same absorber, segmentation and depth.

The energy resolution for a single incident particle of each segment of the calorimeter was measured using a testbeam and can be parametrized as

$$\frac{\sigma}{E_{\mathrm{T}}} = \frac{a}{\sqrt{E_{\mathrm{T}}}} \oplus b \tag{3.1}$$

where $\oplus$ represents the square root of the quadratic sum, and $E_{\mathrm{T}} = E \sin\theta$ where θ is the particle incident angle. The first term a comes from sampling fluctuations and photostatistics of PMTs. The constant term b comes from the intercalibration between the different towers and the non-uniform response of the calorimeter. For the CEM, the energy resolution of high-energy electrons and photons has $a = 14\%(GeV/c)^{1/2}$ and $b = 2\%$. Charged pions were used to obtain the energy resolution in the hadronic calorimeters. For the CHA, that resulted in $a = 50\%(GeV/c)^{1/2}$ and $b = 3\%$, and for the WHA $a = 75\%(GeV/c)^{1/2}$ and $b = 4\%$.

3.2.3.2 Plug Calorimeters

The plug calorimeters cover the region $1.1 < |\eta| < 3.6$. They are built with the same technology of the central calorimeters and replaced the gas detectors used during Run I. The detectors are segmented as projective towers, but the segmentation varies in different regions: ϕ segmentation varies between 7° and 15°, while the segmentation in pseudorapidity changes from 0.1 to 0.6 as η increases. The light produced in the scintillator plates is read by one PMT for every tower. The plug calorimeters are divided in two main parts: the electromagnetic calorimeter (PEM) and the hadronic calorimeter (PHA).

The PEM [44] is built with lead as passive material and has a depth of $23X_0$. The PHA [45], with a depth of $6.8\lambda_I$, uses iron as absorber. The segmentation in PEM and PHA are the same, and the towers in both calorimeters match, except the lowest η towers in the PEM which do not have a corresponding PHA tower.

The energy resolution of the plug calorimeters is parametrized as a function of the absolute energy of the incident particle as

$$\frac{\sigma}{E} = \frac{\alpha}{\sqrt{E}} \oplus \beta \tag{3.2}$$

---

[13] The interaction length λ_I is defined as the average path length of a particle before undergoing an inelastic interaction with a nucleus of the medium it is transversing.

The results in the testbeam were $\alpha = 16\%(GeV/c)^{1/2}$, $\beta = 1\%$ for the PEM, and $\alpha = 80\%(GeV/c)^{1/2}$, $\beta = 5\%$ for the PHA.

	Coverage	**Segmentation** ($\eta \times \phi$)	**Thickness**	**Resolution** (E in GeV)
CEM	$\|\eta\| < 1.1$	0.1×0.26	$18X_0, 1\lambda_I$	$14\%/\sqrt{E_{\mathrm{T}}} \oplus 2\%$
CHA	$\|\eta\| < 0.9$	0.1×0.26	$4.7\lambda_I$	$50\%/\sqrt{E_{\mathrm{T}}} \oplus 3\%$
WHA	$0.9 < \|\eta\| < 1.3$	0.1×0.26	$4.7\lambda_I$	$75\%/\sqrt{E_{\mathrm{T}}} \oplus 4\%$
PEM	$1.1 < \|\eta\| < 3.6$	$(0.1 - 0.6) \times (0.13 - 0.26)$	$23X_0, 1\lambda_I$	$16\%/\sqrt{E} \oplus 1\%$
PHA	$1.2 < \|\eta\| < 3.6$	$(0.1 - 0.6) \times (0.13 - 0.26)$	$6.8\lambda_I$	$80\%/\sqrt{E} \oplus 5\%$

Table III: Summary of the characteristics of the CDF Run II calorimeters.

3.2.3.3 Shower Profile Detectors

Two types of shower profile detectors help to distinguish between electrons, photons and $\pi^0 \rightarrow \gamma\gamma$: the shower maximum detectors and the pre-radiator detectors. The Central Shower Maximum (CES) [42] and the Plug Shower Maximum (PES) [46] are located at $6X_0$ in the CEM and PEM, respectively. They sit at the depth at the expected maximum of the lateral shower profile. The Central Pre-Radiator (CPR) [47] and the Plug Pre-Radiator (PPR) [48] are located in the inner face of the central and plug calorimeters, respectively. They consist of several multiwire proportional chambers which sample the electromagnetic showers that start in the solenoid material in front of them.

3.2.4 Muon Detectors

The outermost subdetectors are a set of drift chambers and scintillators devoted to the detection of muons. The different muons subsystems [49] are: the Central Muon Detector (CMU), the Central Muon Upgrade Detector (CMP, CSP), the Central Muon Extension Detector (CMX, CSX) and the Intermediate Muon Detector (IMU). Altogether they cover $|\eta| < 2.0$. The signals in the muon detectors are matched with hits in the COT to reconstruct the full trajectory of the muons.

3.2.5 DAQ and Triggers

With bunches separated by 396 ns, the crossing rate of physics bunches at the Tevatron reach 2.5 MHz. Data can only be written to tape at a rate of $\sim 75\ Hz$ and only a fraction of the collision generate interesting events. This enormous reduction factor is carried out by the trigger systems that select interesting physics events to be recorded and copied to tape. The CDF trigger system has a three-level architecture with each level providing a rate reduction sufficient to allow processing in the next level with minimal deadtime. Figure 3.8 shows a general sketch of the data flow across the different trigger levels. The first and the second level are hardware based and the third one is a processor farm. The trigger chain is synchronized by the Trigger Supervisor Interface (TSI) that provides clock signals and distributes trigger decisions between the different levels.

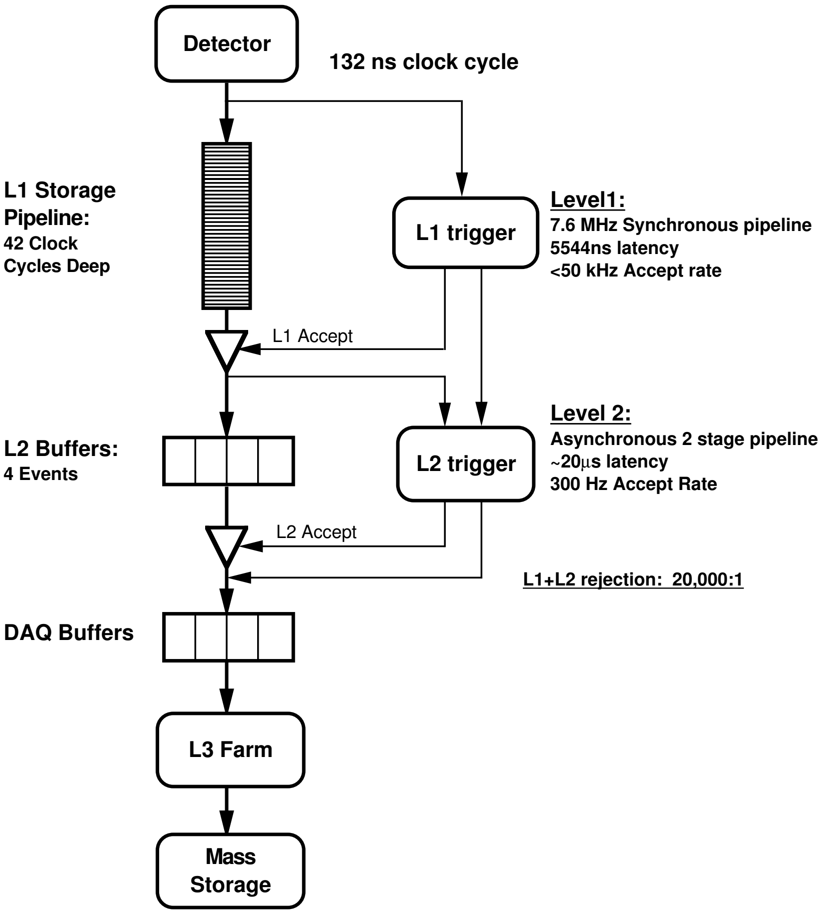

Detector
132 ns clock cycle
L1 Storage Pipeline:
42 Clock Cycles Deep
L1 trigger
Level1:
7.6 MHz Synchronous pipeline
5544ns latency
<50 kHz Accept rate
L1 Accept
L2 Buffers:
4 Events
L2 trigger
Level 2:
Asynchronous 2 stage pipeline
~20µs latency
300 Hz Accept Rate
L2 Accept
L1+L2 rejection: 20,000:1
DAQ Buffers
L3 Farm
Mass Storage

Figure 3.8: General diagram of CDF Run II trigger and data acquisition systems. The figure corresponds to the original design of the system conceived to run at 132 ns clock cycles that was the expected bunch spacing at the Tevatron.

Level 1 uses hardware to find physics objects based on partial detector information. It is a synchronous system with an event read and a decision made every 396 ns, with a buffer of 42 clock cycles. After Level 1 requirements, the rate of events is reduced to less than 50 kHz. The system consists in three parallel processing streams. One stream finds calorimeter based objects, a second stream finds muons, and the third one finds tracks in the COT. As some

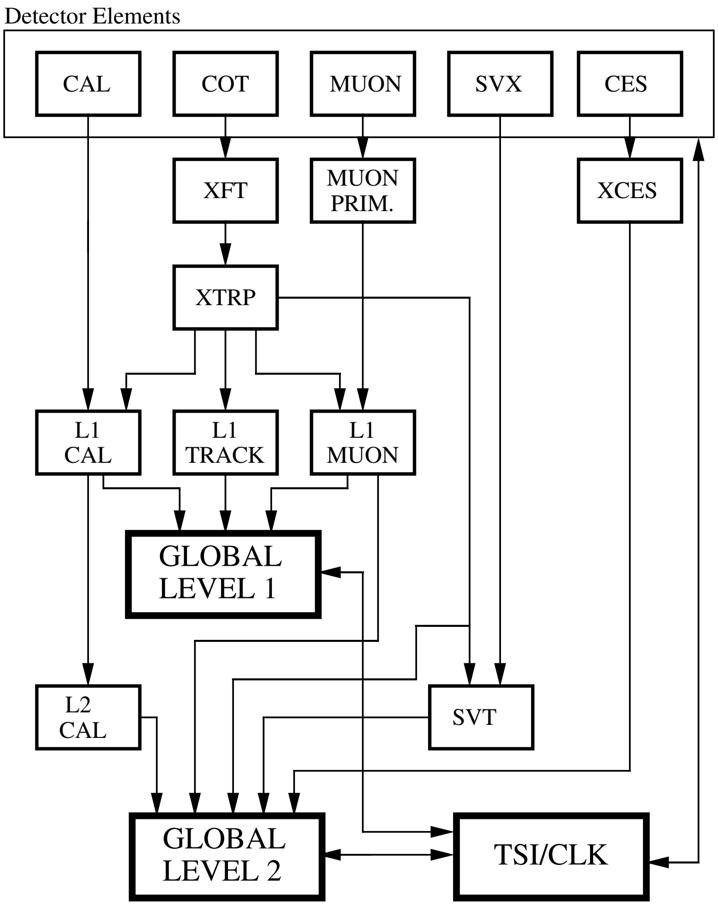

Detector Elements
CAL
COT
MUON
SVX
CES
XFT
MUON PRIM.
XCES
XTRP
L1 CAL
L1 TRACK
L1 MUON
GLOBAL LEVEL 1
L2 CAL
SVT
GLOBAL LEVEL 2
TSI/CLK

Figure 3.9: Diagram of Level 1 and Level 2 trigger systems and data flow between them.

calorimeter and muon triggers require the presence of a track, i.e. to find an electron or a muon, track information is also sent to the calorimeter and muon streams (see Fig. 3.9). Level 1 trigger decisions are formed using simple logic ANDs and ORs of objects from these streams.

At Level 2, the data is transferred to programmable processors. The input data consist on Level 1 tracks, Level 1 muons, shower maximum calorimeters and data from the silicon detectors. Level 2 trigger systems consists on several asynchronous subsystems: Cluster Finder (L2CAL) runs over calorimeter data to find clusters of energy; the Silicon Vertex Tracker (SVT) [50] uses silicon systems data to detect secondary vertices from B hadron decays; and algorithms that run on shower maximum detectors data to find isolated energy clusters, and use track information to better distinguish between electrons and photons.

The last level of the trigger is software based. After the Level 2 decision, the data blocks from different detector components are read by the Event Builder and the Reformatter builds complete events into the appropriate data structure for analysis. During this process, the event integrity is checked and warnings are issued if the error rate is too high. Then Level 3 trigger algorithms process the event. Level 3 algorithms take advantage of the full detector information and improved resolution not available to the lower trigger levels, i.e. 3-dimensional track reconstruction, tighter matching of tracks to calorimeter, muon systems information, and jets reconstructed with the same jet algorithms used in the offline analysis with vertex position at $z = 0$.

A trigger path is a set of trigger requirements that an event must satisfy at Level 1, Level 2 and Level 3 to be recorded. There are more than 200 trigger paths implemented in CDF. Events that pass any of these trigger paths are sent to the Consumer Server Logger (CSL) and from there to mass storage. Datasets are groups of trigger paths that identify the physics objects required by the trigger: B hadron, high-p_{T} electron, high-p_{T} muon, high-p_{T} photon, taus, jets, missing transverse energy, J/Ψ, etc. and combinations of them. Every event can be copied to one or more datasets depending on the trigger paths it passed. The data used in this analysis was recorded using the ELECTRON_CENTRAL_18 trigger path. This trigger path selects events with a high-p_{T} energy cluster in the central calorimeter and a COT track pointing to it.

3.2.6 Luminosity Measurement

At CDF the beam luminosity is determined from the rate of inelastic $p\overline{p}$ scattering, which has a relatively large cross section $\sigma_{in} = 61$ mb and provides enough interactions for a precise measurement. The $p\overline{p}$ scattering rate is measured with the Cherenkov Luminosity Counters (CLC) [51]. The CLC consists of two modules installed around the beampipe at each end of the detector (East and West) which provide coverage in the region $3.75 < |\eta| < 4.75$. Each module contains 48 long conical gas-filled Cherenkov counters pointing to the center of the detector. The counters are arranged in three concentric layers of 16 counters each, around the beampipe. Each counter is made of highly reflective aluminized Mylar with a light collector that gathers the Cherenkov light into photomultiplier tubes. The gas used in the counters is Isobutane, a good radiator and highly transparent. Coincidence between East and West modules is required when counting interactions. Since the number of $p\overline{p}$ interactions per bunch crossing μ follows Poisson statistics, it is determined by measuring the fraction of empty bunch crossings $N_{n=0}$ over the total number of crossings N_{total}:

$$\mathcal{P}(n=0) = \mathrm{e}^{-\mu} = \frac{N_{n=0}}{N_{\mathrm{total}}} \tag{3.3}$$

The instantaneous luminosity L is derived from the number of inelastic $p\overline{p}$ events per bunch crossing measured with the CLC (μ), the CLC acceptance (ϵ_{CLC}), the rate of bunch crossings (f_{BC}) and the inelastic $p\overline{p}$ cross section(σ_{in}), according to the expression:

$$L = \frac{\mu \cdot f_{\mathrm{BC}}}{\epsilon_{\mathrm{CLC}} \cdot \sigma_{in}} \tag{3.4}$$

The CLC acceptance is determined using both data and simulation with the formula:

$$\epsilon_{\mathrm{CLC}} = \left(\frac{N_{\mathrm{CLC}}}{N_{\mathrm{CLC+Plug}}}\right)_{\mathrm{data}} \times \left(\frac{N_{\mathrm{CLC+Plug}}}{N_{\mathrm{inelastic}}}\right)_{\mathrm{simulation}} \tag{3.5}$$

where $N_{\mathrm{CLC+Plug}}$ is the number of inelastic interactions tagged by the CLC and the plug calorimeters, N_{CLC} is the subset of events with East-West coincidence that pass the selection criteria and $N_{\mathrm{inelastic}}$ is the number of total inelastic collisions. The obtained acceptance value is $\epsilon_{\mathrm{CLC}} = 0.60 \pm 0.03$.

σ_{in} is obtained from the combined measurements of CDF and E811 at $\sqrt{s} = 1.8$ TeV and extrapolated to 1.96 TeV [52]. The total uncertainty on the luminosity measurement is 5.8%. It is dominated by the uncertainty on the CLC acceptance (4.4%) and in σ_{in} (4.0%). Other sources of uncertainty with smaller contributions ($< 2\%$) come from the detector calibration and stability versus time.

3.3 Reconstruction of Physics Objects

In this section we describe how the physics objects relevant for the analysis are reconstructed from the signals collected by the different subdetectors in CDF.

3.3.1 Electrons

Electron candidates are reconstructed starting from “seed” calorimeter towers with transverse electromagnetic energy $E_{\mathrm{T,EM}} > 10$ GeV. In the central calorimeters, electromagnetic clusters must be contained in the same wedge in ϕ and can consist at most of the seed and its nearest neighbors in η on either side. Neighbor towers must satisfy $E_{\mathrm{T}} > 0.1$ GeV and ${E_{\mathrm{T}}}^{neighbor}/{E_{\mathrm{T}}}^{seed}$ to be less than a given value. Every tower can only be in one cluster. In the plug calorimeters, electromagnetic clusters have a square 2×2 tower configuration. The energy clusters in the calorimeters are then matched to clusters in the shower maximum detectors and to the track with maximum p_{T} from all the tracks that point to the shower maximum cluster, if any. Variables on the profile of the shower and the matching between the energy cluster and the shower maximum cluster will be used to select the electrons in the analysis (see the event selection in Section 4.3.3).

Electromagnetic clusters are treated as massless objects. The total energy of the cluster is $E = E_{\mathrm{EM}} + E_{\mathrm{HAD}}$, where E_{EM} and E_{HAD} are the sum of the electromagnetic and hadronic energy of the towers in the clusters, respectively. The transverse energy of the electron is defined: $E_{\mathrm{T}} = E_{\mathrm{T,EM}} + E_{\mathrm{T,HAD}}$. ϕ and η coordinates of the cluster are determined following the *snowmass* scheme:

$$\eta = \frac{E_{\mathrm{EM}} \times \eta_{\mathrm{EM}} + E_{\mathrm{HAD}} \times \eta_{\mathrm{HAD}}}{E} \tag{3.6}$$

$$\phi = \frac{E_{\mathrm{EM}} \times \phi_{\mathrm{EM}} + E_{\mathrm{HAD}} \times \phi_{\mathrm{HAD}}}{E} \tag{3.7}$$

where η_{EM} and ϕ_{EM} are pondered over the towers in the cluster:

$$\eta_{\mathrm{EM}} = \frac{\sum_i E^i_{\mathrm{EM}} \times \eta^i}{\sum_i E^i_{\mathrm{EM}}} \tag{3.8}$$

$$\phi_{\mathrm{EM}} = \frac{\sum_i E^i_{\mathrm{EM}} \times \phi^i}{\sum_i E^i_{\mathrm{EM}}} \tag{3.9}$$

and, similarly, η_{HAD} and ϕ_{HAD} are obtained according to:

$$\eta_{\mathrm{HAD}} = \frac{\sum_i E^i_{\mathrm{HAD}} \times \eta^i}{\sum_i E^i_{\mathrm{HAD}}} \tag{3.10}$$

$$\phi_{\mathrm{HAD}} = \frac{\sum_i E^i_{\mathrm{HAD}} \times \phi^i}{\sum_i E^i_{\mathrm{HAD}}} \tag{3.11}$$

3.3.2 Jets

Jets are defined as collimated flows of particles from the fragmentation of partons into hadrons. Jets appear as clusters of energy depositions in the calorimeters. The jet reconstruction algorithms use the so-called physics towers. A physics tower is created from every detector tower in the following way:

$$\begin{aligned} p_x &= E_{\rm EM}\sin\theta_{\rm EM}\cos\phi_{\rm EM} + E_{\rm HAD}\sin\theta_{\rm HAD}\cos\phi_{\rm HAD} && (3.12)\\ p_y &= E_{\rm EM}\sin\theta_{\rm EM}\sin\phi_{\rm EM} + E_{\rm HAD}\sin\theta_{\rm HAD}\sin\phi_{\rm HAD} && (3.13)\\ p_z &= E_{\rm EM}\cos\theta_{\rm EM} + E_{\rm HAD}\cos\theta_{\rm HAD} && (3.14)\\ E &= E_{\rm EM} + E_{\rm HAD} && (3.15)\end{aligned}$$

where $E_{\rm EM}$ ($E_{\rm HAD}$) is the energy deposited in the electromagnetic (hadronic) part of the tower, and $\theta_{\rm EM}, \phi_{\rm EM}$ ($\theta_{\rm HAD}, \phi_{\rm HAD}$) is the direction from the primary vertex of the interaction to the typical shower maximum position of an electromagnetic (hadronic) shower (Fig. 3.10). Note that, as shown in the figure, if the vertex of the interaction is not at the center of the detector $\theta_{\rm EM}$ and $\theta_{\rm HAD}$ can be different. The four-momentum of every physics tower is then:

$$\mathbf{p} = (p_x, p_y, p_z, E) \tag{3.16}$$

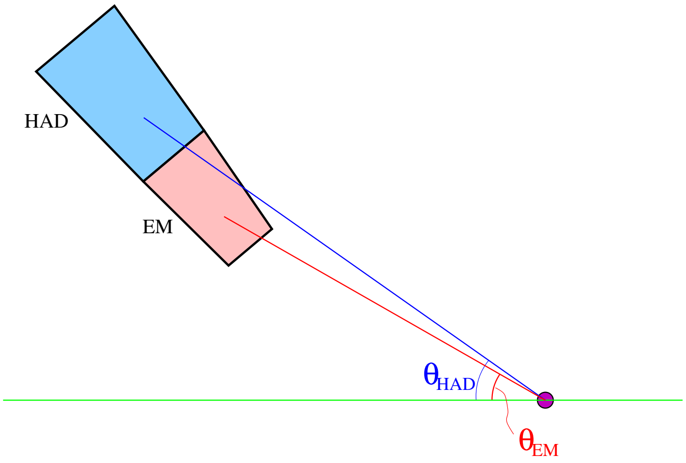


Figure 3.10: Schematic of a single CDF calorimeter tower. $\theta_{\rm EM}$ is the angle from the interaction point to the position of the maximum of a typical electromagnetic shower developed in the tower. $\theta_{\rm HAD}$ is the angle from the interaction point to the position of the maximum of a typical hadronic shower from a pion developed in the tower.

Towers belonging to the identified electrons are excluded from the list of available towers used to reconstruct jets. Jets are searched for using the Midpoint algorithm, as described in Section 2.9.1, with a cone radius of $R = 0.7$ in $y - \phi$ space and a merging/splitting fraction of 0.75. Towers with $p_{\rm T} > 0.1$ GeV are considered, and towers with $p_{\rm T} > 1$ GeV are used as initial seeds. The measured transverse momentum $p_{\rm T,cal}^{\rm jet}$ and the rapidity $y^{\rm jet}$ of the jets are determined with the E-scheme (Eq. 2.30):

$$p_{\rm T,cal}^{\rm jet} = |\vec{p}_{\rm cal}| \times \sin\theta \ , \quad y^{\rm jet} = \frac{1}{2}\ln\frac{E+p_z}{E-p_z} \tag{3.17}$$

where θ is the polar angle. Rapidity y is used instead of pseudorapidity η because in the E-scheme jets are massive and only Δy distances are invariant under Lorentz transformations.

The jets at particle level, in Monte Carlo generated events, are reconstructed from the stable particles before interacting with the calorimeter(14) At particle level, the Midpoint algorithm uses all the particles in the final state (besides the electrons from the Z/γ^* boson decay), and particles with $p_\mathrm{T} > 1$ GeV as seeds.

3.3.2.1 Corrections of the Jets Transverse Momentum

The transverse momentum of the jets measured in the calorimeter $p_\mathrm{T,cal}^\mathrm{jet}$ is systematically lower than the corresponding jet at the level of stable particles $p_\mathrm{T}^\mathrm{jet}$. The energy of the jet measured in the calorimeter is underestimated due to the presence of non-instrumented material and the non-compensated nature of the calorimeters. In hadron colliders, at high instantaneous luminosities, pile-up contributions to $p_\mathrm{T,cal}^\mathrm{jet}$ from multiple $p\overline{p}$ interactions in the same bunch crossing can also occur. An average correction in $p_\mathrm{T}^\mathrm{jet}$ is applied to correct for those effects as [53]:

$$p_\mathrm{T,corr}^\mathrm{jet} = (p_\mathrm{T,cal}^\mathrm{jet} \times C_y - C_\mathrm{MI}) \times C_\mathrm{Abs} \tag{3.18}$$

where the correction factors C correspond to:

- C_y is a correction depending on y^jet. The response of the calorimeter changes along y, due to the different response levels of the central and plug calorimeters and the cracks in the boundaries between the calorimeters. To obtain a homogeneous calorimeter response in y^jet, $p_\mathrm{T}^\mathrm{jet}$ is calibrated with respect to the central calorimeter, where we have also track information that can be used as cross-check. The C_y corrections are obtained imposing dijet balance in two-jet events with one jet in the central calorimeter (Fig. 3.11). The y-dependent corrections are determined separately for data and Monte Carlo simulated events, and for different $p_\mathrm{T}^\mathrm{jet}$ bins. After applying C_y, the dijet balance should become flat in y at 1.0. The remaining differences are due to the limitations of the parameterizations in y and p_T, and are taken as systematic uncertainties on the correction. The total uncertainty on C_y ranges between 0.5% and 3% depending on p_T and y.

- C_MI is the transverse momentum subtracted from $p_\mathrm{T,cal}^\mathrm{jet}$ to account for extra energy in the jet cone coming from multiple $p\overline{p}$ interactions occurred in the same bunch crossing. Extra $p\overline{p}$ interactions are identified via the presence of additional reconstructed vertices. The average transverse momentum ξ inside a random cone of radius $R = 0.7$ in $y-\phi$ as

---

(14) In the Monte Carlo samples, particles with lifetimes above 10^{-11} seconds in the final state are considered stable particles

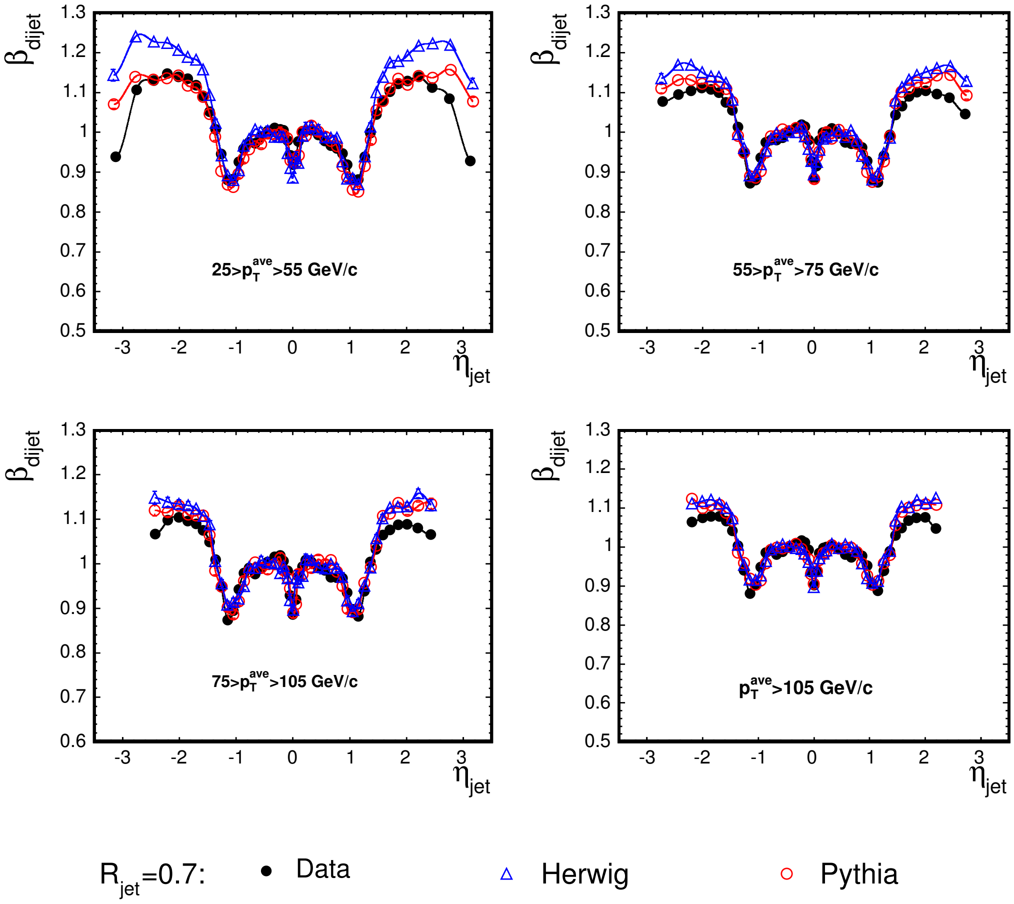


Figure 3.11: Dijet balance $\beta_{\text{dijet}} = p_{\mathrm{T}}{}^{\text{probe}}/p_{\mathrm{T}}{}^{\text{trigger}}$ as a function of η^{jet} for data, Pythia and Herwig MC samples in different $p_{\mathrm{T}}^{\text{jet}}$ bins. The dijet balance shows the non-homogeneous response of the calorimeter and the gaps at $\eta = 0$ and $|\eta| = 1.1$.

a function of the number of reconstructed vertices N_v is measured in a minimum bias data sample (Fig. 3.12). The results indicates that a value of $\xi = 1.06 \pm 0.32$ GeV/c must be subtracted from $p_{\mathrm{T,cal}}^{\mathrm{jet}}$ for each additional vertex in the event:

$$p_{\mathrm{T}}' = p_{\mathrm{T}} - \xi(N_v - 1) \tag{3.19}$$

This method relies on the accuracy of the vertex reconstruction algorithm. The efficiency of the vertex reconstruction is measured in several data samples with different topologies ($W \rightarrow e\nu$, inclusive jets and minimum bias samples) from where a 15% uncertainty on the subtracted p_{T} is determined.

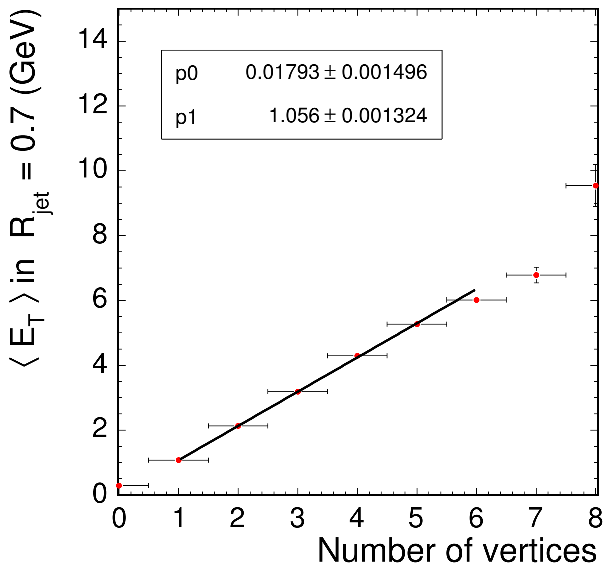


Figure 3.12: Mean E_{T} in a random cone of $R = 0.7$ as a function of the number of reconstructed vertices.

- C_{Abs} is the average absolute jet correction applied to compensate the underestimation of the measured transverse momentum of the jets with respect to the hadron level jets. The correction is obtained from the calorimeter simulation implemented in the Monte Carlo. For each particle jet, $p_{\mathrm{T}}^{\mathrm{jet}}$ is determined as a function of the $p_{\mathrm{T,cal}}^{\mathrm{jet}}$ measured in the calorimeter (Fig. 3.13). The uncertainty on the correction is computed from the difference between data and Monte Carlo in the calorimeter response to single particles, and varies between 2% at $p_{\mathrm{T}}^{\mathrm{jet}} \sim 30$ GeV/c, and 2.7% at $p_{\mathrm{T}}^{\mathrm{jet}} \sim 00$ GeV/c. The uncertainty on C_{Abs} constitutes the main source to the total jet energy uncertainty [53].

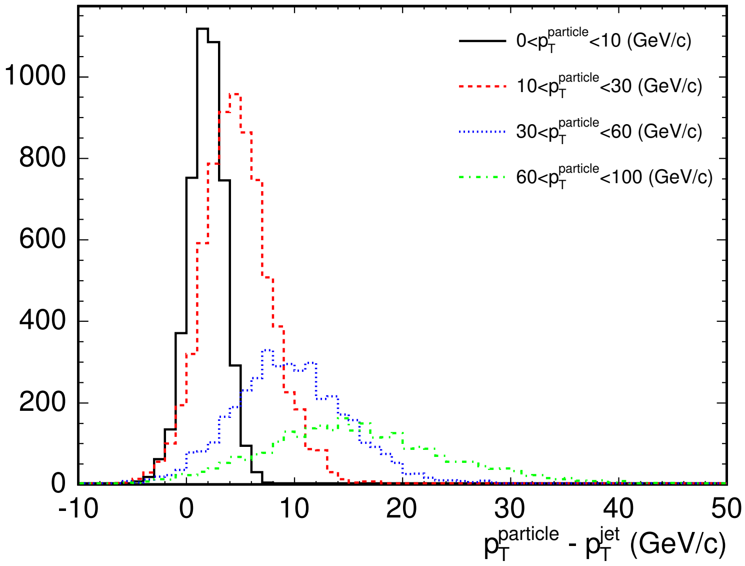


Figure 3.13: Difference between the $p_{\rm T}$ of the particle jet and the $p_{\rm T}$ of the jet reconstructed in the calorimeter for different bins in $p_{\rm T}^{\rm jet}$.

Chapter 4

Inclusive Jet Cross Sections in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ Production

This chapter is devoted to explain in detail all the steps that we followed to measure the jet cross section in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ events. The chapter describes the dataset used, the calculation of the efficiencies, the event selection, the estimation of the background contributions, the unfolding of the cross section, and finishes with a discussion on the systematic uncertainties.

4.1 Definition of the Dataset

We use data collected with the CDF Detector between February 2002 and January 2007 corresponding to a total integrated luminosity of 1.73 fb^{-1}. Data have been selected using the high-$p_{\rm T}$ electron trigger ELECTRON_CENTRAL_18. The requirements at the different levels of the trigger are listed in Table IV. The trigger selection requires a cluster of energy in the central electromagnetic calorimeter and a reconstructed track pointing to it. We only consider data for which all the subsystems used in the analysis were fully operational (see Appendix A for a detailed description of the Data Quality Monitoring systems).

4.1.1 Electron Trigger Efficiency

The efficiency of the ELECTRON_CENTRAL_18 trigger to select central electrons has been measured independently for every level of the trigger and separately for the calorimeter and the track requirements.

To measure the tracking efficiency $\epsilon^{\rm trk}$ we use a sample of $W \rightarrow e\nu$ candidates selected with

ELECTRON_CENTRAL_18 trigger path	
Level 1	Cluster in the central calorimeter with $E_{\mathrm{T}} > 8$ GeV
L1_CEM8_PT8	$E_{HAD}/E_{EM} \leq 0.125$
	A track pointing to the energy cluster with:
	• $p_{\mathrm{T}}^{\mathrm{trk}} > 8$ GeV/c
	• Hits in ≥ 3 superlayers in the COT
	• ≥ 11 hits per superlayer in the COT
Level 2	Cluster in the central calorimeter with $E_{\mathrm{T}} > 16$ GeV
L2_CEM16_PT8	$E_{HAD}/E_{EM} \leq 0.125$
	A track pointing to the energy cluster with:
	• $p_{\mathrm{T}}^{\mathrm{trk}} > 8$ GeV/c
	• $\lvert\eta^{track}\rvert < 1.3$
Level 3	Number of electromagnetic objects ≥ 1
L3_CEM18_PT9	Electromagnetic object in the central calorimeter with:
	• $E_{\mathrm{T}} > 18$ GeV
	• $E_{HAD}/E_{EM} \leq 0.125$
	• $Lshr \leq 0.4$
	Central track with:
	• ${p_{\mathrm{T}}}^{track} > 9$ GeV/c
	• $z \leq 8\,cm$ to the primary vertex position

Table IV: Selection applied to every level for the high-p_{T} central electron trigger. Lateral sharing $Lshr$ is defined in Section 4.3.3.

a trigger that has the same electron requirements as ELECTRON_CENTRAL_18 except the requirements on the track, thus introducing no bias in the track requirements. The efficiency is determined counting how many electrons from the W candidates (N_W) passed each level of the trigger:

$$\begin{aligned}
\epsilon_{\mathrm{L1}}^{\mathrm{trk}} &= \frac{N_W \text{ \& passed Level 1}}{N_W} \\
\epsilon_{\mathrm{L2}}^{\mathrm{trk}} &= \frac{N_W \text{ \& passed Level 1 \& passed Level 2}}{N_W \text{ \& passed Level 1}} \\
\epsilon_{\mathrm{L3}}^{\mathrm{trk}} &= \frac{N_W \text{ \& passed Level 1 \& passed Level 2 \& passed Level 3}}{N_W \text{ \& passed Level 1 \& passed Level 2}}
\end{aligned}$$

The measured tracking efficiencies are $\epsilon_{\mathrm{L1}}^{\mathrm{trk}} = 0.983 \pm 0.003$ for Level 1, $\epsilon_{\mathrm{L2}}^{\mathrm{trk}} = 0.999 \pm 0.001$ for Level 2 and $\epsilon_{\mathrm{L3}}^{\mathrm{trk}} = 0.997 \pm 0.001$ for Level 3. The efficiency is measured as a function of several variables: E_{T}^{e}, ϕ^{e}, η^{e}, jet multiplicity. No dependence is observed except for η^{e}. A systematic uncertainty of 1.5% in the efficiency is introduced to take this effect into account.

The efficiency of the calorimeter trigger selection ϵ^{cal} is measured using an unbiased high-p_{T} muon sample. To determine the Level 1 efficiency, we look for events with at least one energy cluster in the central electromagnetic calorimeter that passes the Level 1 requirements. In order to study Level 2 calorimeter trigger efficiency we count the number of events that have at least 1 electron in the central calorimeter (N_e) (see electron selection cuts in Section 4.3.3) and passed Level 1:

$$\epsilon_{\mathrm{L2}}^{\mathrm{cal}} = \frac{N_e \text{ \& passed Level 1 \& passed Level 2}}{N_e \text{ \& passed Level 1}} \tag{4.1}$$

For Level 3, we use a special trigger (L1_CEM8, L2_CEM16, L3_CEM18) that has the same calorimeter requirements as ELECTRON_CENTRAL_18 but no requirements on the track of the electrons:

$$\epsilon_{\mathrm{L3}}^{\mathrm{cal}} = \frac{N_e \text{ \& passed L1\_CEM8 \& passed L2\_CEM16 \& passed L3\_CEM18}}{N_e \text{ \& passed L1\_CEM8 \& passed L2\_CEM16}} \tag{4.2}$$

For electrons with $E_{\mathrm{T}}^{e} > 25$ GeV the calorimeter trigger efficiencies for the three levels, $\epsilon_{\mathrm{L1}}^{\mathrm{cal}}$, $\epsilon_{\mathrm{L2}}^{\mathrm{cal}}$ and $\epsilon_{\mathrm{L3}}^{\mathrm{cal}}$ are 100% and flat in E_{T}^{e}.

Combining the tracking and calorimeter efficiencies we obtain a global trigger efficiency per electron of $\epsilon_{\mathrm{E18}} = 0.979 \pm 0.003$.

4.2 Monte Carlo Samples

To correct the measurements to the hadron level, samples of $Z/\gamma^{*}(\rightarrow e^{+}e^{-}) + \text{jets}$ events are produced using PYTHIA 6.216. It generates LO $2 \rightarrow 2$ process *parton* + *parton* $\rightarrow$

$Z/\gamma^* + parton$ and develops a parton shower. $\Lambda_{\rm QCD}$ is set to 0.146 GeV to provide a proper modeling of the energy flows measured in the data and CTEQ5L PDFs [54] are used. All the event samples created with PYTHIA and CTEQ5L has been reweighted in $\hat{p}_{\rm T}$ and $E_{\rm T}^{e,\rm fwd}$ as it is described in detail in Section 4.6. The generation has been carried out for different $\hat{p}_{\rm T}$ ranges to ensure sufficient statistics in all the range of jet transverse momenta. Finally, the events are passed through the CDF detector simulation to reproduce the detector response. Several studies have been carried out to examine carefully the detector simulation [53] and the jet and electron reconstruction (Section 4.3.4) in the Monte Carlo generated events.

Different tunings of the underlying event model are used in the analysis. The sample of events used in the unfolding of the measurements back to the hadron level are generated with TUNE A, and TUNE DW is used to estimate the uncertainties in the determination of the non-perturbative contributions to the theoretical cross section. A separated sample with the underlying event switched off (*no UE*) is used to determine the underlying event effects in the cross section.

PYTHIA Monte Carlo is also used to estimate the non-QCD background contributions to the $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ signal (Section 4.5.2): $Z/\gamma^*(\rightarrow \tau^+\tau^-) + \text{jets}$, dibosons ($WW$, ZZ, ZW), $t\bar{t}$ and $Z/\gamma^*(\rightarrow e^+e^-) + \gamma$. All the background samples are generated using PYTHIA-TUNE A.

4.3 Electron Selection and Z/γ^* Reconstruction

Z/γ^* boson candidates are identified via the presence of two high-$p_{\rm T}$ electrons. Due to the geometry of the electromagnetic calorimeters in CDF, electrons are required to be either both in the central region of the calorimeter ($|\eta| < 1.0$, Central-Central combinations, CC), or one electron in the central region and the other in the forward (plug) region ($1.2 < |\eta| < 2.8$, Central-Forward combinations, CF). Some steps in the analysis require the CC and CF events to be treated differently, but, otherwise explicitly indicated, the final results are always the combination of CC and CF configurations.

4.3.1 Event Trigger Efficiency

The electron selection used in this analysis (see below, Section 4.3.3) is more restrictive than the trigger selection. In the case of the CC combination, the trigger efficiency per event needs special care to properly compute the relevant combinations because any of the two electrons could fire the trigger.

If $\epsilon_{\rm E18}$ is the trigger efficiency per electron, the probability that the event is triggered by both electrons is $(\epsilon_{\rm E18})^2$, and the probability that only one of the two electrons triggers the event is $\epsilon_{\rm E18}\,(1-\epsilon_{\rm E18})$. The global trigger efficiency for the CC selection becomes the combination of these three possible cases.

$$\epsilon_{\rm trig}^{\rm CC} = (\epsilon_{\rm E18})^2 + 2\cdot\epsilon_{\rm E18}\,(1-\epsilon_{\rm E18}) = \epsilon_{\rm E18}\,(2-\epsilon_{\rm E18}) = 0.9996 \pm 0.0001 \tag{4.3}$$

For the CF configuration, the global trigger efficiency is simply:

$$\epsilon_{\rm trig}^{\rm CF} = \epsilon_{\rm E18} = 0.979 \pm 0.003 \tag{4.4}$$

Whenever necessary, the trigger efficiencies for CC and CF are combined bin by bin in the distributions taking into account the proportion of CC and CF selected events in every bin. All the trigger efficiency values are summarized in Table V.

	Trigger efficiency	**Systematic uncertainty**
Per electron	0.979 ± 0.003	1.5%
Per CC event	0.9996 ± 0.0001	0.06%
Per CF event	0.979 ± 0.003	1.5%

Table V: Summary of the trigger efficiency values and uncertainties per electron and per event. See Section 4.7 for the determination of the systematic uncertainties.

4.3.2 Electron Energy Corrections

The energy of the electrons is determined adding the energy of the calorimeter towers that form the electron cluster as described in Section 3.3.1. This energy is corrected afterward for several effects:

- a *face correction* takes into account the variation of the response of the calorimeter from wedge to wedge in ϕ.
- the energy deposited in the pre-radiator detectors is added to the energy of the electron.
- a parametrization has been done using testbeam data to estimate the amount of shower energy that is not contained in the energy cluster. The amount of this *leakage energy* depends on the position of the center of the electron shower inside the calorimeter tower.
- the energy of the electrons is scaled separately in data and in Monte Carlo events to obtain a reconstructed invariant mass M_{ee} that peaks at 91 GeV/c^2. The obtained scale factors are reported in Table VI for different η regions, and different data taking periods.

			Data			MC
			0d	0h	0i	
-2.8	$< \eta^e <$	-1.78	1.007	1.014	1.012	0.998
-1.78	$< \eta^e <$	-1.2	1.015	1.014	1.012	0.996
-1.0	$< \eta^e <$	1.0	1.000	1.0025	1.0042	0.996
1.2	$< \eta^e <$	1.78	1.020	1.018	1.016	0.996
1.78	$< \eta^e <$	2.8	1.010	1.018	1.016	0.998

Table VI: Energy scale factors for central and forward electrons. 0d, 0h and 0i refer to different periods of data taking containing approximately the same amount of integrated luminosity.

4.3.3 Event Selection

In this analysis we use three sets of electron identification criteria: for *tight*, *loose* and *plug* electrons. The individual cuts on every set are shown in Table VII. Tight and loose cuts are used to select electrons in the central region ($|\eta| < 1.0$), and plug cuts for electrons in the forward region ($1.2 < |\eta| < 2.8$). All electrons are required to have $E_{\mathrm{T}} > 25$ GeV. The central electrons (tight and loose) must have a reconstructed track of $p_{\mathrm{T}}^{\mathrm{trk}} > 10$ GeV/c pointing to the energy cluster in the calorimeter. The track must also fulfill certain quality cuts that require a given number of hits in every superlayer (SL) of the COT and to intersect the beamline (z_0^{trk}) within 60 cm of the nominal interaction point. Fiduciality cuts avoid reconstructing electrons near non-instrumented regions of the calorimeters and shower maximum detectors. χ^2 cuts compare the shower profile to typical electromagnetic showers. Lateral sharing ($Lshr$) measures how well the lateral shower development matches that expected for an electromagnetic cascade. It is defined as $Lshr = 0.14 \times \frac{\sum_i M_i - P_i}{\sqrt{(0.14 \times E_{\mathrm{EM}})^2 + \sum_i (\Delta P_i)^2}}$ and compares the measured energy M in the tower adjacent to the seed tower with the predicted energy P in those towers. An algorithm uses the track information to reject electrons from $\gamma \rightarrow e^+e^-$ conversions. An additional set of electron selection cuts, *very tight*, are used to measure the electron ID efficiency (Section 4.3.4). It consists of the same tight cuts with more stringent requirements in E/p^{trk} ($E/p^{\mathrm{trk}} < 1.2$ or $p_{\mathrm{T}}^{\mathrm{trk}} > 50$ GeV/c) and $Lshr$ ($Lshr < 0.1$).

Events are selected using the following selection criteria:

- at least one reconstructed primary vertex with z-position within 60 cm of the nominal interaction point.
- one central electron passing tight cuts.
- a second central electron passing loose cuts (CC configuration) or a forward electron

Variable	**Tight cuts**	**Loose cuts**	**Plug cuts**
$\lvert\eta\rvert$	< 1.0		> 1.2 and < 2.8
$\lvert z_0^{\mathrm{trk}}\rvert$	$\leq 60\,cm$		
$p_{\mathrm{T}}^{\mathrm{trk}}$	≥ 10 GeV/c		
E_{T}	> 25 GeV		
Track quality	≥ 3 axial SL with ≥ 5 hits per SL ≥ 2 stereo SL with ≥ 5 hits per SL		
Had/Em	$\leq 0.055 + (0.00045 \times E)$		≤ 0.05
CES Fiduciality	$= 1$		
E/p^{trk}	≤ 2 or $p_{\mathrm{T}}^{\mathrm{trk}} > 50$ GeV/c		
$Lshr$	≤ 0.2		
PEM χ^2			≤ 10
CES χ^2	≤ 10		
CES $\Delta x \cdot Q$	≥ -3.0 cm and ≤ 1.5 cm		
CES Δz	< 3.0 cm		
Conversion	$=$ FALSE		

Table VII: List of electron selection cuts. Fiduciality cuts ensure that the energy is measured in a well instrumented region of the calorimeter.

passing plug cuts (CF configuration).

- the invariant mass of the two electrons must be in the range $66 < M_{ee} < 116$ GeV/c^2.

To avoid biasing the measurements of high-p_{T} events, electron pairs are not required to have opposite charge. The accuracy in the determination of the charge of a particle decreases rapidly with the p_{T} of the particle when its track becomes straighter. Such requirement would lead to a loss of efficiency at high-p_{T}. Also, no requirement on electron isolation is imposed, because when the p_{T} of the Z/γ^* bosons increases, electrons are boosted and come close to each other. Figure 4.1 shows the separation between the two electrons in $\eta - \phi$ as a function of $p_{\mathrm{T,corr}}^{\mathrm{jet}}$ in data and Monte Carlo. A cut on isolation of the electrons would immediately translate into an effective cut on minimum ΔR between the two electrons which suppresses high-$p_{\mathrm{T}}^{\mathrm{jet}}$ events. An isolation requirement introduces an additional sensitivity to small imperfections in the description of the soft-gluon radiation implemented in the Monte Carlo samples.

The measurements are defined in a limited but well defined kinematic range for the Z/γ^* decay products: $E_{\mathrm{T}}^{e} > 25$ GeV, $\lvert\eta^{e_1}\rvert < 1.0$, $\lvert\eta^{e_2}\rvert < 1.0$ or $1.2 < \lvert\eta_{e_2}\rvert < 2.8$ and $66 < M_{ee} < 116$ GeV/c^2, where final-state QED radiation effects are corrected for.

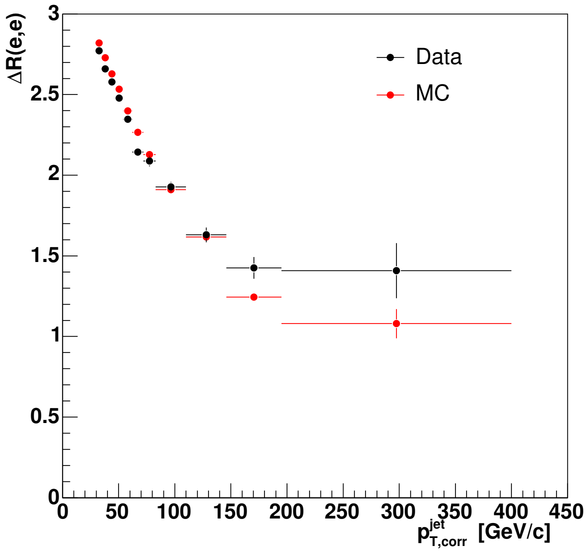


Figure 4.1: ΔR between the two electrons as function of the p_{T} of the leading jet.

4.3.4 Electron ID Efficiency

The electron ID efficiency could be affected by the presence of hadronic jets in the event. The efficiency is estimated, in data and Monte Carlo samples, as a function of jet multiplicity and $p_{\mathrm{T,corr}}^{\mathrm{jet}}$, and a relative scale factor Data/MC is determined.

In the measurement of the efficiency we use a very pure $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ sample obtained using a special event selection:

- events are required to have at least one reconstructed primary vertex with z-position within 60 cm of the nominal interaction point.
- one central electron passing *very tight* cuts, and a *probe* electron passing only geometrical and kinematic cuts ($E_{\mathrm{T}}^{e} > 25$ GeV and $|\eta| < 1.0$ or $1.2 < |\eta| < 2.8$), with an invariant mass M_{ee} between 86 and 96 GeV/c^2.
- the missing transverse energy $\not\!\!E_{\mathrm{T}}$[15] in the event must satisfy $\not\!\!E_{\mathrm{T}} < 20$ GeV to avoid contamination form $W(\rightarrow e + \nu) + \text{jets}$ events.

The very tight requirements on the first electron ensure the presence of a real electron in the event, and the narrower invariant mass window ($86 - 96$ GeV/c^2) makes the background rejection more efficient.

---

[15] The missing transverse energy is defined as the norm of $-\sum E_{\mathrm{T}}^{i} \cdot \vec{n}_i$, where $\vec{n}_i$ is the unit vector on the azimuthal plane that points from the beam line to the ith calorimeter tower.

The corresponding electron ID efficiency is measured applying the tight, loose and plug selection cuts to the probe electron. In the case of the plug electrons, the efficiency ϵ_P is simply given by the fraction of probe electrons in the forward region that pass the plug cuts, i.e. the ratio of the number of very tight and plug electron pairs $N_{\mathrm{vT-P}}$ over the number of very tight and probe forward electrons pairs $N_{\mathrm{vT-pF}}$:

$$\epsilon_P = \frac{N_{\mathrm{vT-P}}}{N_{\mathrm{vT-pF}}} \tag{4.5}$$

For the central electrons efficiencies, the measurement requires further considerations, related to the fact that the very tight electrons are a subsample of the tight electrons, which at the same time are a subsample of the loose electrons. If $\mathcal{P}_{\mathrm{vT}}$ is the probability for an electron to pass the very tight cuts, $\mathcal{P}_{\mathrm{T}}$ the probability to pass the tight cuts but not the very tight cuts, and $\mathcal{P}_{\mathrm{L}}$ the probability to pass the loose cuts but not the tight cuts, these probabilities are related to the measured efficiencies as:

$$\begin{aligned} \mathcal{P}_{\mathrm{vT}} &= \epsilon_{\mathrm{vT}} && (4.6)\\ \mathcal{P}_{\mathrm{T}} &= \epsilon_{\mathrm{T}} - \epsilon_{\mathrm{vT}} && (4.7)\\ \mathcal{P}_{\mathrm{L}} &= \epsilon_{\mathrm{L}} - \epsilon_{\mathrm{T}} && (4.8) \end{aligned}$$

The number of electron pairs for the different combinations are defined in terms of probabilities and efficiencies as:

$$\begin{aligned} N_{\mathrm{vT-vT}} &= N \cdot \mathcal{P}_{\mathrm{vT}} \cdot \mathcal{P}_{\mathrm{vT}} = N \cdot \epsilon_{\mathrm{vT}}^2 \\ N_{\mathrm{vT-T}} &= N \cdot [\mathcal{P}_{\mathrm{vT}} \cdot \mathcal{P}_{\mathrm{vT}} + 2 \cdot \mathcal{P}_{\mathrm{vT}} \cdot \mathcal{P}_{\mathrm{T}}] \\ &= N \cdot \epsilon_{\mathrm{vT}} \cdot (2\epsilon_{\mathrm{T}} - \epsilon_{\mathrm{vT}}) \\ N_{\mathrm{vT-L}} &= N \cdot [\mathcal{P}_{\mathrm{vT}} \cdot \mathcal{P}_{\mathrm{vT}} + 2 \cdot \mathcal{P}_{\mathrm{vT}} \cdot \mathcal{P}_{\mathrm{T}} + 2 \cdot \mathcal{P}_{\mathrm{vT}} \cdot \mathcal{P}_{\mathrm{L}}] \\ &= N \cdot \epsilon_{\mathrm{vT}} \cdot (2\epsilon_{\mathrm{L}} - \epsilon_{\mathrm{vT}}) \\ N_{\mathrm{vT-pC}} &= N \cdot [\mathcal{P}_{\mathrm{vT}} \cdot \mathcal{P}_{\mathrm{vT}} + 2 \cdot \mathcal{P}_{\mathrm{vT}} \cdot (1 - \mathcal{P}_{\mathrm{vT}})] \\ &= N \cdot \epsilon_{\mathrm{vT}} \cdot (2 - \epsilon_{\mathrm{vT}}) \end{aligned}$$

where *pC* stands for *probe Central electrons*, and N is the total number of events in the sample. Resolving the equations, ϵ_{T} and ϵ_{L} can be expressed as:

$$\begin{aligned} \epsilon_{\mathrm{T}} &= \frac{N_{\mathrm{vT-vT}} + N_{\mathrm{vT-T}}}{N_{\mathrm{vT-vT}} + N_{\mathrm{vT-pC}}} && (4.9)\\ \epsilon_{\mathrm{L}} &= \frac{N_{\mathrm{vT-vT}} + N_{\mathrm{vT-L}}}{N_{\mathrm{vT-vT}} + N_{\mathrm{vT-pC}}} && (4.10) \end{aligned}$$

The efficiencies per electron as a function of $p_{\mathrm{T,corr}}^{\mathrm{jet}}$ in events with at least one jet in data and in Monte Carlo events and the relative scale factors are shown in Fig. 4.2 and summarized in Table VIII. Combining the efficiencies on identifying single electrons, we determine the efficiency of reconstructing a Z/γ^* boson. In the CC case it is calculated as follows:

$$\epsilon^{\mathrm{CC}}_{Z/\gamma^*} = \epsilon_{\mathrm{T}} \cdot \epsilon_{\mathrm{T}} + 2\epsilon_{\mathrm{T}} \cdot (\epsilon_{\mathrm{L}} - \epsilon_{\mathrm{T}}) = \epsilon_{\mathrm{T}} \cdot (2\epsilon_{\mathrm{L}} - \epsilon_{\mathrm{T}}) \tag{4.11}$$

And for the CF case:

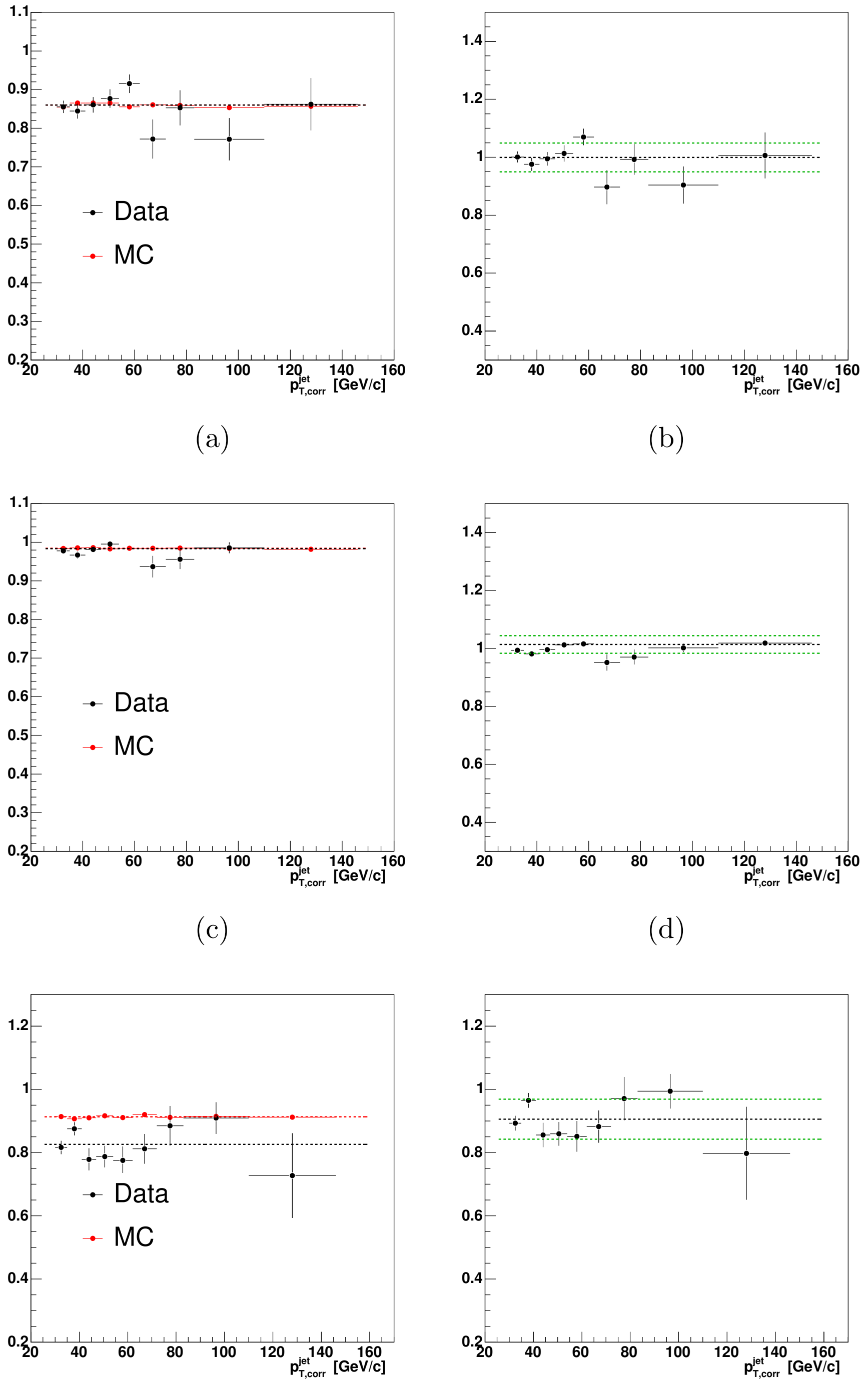
$$\epsilon^{\mathrm{CF}}_{Z/\gamma^*} = \epsilon_{\mathrm{T}} \cdot \epsilon_{\mathrm{P}} \tag{4.12}$$

$Z/\gamma^*(\to e^+e^-)+\geq 1$ jet	**Data**	**Monte Carlo**	$\kappa^{\mathrm{MC}}_{\mathrm{Data}}$ **=Data/MC**	**Sys1**	**Sys2**	**Total Sys**
Tight	0.856 ± 0.009	0.861 ± 0.001	0.994 ± 0.012	5%	0.3%	5.01%
Loose	0.976 ± 0.004	0.9842 ± 0.0004	0.992 ± 0.003	3%	0.4%	3.03%
Plug	0.821 ± 0.012	0.914 ± 0.001	0.898 ± 0.013	7%	1.3%	7.12%
Z/γ^* (CC)	0.939 ± 0.014	0.953 ± 0.002	0.985 ± 0.015	5%	0.5%	5.02%
Z/γ^* (CF)	0.703 ± 0.013	0.787 ± 0.001	0.893 ± 0.016	5%	1.33%	5.17%
$Z/\gamma^*(\to e^+e^-)+\geq 2$ jets	**Data**	**Monte Carlo**	$\kappa^{\mathrm{MC}}_{\mathrm{Data}}$ **=Data/MC**	**Sys1**	**Sys2**	**Total Sys**
Tight	0.819 ± 0.034	0.851 ± 0.002	0.963 ± 0.041	5%	0.3%	5.01%
Loose	0.965 ± 0.015	0.981 ± 0.001	0.984 ± 0.016	3%	0.4%	3.03%
Plug	0.814 ± 0.042	0.907 ± 0.004	0.898 ± 0.046	7%	1.3%	7.12%
Z/γ^* (CC)	0.910 ± 0.054	0.945 ± 0.004	0.963 ± 0.057	5%	0.5%	5.02%
Z/γ^* (CF)	0.678 ± 0.043	0.782 ± 0.004	0.867 ± 0.055	5%	1.33%	5.17%
$Z/\gamma^*(\to e^+e^-)+\geq 3$ jets	**Data**	**Monte Carlo**	$\kappa^{\mathrm{MC}}_{\mathrm{Data}}$ **=Data/MC**	**Sys1**	**Sys2**	**Total Sys**
Tight	0.83 ± 0.16	0.834 ± 0.013	1.00 ± 0.20	5%	0.3%	5.01%
Loose	1.0 ± 0.5	0.980 ± 0.005	1.021 ± 0.005	3%	0.4%	3.03%
Plug	1.0 ± 0.5	0.911 ± 0.013	1.100 ± 0.016	7%	1.3%	7.12%
Z/γ^* (CC)	0.97 ± 0.23	0.939 ± 0.020	1.04 ± 0.25	5%	0.5%	5.02%
Z/γ^* (CF)	0.83 ± 0.16	0.779 ± 0.016	1.07 ± 0.21	5%	1.33%	5.17%

Table VIII: Efficiencies in data and Monte Carlo for $Z/\gamma^*+\geq$1, 2 and 3 jets, and the relative scale factor $\kappa^{\mathrm{MC}}_{\mathrm{Data}}$. *Sys1* are the uncertainties assigned due to the dependence on the p_{T} of the jets, and *Sys2* are other uncertainties on the global electron efficiencies. Both are explained in Section 4.7.

The efficiencies on reconstructing the Z/γ^* boson for the CC and CF case in events with at least one jet in data and in Monte Carlo events and their relative scale factor $\kappa^{\mathrm{MC}}_{\mathrm{Data}}$ are shown in Fig. 4.3 and the values are reported in Table VIII. There is no indication of any dependence of the efficiency on $p^{\mathrm{jet}}_{\mathrm{T,corr}}$. However, due to the limited statistics in the measurements, effects smaller than 5% cannot be excluded and it is taken as a systematic uncertainty.

In the cases of $Z/\gamma^*(\to e^+e^-)+\geq 2$ jets and $Z/\gamma^*(\to e^+e^-)+\geq 3$ jets, there is not enough statistics to carry out a precise study of the $p^{\mathrm{jet}}_{\mathrm{T,corr}}$ dependence of the efficiency. However, the global efficiency values shown in Table VIII are in agreement with the efficiencies found

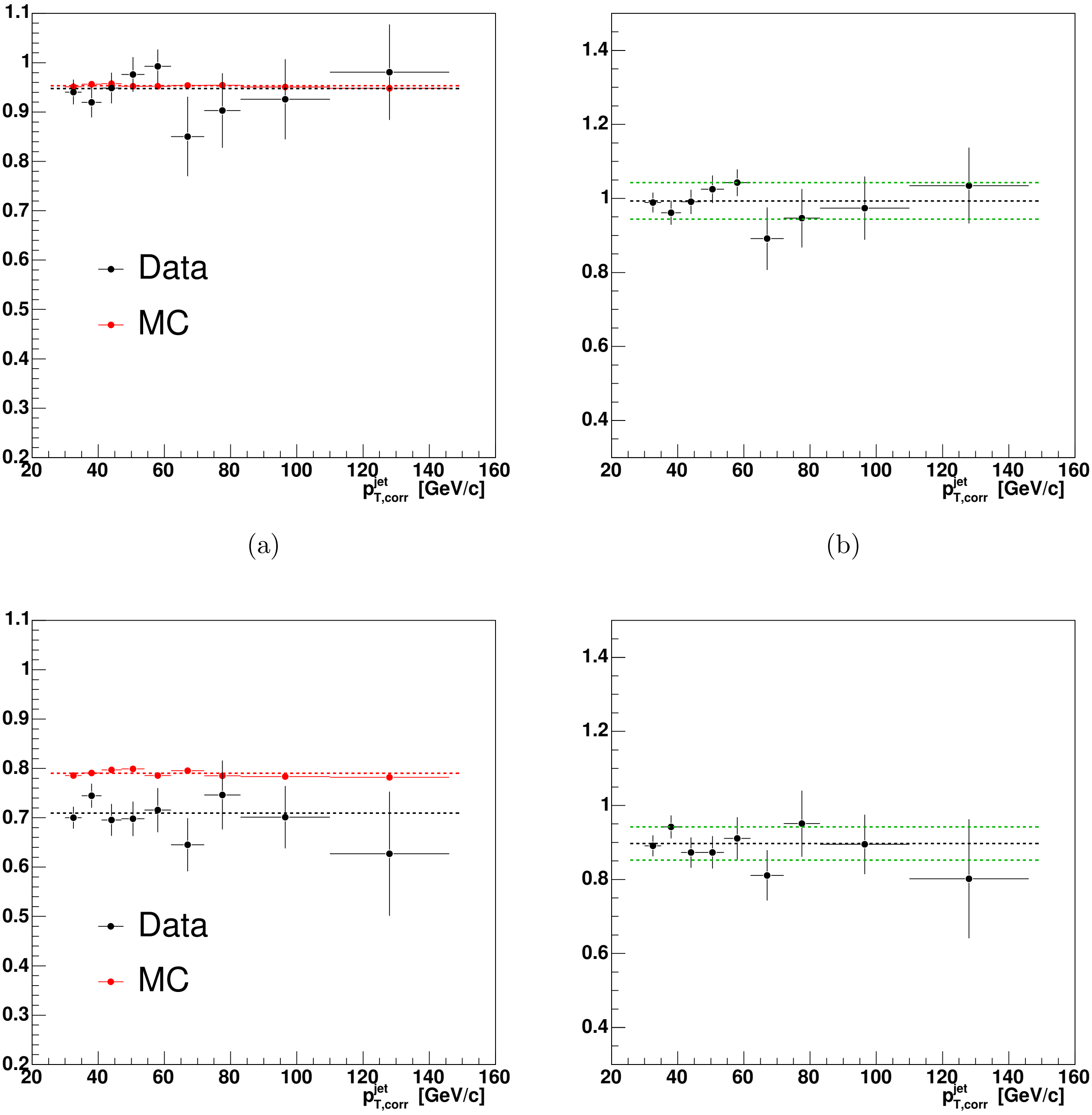


(a) (b)

(c) (d)

(e) (f)

Figure 4.2: Electron ID efficiency as a function of the p_T of the jet in events with at least one jet, in the data and in the Monte Carlo for (a) tight, (c) loose and (e) plug electrons. Scale factor Data/MC for (b) tight, (d) loose and (f) plug electrons. The dashed lines correspond to an uncertainty of 5% for the tight electrons, 3% for the loose electrons and 7% for the plug electrons.



(a) (b)

(c) (d)

Figure 4.3: Z/γ^* boson ID efficiency in (a) CC and (c) CF cases, as a function of the p_T of the jet in events with at least one jet, in the data and in the Monte Carlo. Scale factor Data/MC for (b) CC and (d) CF events. The dashed lines correspond to an uncertainty of 5% on both cases.

in Z/γ^* events with $\geq$ 1 jet. They also agree with the efficiency values obtained in previous studies at CDF in inclusive Z/γ^* samples [55]. Therefore, the values obtained for $Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet will be used in all the cases.

4.3.5 Inclusive $Z/\gamma^* \rightarrow e^+e^-$ Cross Section

As an ultimate check of our understanding of the Z/γ^* selection and reconstruction, we measured the inclusive $Z/\gamma^* \rightarrow e^+e^-$ production in $66 < M_{ee} < 116$ GeV/c^2. We consider the following processes as background contributors:

- QCD backgrounds estimated from data: QCD inclusive jets and $W(\rightarrow e+\nu)+\text{jets}$
- Non-QCD backgrounds estimated from Monte Carlo samples: $Z/\gamma^* \rightarrow \tau^+\tau^-$, WW, WZ, ZZ and $t\bar{t}$ production,

where the technique used to estimated the number of expected background events is extensively described in Section 4.5. Therefore, this measurement is also a cross check of the background estimation procedure. In 1.73 fb^{-1} of data we found 90491 Z/γ^* boson candidates, where we expect 3557 ± 75 events from the background processes that represent 4% of the sample.

In this particular case, and only to compare the measurement with previous CDF results, the inclusive cross section is not restricted in the phase space of the electrons considered for our analysis, and instead is corrected for the kinematic cuts applied in the event selection using the acceptance $\mathcal{A}$:

$$\mathcal{A} = \frac{N_Z^{\text{gen}}\ (66 < M_{ee} < 116 \text{ GeV/c}^2)}{N_Z^{\text{cal}}\ (66 < M_{ee} < 116 \text{ GeV/c}^2, E_{\text{T}}^e, \eta^e)} \qquad (4.13)$$

where the numerator is the number of generated $Z/\gamma^* \rightarrow e^+e^-$ events within the mass range; and the denominator is the number of reconstructed events at calorimeter level that pass our event selection: $66 < M_{ee} < 116$ GeV/c$^2, E_T^e > 25$ GeV$, |\eta^{e_1}| < 1.0, |\eta^{e_2}| < 1.0$ or $1.2 < |\eta^{e_2}| < 2.8$. The acceptance is determined from a Monte Carlo sample of Pythia 6.216 $Z/\gamma^* \rightarrow e^+e^-$ events that gives a value of $\mathcal{A} = 4.662 \pm 0.009$.

The inclusive cross section is determined with the following formula:

$$\sigma = \frac{\mathcal{A}}{\mathcal{L}} \left(\frac{N_Z^{\text{cal}} - N_{\text{QCD bkg}}}{\kappa_{\text{Data}}^{\text{MC}} \cdot \epsilon_{\text{trig}}} - N_{\text{non-QCD bkg}} \right) \qquad (4.14)$$

After the number of expected background events $N_{\text{QCD bkg}}$ and $N_{\text{non-QCD bkg}}$ is subtracted from the observed events in the data N_Z^{cal}, the cross section is corrected for the acceptance $\mathcal{A}$ and divided by the integrated luminosity $\mathcal{L}$. The quantities obtained from the data (N_Z^{cal} and

$N_{\text{QCD bkg}}$) are corrected for the remaining differences between data and Monte Carlo in the selection efficiency $\kappa^{\text{MC}}_{\text{Data}}$ (Section 4.3.4) and for the trigger inefficiencies ϵ_{trig} (Section 4.3.1). Figure 4.4 shows the invariant mass of the two electrons in the inclusive selection. We obtain an inclusive cross section of:

$$\sigma(p\overline{p} \rightarrow Z/\gamma^* \rightarrow e^+e^-) = 254.1 \pm 3.3 \text{ pb} \qquad 66 < M_{ee} < 116 \text{ GeV/c}^2 \tag{4.15}$$

where only statistical uncertainties have been quoted. The number is in agreement with the previously measured cross section at CDF of $255.8 \pm 3.9(\text{stat.})^{+5.5}_{-5.4}(\text{syst.})$ pb [55] and with the theoretical NNLO prediction of 251.3 ± 5.0 pb [56]. The good agreement in the measured inclusive cross section illustrates the solid understanding of the Z/γ^* boson selection, the efficiencies determination and the background estimation procedures (see Section 4.5).

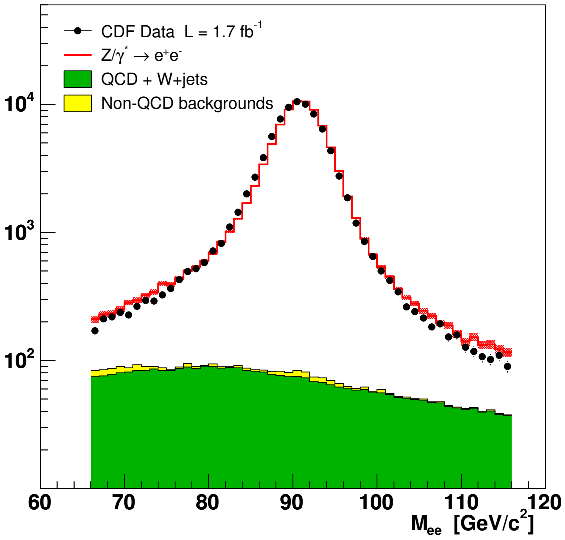


Figure 4.4: Z/γ^* invariant mass for the inclusive selection.

4.4 Jet Reconstruction and Selection

Jets are search for using the Midpoint algorithm (Section 2.9). In data and Monte Carlo simulated events, jets are reconstructed using calorimeter towers after excluding those towers associated with the reconstructed electrons from the Z/γ^* boson in the final state. Searching for jets without first removing the electron clusters would bias the kinematic distributions of the jets in a non trivial way. The jet algorithm would first identify a jet around the electron clusters and, if these clusters are close to real hadronic jets, in the merging/splitting process the algorithm could decide to merge them into a single jet together with the electron, thus affecting the number of jets in the event. Moreover, since the merging of two jets depends on the energy that their cones share, and the electron spectrum peaks at 45 GeV (Jacobian peak), this would also bias the resulting p_{T} spectrum of the jets. Figure 4.5 shows the difference between the raw cross section (not corrected for inefficiencies and without background subtraction) for jets clustered with the electron towers being removed and jets clustered with the presence of the electromagnetic clusters of the electrons from the Z/γ^* decay. The difference is about 5%-10% at low $p_{\mathrm{T}}^{\mathrm{jet}}$ and negligible at high-$p_{\mathrm{T}}^{\mathrm{jet}}$. In addition, and for a proper comparison with theoretical predictions, we also require a minimum distance between the jets and the electrons of $\Delta R_{\mathrm{e-jet}} > 0.7$.

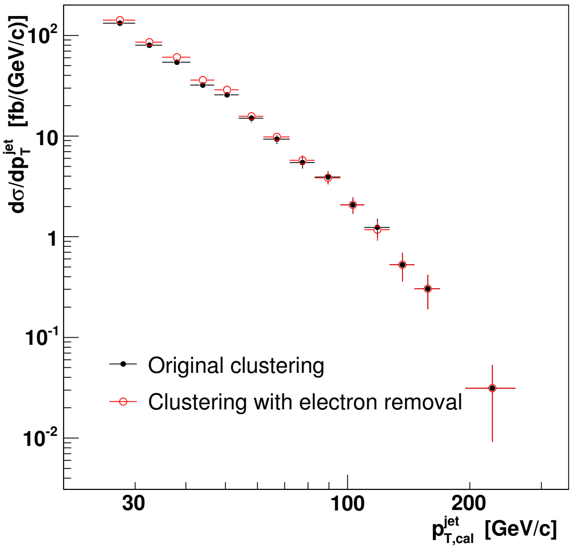


Figure 4.5: Detector level cross section not corrected for inefficiencies and without background subtraction, using only the first 368 pb^{-1} of data, for jets clustered with the electron towers (full circles) and jet clustered after removing the electron towers from the calorimeter (open circles). The maximum difference is 10% at low $p_{\mathrm{T}}^{\mathrm{jet}}$.

After the reconstruction, jets are corrected on average in $p_{\mathrm{T,cal}}^{\mathrm{jet}}$ (Section 3.3.2). We select jets with $p_{\mathrm{T,corr}}^{\mathrm{jet}} > 30$ GeV/c and $|y^{\mathrm{jet}}| < 2.1$. The $|y^{\mathrm{jet}}|$ cut ensures that jet cones are completely inside the calorimeter coverage and avoids contributions from proton and antiproton remnants. The same kinematic requirements ($p_{\mathrm{T}}^{\mathrm{jet}}$, $|y^{\mathrm{jet}}|$ and $\Delta R_{\mathrm{e-jet}}$ cuts) are applied to particle jets at the hadron level.

4.4.1 Dependence on Instantaneous Luminosity

We measure jets in a wide range of instantaneous luminosities ($20 - 300{\cdot}10^{30}$ cm^{-2}s^{-1}). As multiple $p\overline{p}$ interactions increase with luminosity, we check that we properly corrected the jets for this effect such that our measurements are not dependent on the instantaneous luminosity. Figure 4.6 shows the number of reconstructed $Z/\gamma^*(\rightarrow e^+e^-)+ \geq N_{\rm jets}$ events as a function of the jet multiplicity $N_{\rm jets}$ for different instantaneous luminosity ranges. For each range, the number of events is normalized to the number of events in the lowest instantaneous luminosity range ($20 - 40{\cdot}10^{30}$ cm^{-2}s^{-1}) where negligible pile-up is expected. No dependency on the instantaneous luminosity is observed.

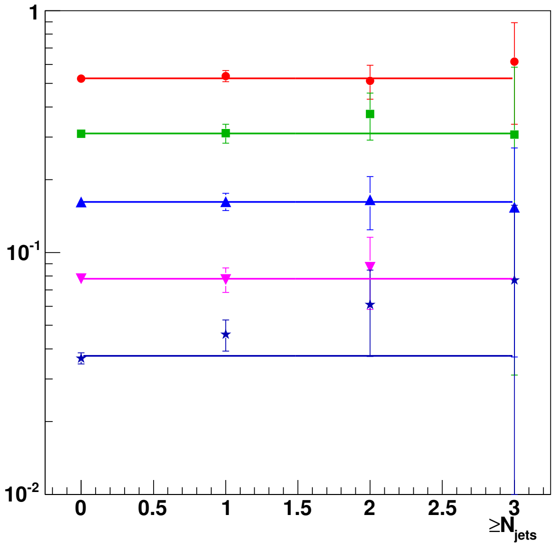


Figure 4.6: Number of $Z/\gamma^*(\rightarrow e^+e^-)+ \geq N_{\rm jets}$ events in different instantaneous luminosity ranges, as a function of the jet multiplicity. In red circles, number of $Z/\gamma^*(\rightarrow e^+e^-)+ \geq N_{\rm jets}$ events with instantaneous luminosity between $40 - 60{\cdot}10^{30}$ cm^{-2}s^{-1} ($N_{Z/\gamma^*}(40-60)$), over $N_{Z/\gamma^*}(20-40)$. In green squares, $N_{Z/\gamma^*}(60-80)/N_{Z/\gamma^*}(20-40)$. In light blue triangles, $N_{Z/\gamma^*}(80-100)/N_{Z/\gamma^*}(20-40)$. In pink inverted triangles, $N_{Z/\gamma^*}(100-120)/N_{Z/\gamma^*}(20-40)$. In dark blue stars, $N_{Z/\gamma^*}(>120)/N_{Z/\gamma^*}(20-40)$.

Two additional observables are measured to cross check the independence on the instantaneous luminosity: the average number of jets per $Z/\gamma^* \rightarrow e^+e^-$ event (Fig. 4.7) and the relative population of the different jet multiplicities (Fig. 4.8) $\mathrm{R}_{(2+3)/1}$ defined as

$$R_{(2+3)/1} = \frac{N_{Z/\gamma^{*}+2\text{ jets}} + N_{Z/\gamma^{*}+3\text{ jets}}}{N_{Z/\gamma^{*}+1\text{ jet}}} \tag{4.16}$$

Both measurements are again compatible with no luminosity dependence.

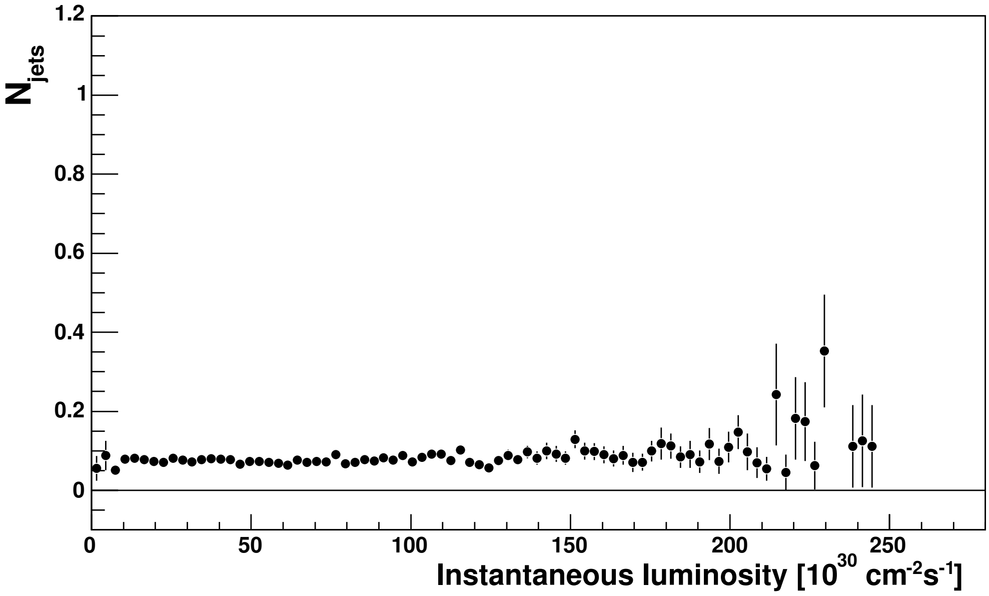


Figure 4.7: Number of jets per $Z/\gamma^{*} \to e^{+}e^{-}$ event as a function of instantaneous luminosity. A fit to a constant gives $\chi^2/\mathrm{ndf} = 68.1/67$.

4.5 Background Estimation

The two major background contributors, QCD inclusive jets and $W(\to e+\nu)+\text{jets}$, are estimated using a data-driven method to identify jets faking electrons. As already mentioned, background contributions from other electroweak processes, such as $Z/\gamma^{*}(\to \tau^{+}\tau^{-})+\text{jets}$, dibosons ($WW$, WZ and WZ), or $t\bar{t}$ events, are estimated using Monte Carlo samples.

4.5.1 QCD and $W(\to e+\nu)+\text{jets}$ Backgrounds

In QCD events, two jets fake the two electrons used to reconstruct Z/γ^{*} boson candidate, and in $W(\to e+\nu)+\text{jets}$ events, the real electron plus one jet faking an electron, are used to reconstruct a Z/γ^{*} candidate. A data-driven method to measure the probability per jet of faking an electron is used. The fake rates are then convoluted with the observed

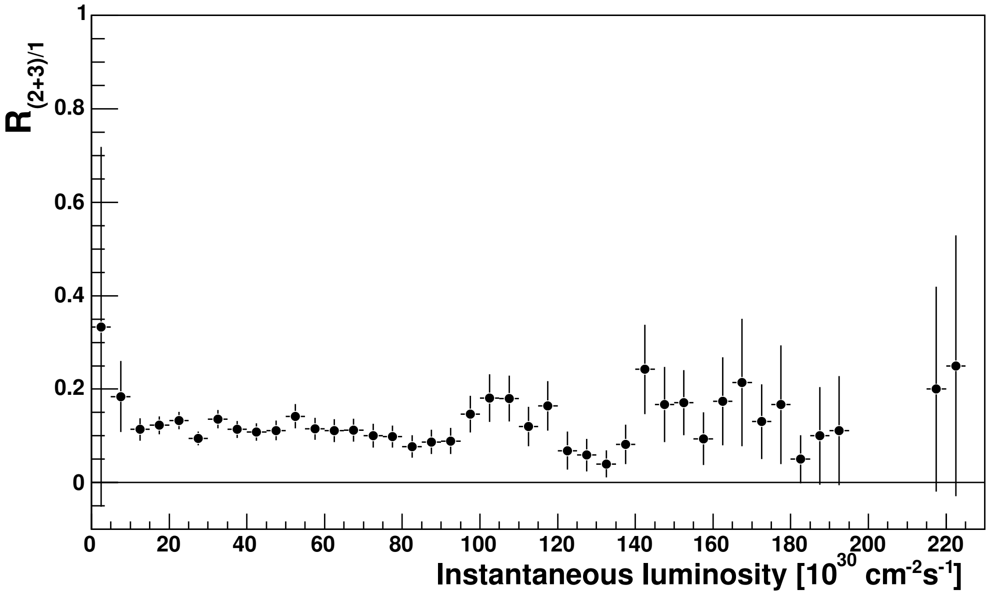


Figure 4.8: Proportion of $Z/\gamma^* \rightarrow e^+e^- + 2$ jets and $Z/\gamma^* \rightarrow e^+e^- + 3$ jets events over $Z/\gamma^* \rightarrow e^+e^- + 1$ jet events as a function of instantaneous luminosity. $\text{R}_{(2+3)/1}$ is defined in Eq. 4.16. The fit to a constant gives $\chi^2/\text{ndf} = 23.1/35$.

jet distributions to give an estimate of the total background. Although the fake probability is small ($\sim 10^{-4} - 10^{-2}$) contributions cannot be neglected due to the large QCD and $W(\rightarrow e + \nu) + \text{jets}$ cross sections compared to that for $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ processes.

4.5.1.1 Fake Rate

The rate of jets faking electrons f_e^{jet} is defined as the probability that a jet passes a set of given electron selection cuts, and it is obtained from the expression:

$$f_e^{\text{jet}} = \frac{N(\text{jets that pass electron cuts})}{N(\text{jets suitable to pass electron cuts})} \tag{4.17}$$

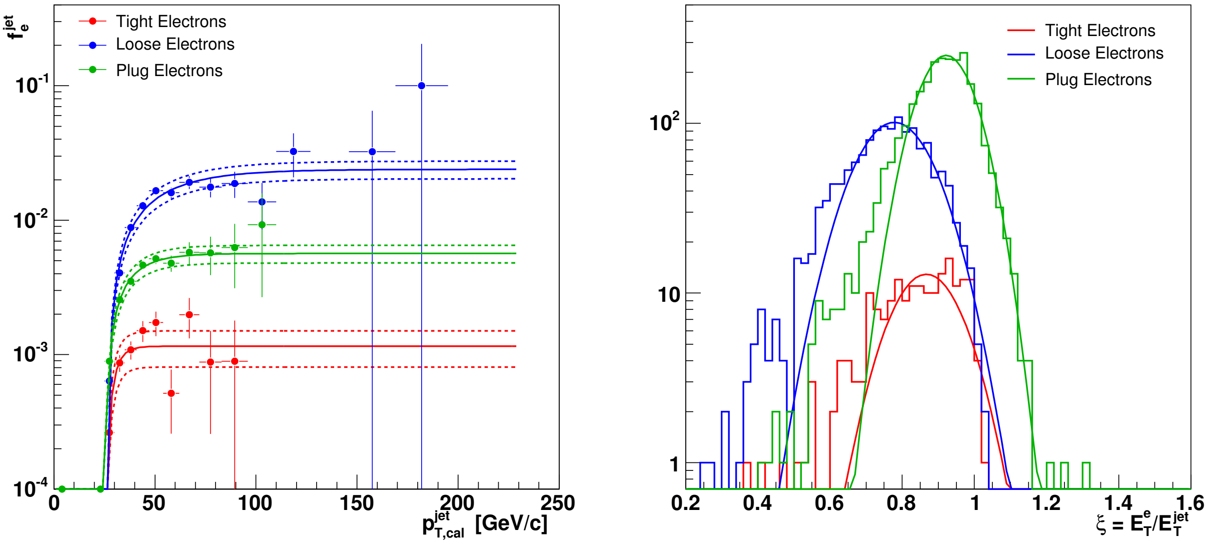
measured as a function of $p_{\text{T,cal}}^{\text{jet}}$. The jets in numerator and denominator above are different for each type of electron selection criteria:

- the numerator is defined as the number of jets reconstructed as electromagnetic objects that pass the tight, loose or plug electron cuts for the tight, loose and plug electron fake rates, respectively.
- the denominator is defined depending on the selection considered. For the tight and loose fake rates (CC selection), the denominator is defined as the number of central jets

($|y| < 1.0$) with at least one track with $p_{\mathrm{T}} > 10$ GeV/c. For the plug fake rate (CF selection), it is the number of jets in the plug region ($1.2 < |y| < 2.8$).

To measure the fake rates a "pure" jet sample is needed. We use an inclusive jet sample selected with a trigger threshold of $E_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV. Events with more than one electron or $\not{E}_{\mathrm{T}} > 15$ GeV are excluded to reject events containing Z/γ^* or W boson candidates. The fake rates are then measured as defined in Eq. 4.17. Figure 4.9(a) shows f_e^{jet} for the three types of electron selection criteria parametrized with an exponential. As expected, given the more restrictive cuts applied on their identification, tight electrons have the smallest fake probability.

Due to the differences in the reconstruction algorithms for jets and electrons, a jet with a given $E_{\mathrm{T}}^{\mathrm{jet}}$ would fake an electron with $E_{\mathrm{T}}^{e} < E_{\mathrm{T}}^{\mathrm{jet}}$. The energy scale factor $\xi_e^{\mathrm{jet}} = E_{\mathrm{T}}^{e}/E_{\mathrm{T}}^{\mathrm{jet}}$ for the different types of electrons are shown in Fig. 4.9(b). The histograms are fitted with a Gaussian distribution. The center of the Gaussian fits are summarized in Table IX.



(a) (b)

Figure 4.9: (a) Fake rates f_e^{jet} for the different types of electron selection cuts. The dashed lines illustrate the associated uncertainties: 30% for the tight fake rate, and 15% for the loose and plug rates. (b) Scale factor ξ between $E_{\mathrm{T}}^{\mathrm{jet}}$ and the E_{T} of the electron that it fakes, fitted with a Gaussian distribution.

	ξ energy scale factor
Tight electron	0.8266
Loose electron	0.7576
Plug electron	0.9203

Table IX: Energy scale factors for electrons faked by jets, $\xi_e^{\rm jet} = E_{\rm T}^{\rm e}/E_{\rm T}^{\rm jet}$.

4.5.1.2 QCD and $W(\rightarrow e+\nu)+$jets Background Estimation

Once the fake rates and scale factors are obtained, the high-$p_{\rm T}$ electron sample is used to estimate the background contribution from QCD and $W(\rightarrow e+\nu)+$jets events. Real Z/γ^* events are rejected from the high-$p_{\rm T}$ electron sample selecting events with one and only one tight electron and no other loose or plug electrons$^{(16)}$. For each jet in every event, $E_{\rm T}^{\rm e}$ of the fake electron is determined as $E_{\rm T}^{\rm e} = \xi_e^{\rm jet} \times E_{\rm T}^{\rm jet}$. If the jet fulfills the kinematic cuts applied to the electrons ($E_{\rm T}$ and η requirements) and the invariant mass of the tight electron and the jet is within $66 < M_{e-{\rm jet}} < 116$ GeV/c^2, the electron-jet combination is considered in the background calculation, where it is weighted by the corresponding $f_e^{\rm jet}$ divided by the number of accepted electron-jet combinations in the event $N_{e-{\rm jet}}$:

$$w^{\rm jet}(p_{\rm T,cal}^{\rm jet}) = \frac{f_e^{\rm jet}(p_{\rm T,cal}^{\rm jet})}{N_{e-{\rm jet}}} \tag{4.18}$$

where $f_e^{\rm jet}$ is the loose fake rate for the case of central jets and the plug fake rate for the case of plug jets. Notice that the tight fake rate is implicitly included in the selection when asking for one tight electron in the event.

For each electron-jet combination in the range $66 - 116$ GeV/c^2, the remaining jets in the event are used a background contributions to the jet distributions weighted by $w^{\rm jet}$. The estimated number of QCD and $W(\rightarrow e+\nu)+$jets events in 1.73 fb^{-1} is listed in Table X. For $N_{\rm jets} \geq 1$, QCD and $W(\rightarrow e+\nu)+$jets are 7% of the whole sample and 8% in the case of $N_{\rm jets} \geq 2$.

4.5.2 Non-QCD Backgrounds

Background contributions from other processes (referred as non-QCD backgrounds) are estimated using Pythia-Tune A Monte Carlo described in Section 2.8. The additional processes considered as background contributers are:

---

[16] Notice that this is the same sample as the $Z/\gamma^*(\rightarrow e^+e^-)+$jets signal, but requiring only one tight electron and no other electrons in the event, there is no overlapping with the signal events, and therefore no bias.

- $Z/\gamma^*(\rightarrow \tau^+\tau^-) + \text{jets}$
- diboson production: WW, ZZ and WZ
- $t\bar{t}$ production jets.
- $Z/\gamma^*(\rightarrow e^+e^-) + \gamma$, where the prompt γ can be misidentified as a jet.

The samples are normalized to the Pythia-Tune A cross section. Pythia-Tune A has been found to describe approximately the total cross section measured in the data (see Fig. 4.11 and the discussion in Section 4.6).

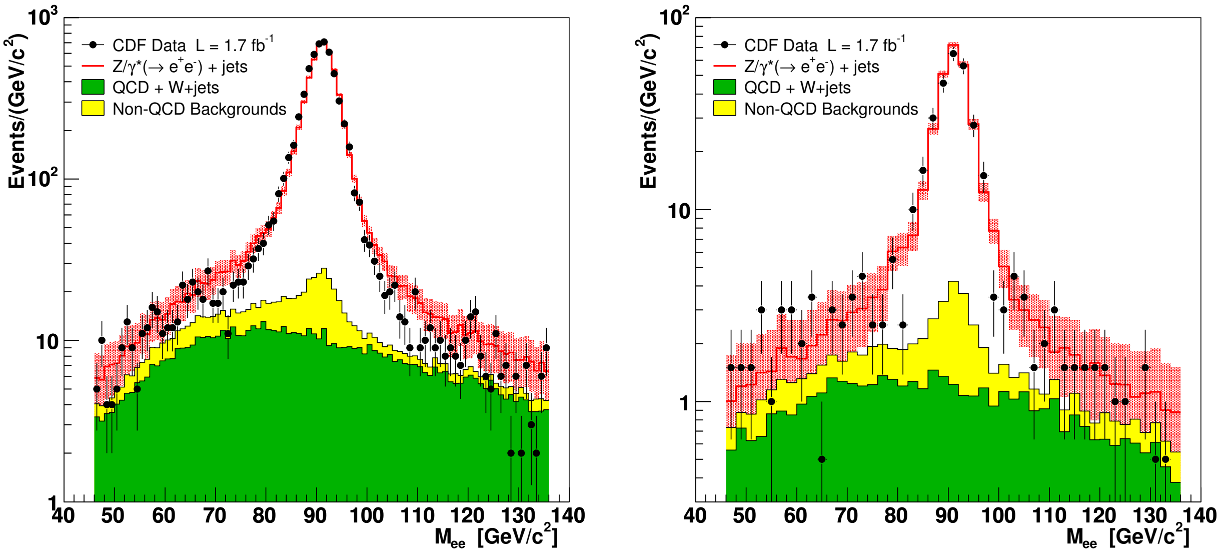
The number of expected events in 1.73 fb^{-1} of data for every background are reported in Table X. The non-QCD backgrounds represent 4% of the sample in the case of $N_{\text{jets}} \geq 1$ and 6% in $N_{\text{jets}} \geq 2$.

	Estimated events	
	$Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet	$Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets
QCD and $W(\rightarrow e+\nu) + \text{jets}$	454 ± 27	56.3 ± 9.5
$Z/\gamma^*(\rightarrow e^+e^-) + \gamma$	181 ± 13	16.0 ± 3.9
Dibosons (WW, ZZ, ZW)	39.2 ± 6.1	13.7 ± 3.6
$t\bar{t}$ production	18.6 ± 4.2	11.1 ± 3.3
$Z/\gamma^*(\rightarrow \tau^+\tau^-) + \text{jets}$	5.8 ± 2.3	0 ± 0
$Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$	5532 ± 69	578 ± 10
Total Expected	6231 ± 75	676 ± 15
Data	6203 ± 79	650 ± 25

Table X: Number of events found in 1.73 fb^{-1} of data compared to the expected background and $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ events.

4.5.3 Validation of the Background Estimation

The measurement of the inclusive $Z/\gamma^* \rightarrow e^+e^-$ cross section (Section 4.3.5) already showed a good background description. We further check the accuracy of the background estimation method in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ considering lower and higher mass regions (*side bands*) around the Z mass window ($46-66$ GeV/c^2 and $116-136$ GeV/c^2) where measurements are more sensitive to the estimation of the background. In the side bands, background events are about 65% of the total, while in the Z mass window they are only about 12%. Figure 4.10 shows data compared to the estimated background and signal in $Z/\gamma^*(\rightarrow e^+e^-)+\geq N_{\text{jets}}$ for $N_{\text{jets}} \geq 1$ and $N_{\text{jets}} \geq 2$. The number of events observed in each region are listed in Table XI. Good agreement is found in the Z mass region and both lower and upper side bands.



(a) (b)

Figure 4.10: Data and signal plus background estimation in the Z mass window and the side bands, in (a) $Z/\gamma^*(\to e^+e^-)+\geq 1$ jet and (b) $Z/\gamma^*(\to e^+e^-)+\geq 2$ jets events.

$Z/\gamma^*(\to e^+e^-)+\geq 1$ jet	**Lower Band**	**Z Window**	**Higher Band**
Data	229 ± 15	6178 ± 79	143 ± 12
Signal + Background	237 ± 15	6179 ± 57	182 ± 12
$Z/\gamma^*(\to e^+e^-)+\geq 2$ jets	**Lower Band**	**Z Window**	**Higher Band**
Data	41 ± 6	631 ± 25	18 ± 4
Signal + Background	31.5 ± 5.8	654 ± 18	23.9 ± 3.9

Table XI: Number of events in data and estimated signal plus background, in the Z mass window ($66 - 116$ GeV/c^2), the lower band of the spectrum ($46 - 66$ GeV/c^2) and the higher band ($116 - 136$ GeV/c^2).

4.6 Unfolding

Raw inclusive jet differential cross sections as a function of a given variable α are defined as:

$$\frac{d\sigma}{d\alpha} = \frac{1}{\mathcal{L}} \times \frac{N^{\rm corr}_{\rm jets}}{\Delta\alpha} \quad , \qquad \alpha \equiv p_{\rm T}^{\rm jet}, y^{\rm jet} \tag{4.19}$$

where $N^{\rm corr}_{\rm jets}$ denotes the total number of jets in the CC plus the CF configurations after the average $p_{\rm T}^{\rm jet}$ correction (Section 3.3.2) in a given bin in α, $\Delta\alpha$ is the size of the bin, and $\mathcal{L}$ is the luminosity. $N^{\rm corr}_{\rm jets}$ has been corrected bin-by-bin for background contributions and trigger inefficiencies:

$$N^{\rm corr}_{\rm jets} = \frac{N_{\rm jets} - N_{\rm QCD\ bkg}}{\kappa^{\rm MC}_{\rm Data} \cdot \epsilon_{\rm trig}} - N_{\rm non-QCD\ bkg} \tag{4.20}$$

where $N_{\rm jets}$ is the number of jets measured in the data in a given bin, $N_{\rm QCD\ bkg}$ is the number of expected QCD and $W(\rightarrow e + \nu) +$ jets background events, $N_{\rm non-QCD\ bkg}$ is the number of non-QCD expected background events, $\epsilon_{\rm trig}$ corrects the values measured in the data for the trigger efficiencies, and $\kappa^{\rm MC}_{\rm Data}$ accounts for the remaining differences between data and Monte Carlo (Section 4.3.4). All the quantities take into account the proportion of CC and CF events in every bin.

The measured raw distributions are then unfolded back to the hadron level. A bin-by-bin unfolding procedure corrects for acceptance and smearing effects, accounting for the efficiency of the selection criteria and the electron and jet reconstruction in the calorimeter. We use the Pythia-Tune A Monte Carlo sample to unfold the measurements. Chapter 5 is devoted to measure observables that guarantee that the Monte Carlo properly simulates those effects. The unfolding factors are determined as follows:

$$\mathcal{U}(\alpha) = \left. \frac{\frac{d\sigma}{d\alpha}\left(Z/\gamma^* \,[\text{hadron level}] + \text{particle jets}\left[p_{\rm T}^{\rm jet}, y^{\rm jet}\right]\right)}{\frac{d\sigma}{d\alpha}\left(Z/\gamma^* \,[\text{reconstructed}] + \text{calorimeter jets}\left[p_{\rm T,corr}^{\rm jet}, y^{\rm jet}\right]\right)} \right|_{\rm MC} \tag{4.21}$$

where the numerator is the differential cross section in α ($p_{\rm T}^{\rm jet}$ or $y^{\rm jet}$) at the hadron level and the denominator is the corresponding raw cross section at detector level. The same kinematic cuts in electrons and jets are applied at hadron level and detector level, thus the cross sections measurements refer to hadron level jets with $p_{\rm T}^{\rm jet} > 30$ GeV and $|y^{\rm jet}| < 2.1$ in the kinematic range of Z/γ^* decay products: $E_{\rm T}^{e} > 25$ GeV, $|\eta^{e_1}| < 1.0$, $|\eta^{e_2}| < 1.0$ or $1.2 < |\eta_{e_2}| < 2.8$, $66 < M_{ee} < 116$ GeV/c^2 and $\Delta R_{\rm e-jet} > 0.7$. To obtain the unfolded cross section, the unfolding factors are applied to the measured raw cross section in data:

$$\left.\frac{d\sigma}{d\alpha}\right|_{\rm measured} = \mathcal{U}(\alpha) \times \left.\frac{d\sigma}{d\alpha}\right|^{\rm Data}_{\rm Raw} \tag{4.22}$$

In order to avoid any bias in the unfolding procedure due to the particular PDFs used in the generation of the Monte Carlo, Pythia samples are reweighted to follow the $p_{\rm T,corr}^{\rm jet}$

distribution measured in the data. Figure 4.11(a) shows the differential cross section as a function of $p_{\mathrm{T,corr}}^{\mathrm{jet}}$ at the detector level for data and PYTHIA-TUNE A Monte Carlo. For the purpose of the unfolding, only the shape of the distribution is taken into account, but in the case of PYTHIA-TUNE A, we observe that also the overall normalization coincides approximately with the data. This is a consequence of the fact that the TUNE A settings were conscientiously tuned to describe the data during the Run I of the Tevatron (Section 2.8.1.1) However, the shape of the cross section in PYTHIA is too "soft", leading to a lower cross section at high $p_{\mathrm{T,corr}}^{\mathrm{jet}}$. It is known that CTEQ5L PDFs produce a softer $p_{\mathrm{T}}^{\mathrm{jet}}$ spectrum than the data. The weights are determined by fitting the ratio of cross sections Data/MC shown in Fig. 4.11(b) with a second order polynomial. Then, PYTHIA generated events are reweighted in $\hat{p}_{\mathrm{T}}$. The reweighting is done separately for the CC and CF configurations, and for the different jet multiplicities. Figure 4.11(c,d) shows the raw cross section for CC selection in data and PYTHIA and their ratio after the reweighting in $\hat{p}_{\mathrm{T}}$. Another bias introduced by the use of a given PDF is observed in the E_{T} spectrum of the forward electrons $E_{\mathrm{T}}^{e,\mathrm{fwd}}$. Figure 4.12(a) shows the ratio Data/MC of the E_{T} of the forward electrons. This ratio is fitted to a straight line and a weight is applied to the E_{T} of the generated forward electrons. The ratio Data/MC of $E_{\mathrm{T}}^{e,\mathrm{fwd}}$ at the detector level after the reweighting is shown in Fig. 4.12(b). The effect of this reweighting in other distributions, other than $E_{\mathrm{T}}^{e,\mathrm{fwd}}$, has been checked to be negligible ($< 1\%$). Unless otherwise explicitly said, all distributions of PYTHIA Monte Carlo used in this analysis have been reweighted in $\hat{p}_{\mathrm{T}}$ and $E_{\mathrm{T}}^{e,\mathrm{fwd}}$. The general formula of the reweighting process can be written as:

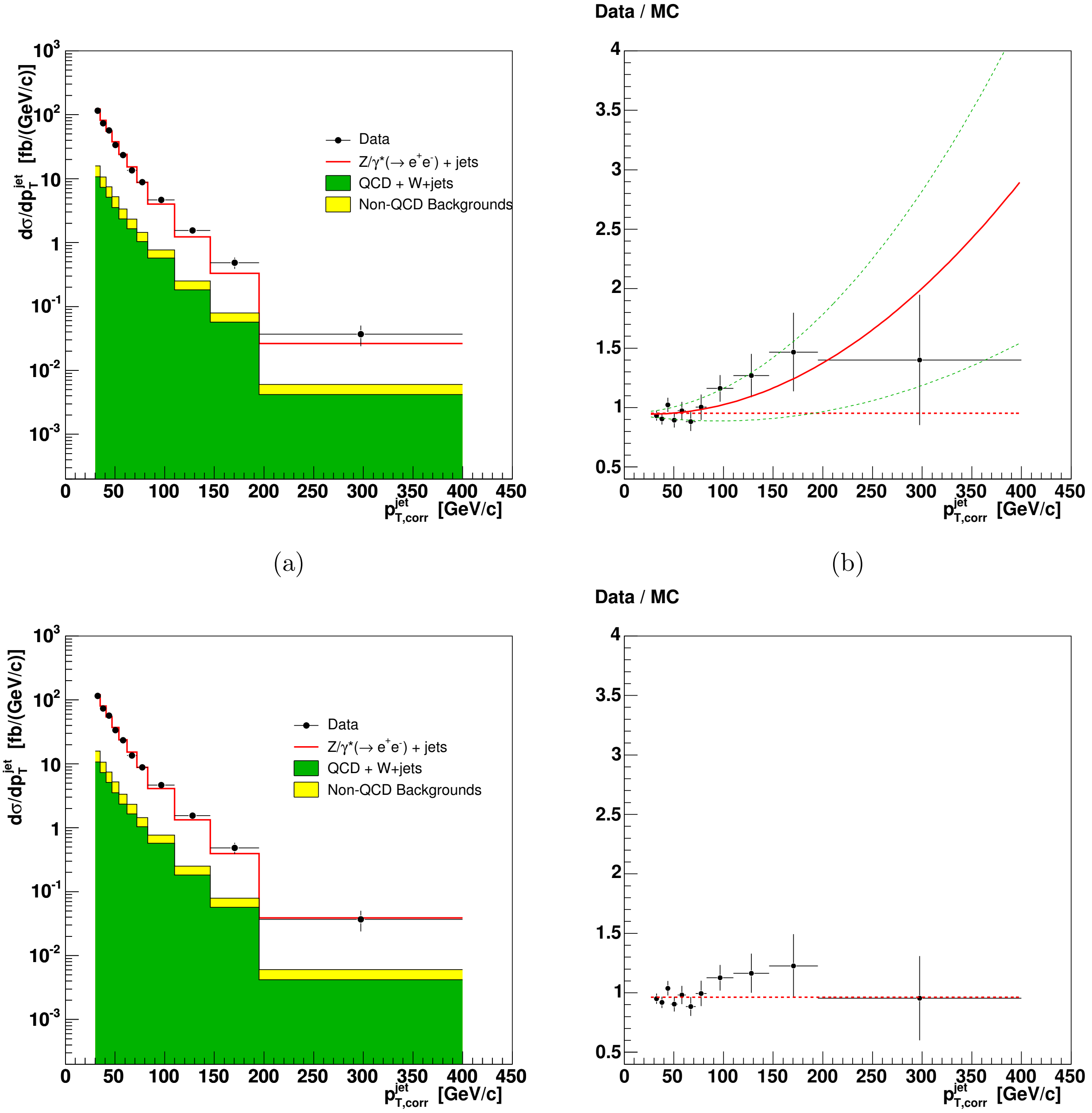
$$w_{\mathrm{PYTHIA}} = \left(a_0 + a_1 \cdot \hat{p}_{\mathrm{T}} + a_2 \cdot \hat{p}_{\mathrm{T}}{}^2\right) \times \left(b_3 + b_4 \cdot E_{\mathrm{T}}^{e,\mathrm{fwd}}\right) \tag{4.23}$$

The values of the p_k factors for the different jet multiplicities and event selections are listed in Table XII.

	a_0	a_1	a_2	b_3	b_4
$Z/\gamma^*(\to e^+e^-)+\geq 1$ jet (CC)	0.9818	-0.000610	$1.373{\cdot}10^{-5}$	1.0	0.0
$Z/\gamma^*(\to e^+e^-)+\geq 2$ jets (CC)	0.9492	-0.002807	$4.867{\cdot}10^{-5}$	1.0	0.0
$Z/\gamma^*(\to e^+e^-)+\geq 3$ jets (CC)	1.0	0.0	0.0	1.0	0.0
$Z/\gamma^*(\to e^+e^-)+\geq 1$ jet (CF)	1.0	0.0	0.0	0.6480	$9.615{\cdot}10^{-3}$
$Z/\gamma^*(\to e^+e^-)+\geq 2$ jets (CF)	1.0	0.0	0.0	0.6480	$9.615{\cdot}10^{-3}$
$Z/\gamma^*(\to e^+e^-)+\geq 3$ jets (CF)	1.0	0.0	0.0	1.0	0.0

Table XII: Weights applied to PYTHIA Monte Carlo samples. The listed factors correspond to Eq. 4.23.

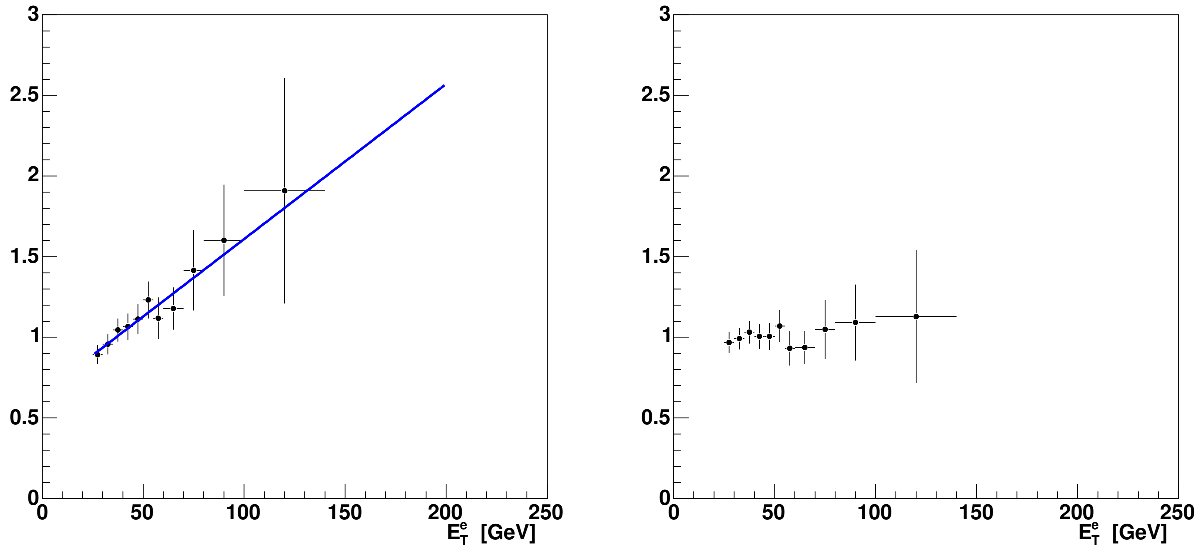
Figure 4.13 shows the unfolding factors for $p_{\mathrm{T}}^{\mathrm{jet}}$ and y^{jet} distributions in $Z/\gamma^* \to e^+e^-$ events with ≥ 1 jet and ≥ 2 jets. The unfolding factors for both distributions are rather



(a) (b)

(c) (d)

Figure 4.11: (a) Detector level cross section for the CC selection in data and Pythia-Tune A Monte Carlo and (b) their ratio Data/MC before reweighting in $\hat{p}_T$. The curve represents the fit to a second order polynomial from which the weights are determined. The dashed lines correspond to the errors in the fit. (c,d) The same cross sections after the reweighting in $\hat{p}_T$ of Pythia Monte Carlo events.



(a)

(b)

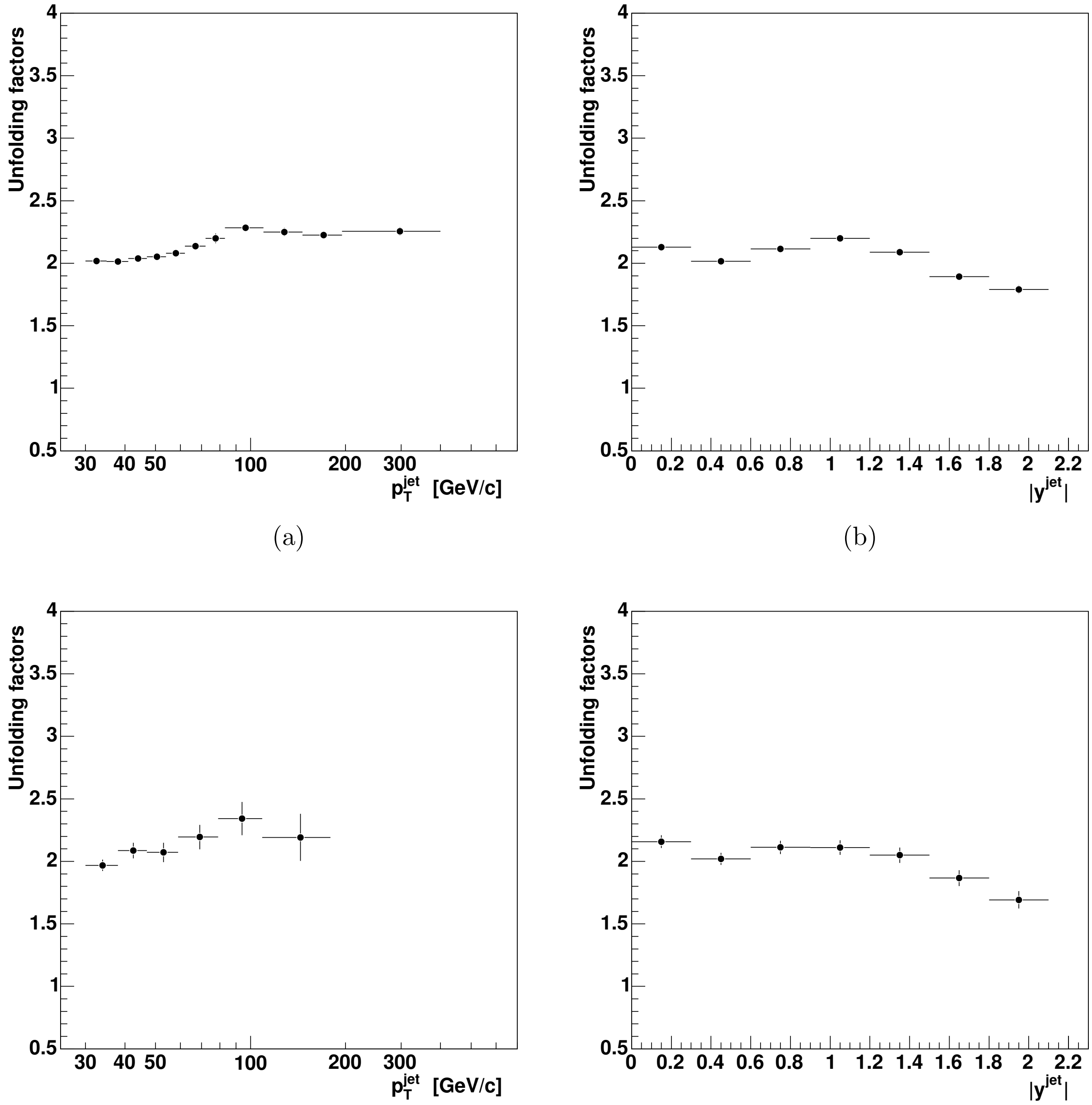
Figure 4.12: Ratio of the forward electron E_{T} Data/MC(Pythia) (a) before and (b) after the reweighting in $E_{\mathrm{T}}^{e,\mathrm{fwd}}$ of the Monte Carlo sample.

flat and have an average value around 2. Although it is not possible to completely separate the effects of the Z/γ^* boson selection and the jet reconstruction, by breaking the unfolding process in several steps it can be estimated that the major contribution to the unfolding factors comes from the selection of the Z/γ^* boson through the identification of the electrons. Appendix B shows the unfolding breakdown studies in detail.

4.7 Systematic Uncertainties

The following sources of systematic uncertainties on the measured cross sections have been considered:

- **Trigger efficiency**: the main uncertainty on the trigger efficiency comes from the dependence of the tracking efficiency at Level 1 on the η of the electron (Section 4.1.1). The associated uncertainty is 1.5%, This uncertainty is propagated in Eq. 4.3 and Eq. 4.4 in Section 4.3.1 to obtain the uncertainty per event for the CC and CF configurations, respectively. The global trigger uncertainty is 0.06% in CC events and 1.5% in CF events (Table V). The final uncertainty in the cross sections is computed according to the proportion of CC and CF events in every bin.



(a)

(b)

(c)

(d)

Figure 4.13: Unfolding factors of the (a) $p_{\mathrm{T}}^{\mathrm{jet}}$ and (b) y^{jet} distributions in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+\geq 1$ jet events. Unfolding factors of the (c) $p_{\mathrm{T}}^{\mathrm{jet}}$ and (d) y^{jet} distributions in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+\geq 2$ jets events.

- **Electron ID efficiencies**: since cross sections are measured in an environment with jets, we checked the dependency of the efficiency versus $p_{\mathrm{T,corr}}^{\mathrm{jet}}$ (Section 4.3.4). The limited statistics of the study reduced any possible dependence to 5% at the most (Fig. 4.3). This upper limit is taken as a systematic uncertainty. An additional $\sim 1\%$ (column *Sys2* in Table VIII) uncertainty accounts for remaining differences between data and Monte Carlo simulation in some electron quantities, as taken from CDF measurements on inclusive Z/γ^* samples [55]. The total systematic uncertainty for the CC configuration is 5.02%, and 5.17% for the CF configuration.

- **Jet energy corrections**: the determination of the jet energy scale (Section 3.3.2.1) dominates the uncertainty on the final cross section. Those uncertainties are specially important at high $p_{\mathrm{T}}^{\mathrm{jet}}$.

  - **Relative jet energy**: the procedure to obtain the C_y corrections is repeated in both data and Monte Carlo varying the fitting techniques. Deviations of β_{dijet} from the unity are taken as uncertainties. They represent an uncertainty on the cross section around 2%-3% and are relatively flat in $p_{\mathrm{T}}^{\mathrm{jet}}$.

  - **Multiple $p\overline{p}$ interactions**: the main uncertainty on the C_{MI} correction comes from the efficiency on determining the number of vertices in the event. This efficiency is measured in different data samples: $W \rightarrow e\nu$, inclusive jets and minimum bias sample. A 15% uncentainty on the pile-up contribution translates into a 1% uncertainty on the measured cross section in the whole range in $p_{\mathrm{T}}^{\mathrm{jet}}$.

  - **Absolute jet energy**: the uncertainty on the absolute correction C_{Abs} varies between 2% at $p_{\mathrm{T}}^{\mathrm{jet}} \sim 30$ GeV/c and 2.7% at $p_{\mathrm{T}}^{\mathrm{jet}} \sim 400$ GeV/c, and is obtained from the difference between data and simulation in the calorimeter response to single particles and the p_{T} spectrum of the particles inside jets [53]. It is the dominant contribution to the cross section uncertainty (3%-12%).

- **QCD and $W(\rightarrow e+\nu) + $jets background estimation**: the main uncertainty in estimating the QCD and $W(\rightarrow e+\nu) + $jets contribution to the $Z/\gamma^*(\rightarrow e^+e^-) + $jets background comes from the determination of the fake rates f_e^{jet}. A 15% uncertainty is taken to cover any possible deviation of the nominal f_e^{jet}. The uncertainty bands in f_e^{jet} are illustrated in Fig. 4.9. The $\not{E}_{\mathrm{T}}$ cut in the event selection of the inclusive jet sample was varied ±30% and its effect on the determination of f_e^{jet} is covered by the 15% uncertainty already assigned. This uncertainty translates directly into a 15% uncertainty in the background estimation. Since the background is small compared to the signal, this uncertainty corresponds only to a 1%-2% uncertainty in the cross section.

- **Non-QCD backgrounds**: the Monte Carlo samples used to estimated the non-QCD backgrounds contributions are normalized to the LO+PS PYTHIA-TUNE A cross section (Section 4.5.2). The estimated uncertainty in the PYTHIA-TUNE A cross sections of the different processes vary between 10% and 30%. As the contribution of the non-QCD backgrounds is very small, a conservative 30% uncertainty on the normalization is used, which translates to a 1% uncertainty in the $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ cross sections.

- **Reweighting of the Monte Carlo**: the unfolding procedure has been repeated using a non-reweighted PYTHIA-TUNE A sample (Section 4.6) to estimate the uncertainty introduced by the use of a reweighted Monte Carlo (Fig. 4.14). As expected, the differences in the unfolding factors are small ($\lesssim$ 1%) because the event weights themselves are small. Furthermore, the value of the weights applied in $\hat{p}_{\mathrm{T}}$ has been varied within the errors of the fitted function used to obtain the weights, as shown in Fig. 4.11(b). The variation in the unfolding factors are also less than 1%.

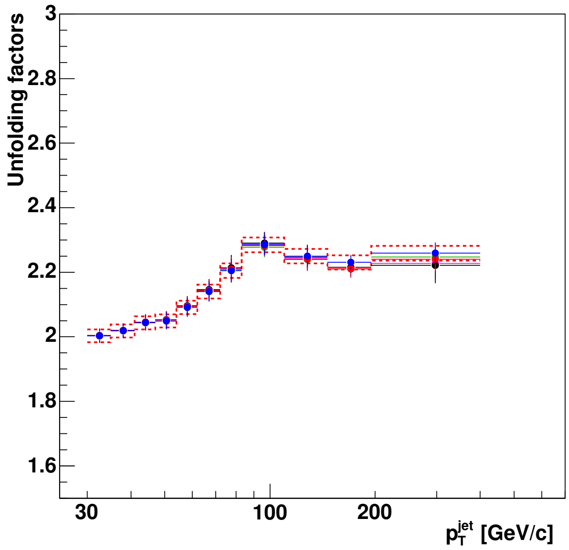


Figure 4.14: Unfolding factors for the $p_{\mathrm{T}}^{\mathrm{jet}}$ distribution in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet events. Black dots represent the unfolding factors using non-reweighted PYTHIA-TUNE A, red dots for PYTHIA-TUNE A reweighted only on $\hat{p}_{\mathrm{T}}$, green and pink dots varying up and down the weighting function, and blue dots for the unfolding factors reweighted on both $\hat{p}_{\mathrm{T}}$ and $E_{\mathrm{T}}^{e,\mathrm{fwd}}$. The dashed lines correspond to the assigned uncertainty of 1%.

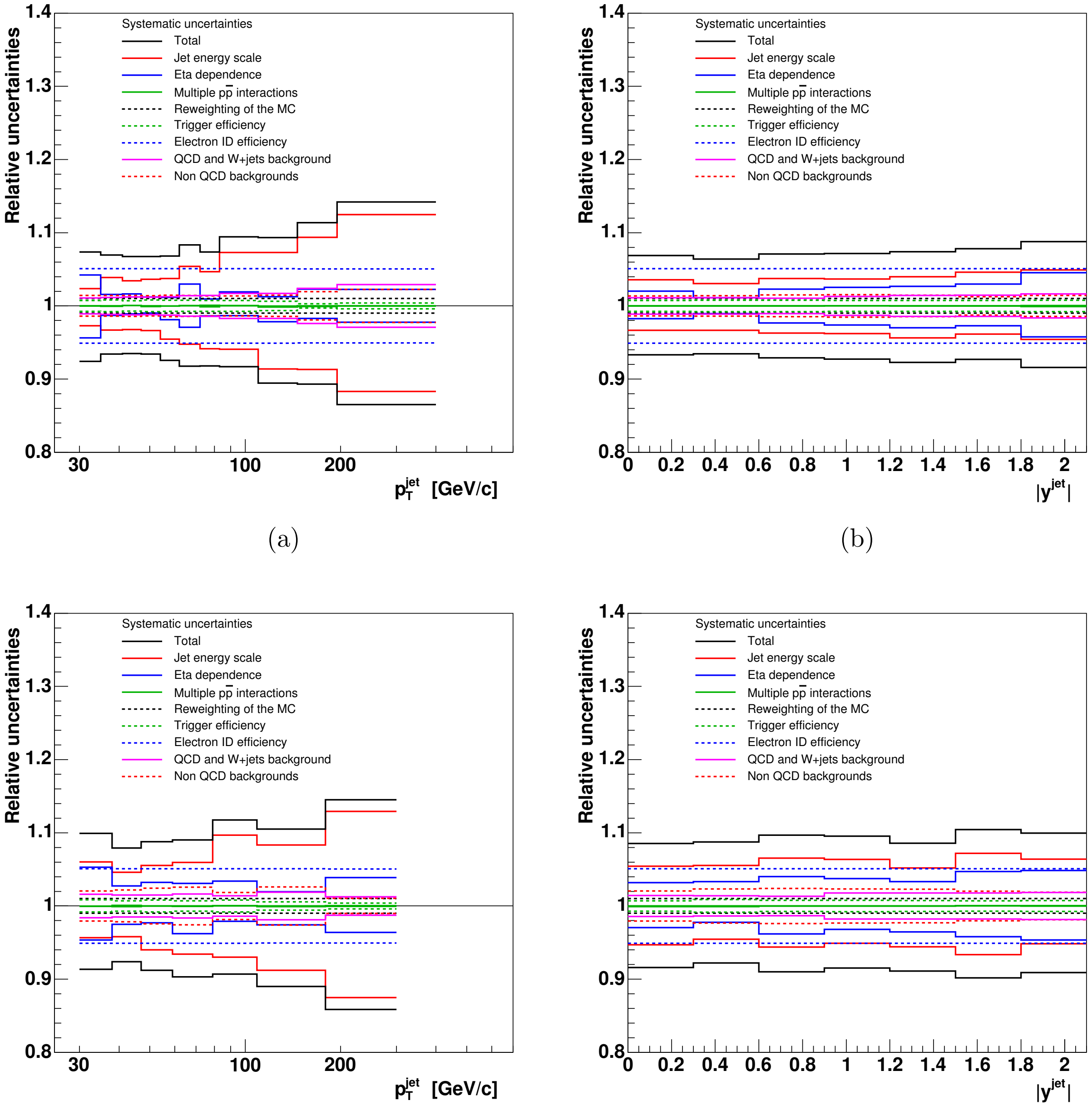
Positive and negative deviations from the nominal value for every systematic uncertainty are added separately in quadrature to determine the total positive and negative systematic uncertainties, respectively, in every bin. An additional ±5.8% uncertainty on the luminosity measurement (Section 3.2.6) must be considered but unless otherwise indicated it is not included in the total uncertainty shown the figures. The systematic uncertainties are summarized in Table XIII and plotted in Fig. 4.15.

$Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet	**CC selection**	**CF selection**
Trigger efficiency	0.06%	1.5%
Electron ID efficiencies	5.02%	5.17%
QCD and $W(\rightarrow e+\nu)+\text{jets}$ background	1%-2%	
Non QCD background	1%	
Reweighting of the Monte Carlo	1%	
Absolute jet energy scale	3%-12%	
η dependence	2%-3%	
Multiple particle interactions	$< 1\%$	
Total	7%-14%	
$Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets	**CC selection**	**CF selection**
Trigger efficiency	0.06%	1.5%
Electron ID efficiencies	5.02%	5.17%
QCD and $W(\rightarrow e+\nu)+\text{jets}$ background	1%-2%	
Non QCD background	2%	
Reweighting of the Monte Carlo	1%	
Absolute jet energy scale	5%-13%	
η dependence	4%	
Multiple particle interactions	$< 1\%$	
Total	10%-15%	

Table XIII: Systematic uncertainties on the cross section measurements.

4.7.1 Correlations between Systematic Uncertainties

A careful consideration of correlations between systematic uncertainties has been carried out. The different sources of systematics are considered independent and fully correlated across p_T^{jet} and y^{jet} bins, and such will be considered when quantifying the agreement with NLO pQCD predictions in Chapter 6.



(a) (b)

(c) (d)

Figure 4.15: Total systematic uncertainties on the (a) p_T^{jet} and (b) y^{jet} inclusive jet differential cross sections in $Z/\gamma^*(\to e^+e^-)+\geq 1$ jet production. Total systematic uncertainties on the (c) p_T^{jet} and (d) y^{jet} inclusive jet differential cross sections in $Z/\gamma^*(\to e^+e^-)+\geq 2$ jets production.

Chapter 5

Jet Shapes and Energy Flows in $Z/\gamma^*(\rightarrow e^+e^-) + \textbf{jets}$

A proper modeling of non-perturbative contributions is crucial to compare measured cross sections to perturbative predictions, where the effects of the underlying event and the fragmentation of the partons into hadrons are not included in the calculation (Section 2.7.3). The contributions from these non-perturbative effects are obtained from samples of Monte Carlo generated events. It is important to validate these samples to ensure a good description of the underlying event and jet fragmentation, as measured in the data.

This chapter describes the measurement of two observables sensitive to the non-perturbative contributions: the shape of the jets and the energy flow in the events. The two measurements are compared to Monte Carlo generated events with different models for the underlying event.

5.1 Jet Shapes

The internal structure of the jets is mainly driven by the parton cascades from the primary parton but it is also sensitive to the details of the underlying event modeling. Therefore, their measurement and comparison to Monte Carlo predictions is a good mean to test and validate the implementation of underlying event models in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ production. The jet shapes have been previously measured in inclusive jet production in CDF using 170 pb^{-1} of data [57]. Results showed that PYTHIA-TUNE A described the data in all the range of $p_{\rm T}^{\rm jet}$, while HERWIG produced too narrow jets at low $p_{\rm T}^{\rm jet}$.

Here, we present differential and integrated jet shapes measurements in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ events (Fig. 5.1) for jets with $p_{\rm T}^{\rm jet} > 30$ GeV/c and $|y^{\rm jet}| < 0.7$. The differential jet shape is defined as the fraction of total $p_{\rm T}$ contained inside an annulus of radius r and width Δr

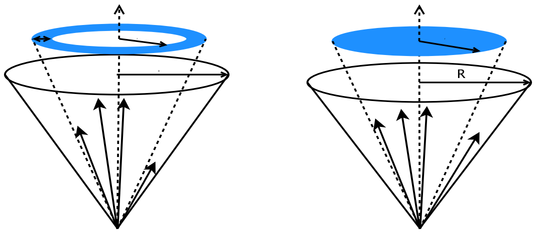
concentric to the jet cone, mathematically:

$$\rho(r) = \frac{1}{N_{\mathrm{jets}}}\frac{1}{\Delta r}\sum_{jets}\frac{p_{\mathrm{T}}(r \pm \Delta r/2)}{p_{\mathrm{T}}(0,R)} \tag{5.1}$$

where R is the radius of the cone of the jet. In our case, $\Delta r = 0.1$ and r varies from 0 to 0.7. In a similar way, the integrated jet shape measures the fraction of total p_{T} contained inside a cone of radius r concentric to the jet cone, formally:

$$\Psi(r) = \frac{1}{N_{\mathrm{jets}}}\sum_{jets}\frac{p_{\mathrm{T}}(0,r)}{p_{\mathrm{T}}(0,R)} \tag{5.2}$$

Notice that $\Psi(r = R) = 1$ by definition. In data and Monte Carlo simulation calorimeter towers are used to reconstruct the jet shapes. The energy of the towers associated to the jet cluster are summed according to their distance r in $y - \phi$ to the jet axis. For jets in the central region $|y^{\mathrm{jet}}| < 0.7$, that is covered by the COT, the jet shapes can also be measured using tracks and their distance r to the center of the jet.



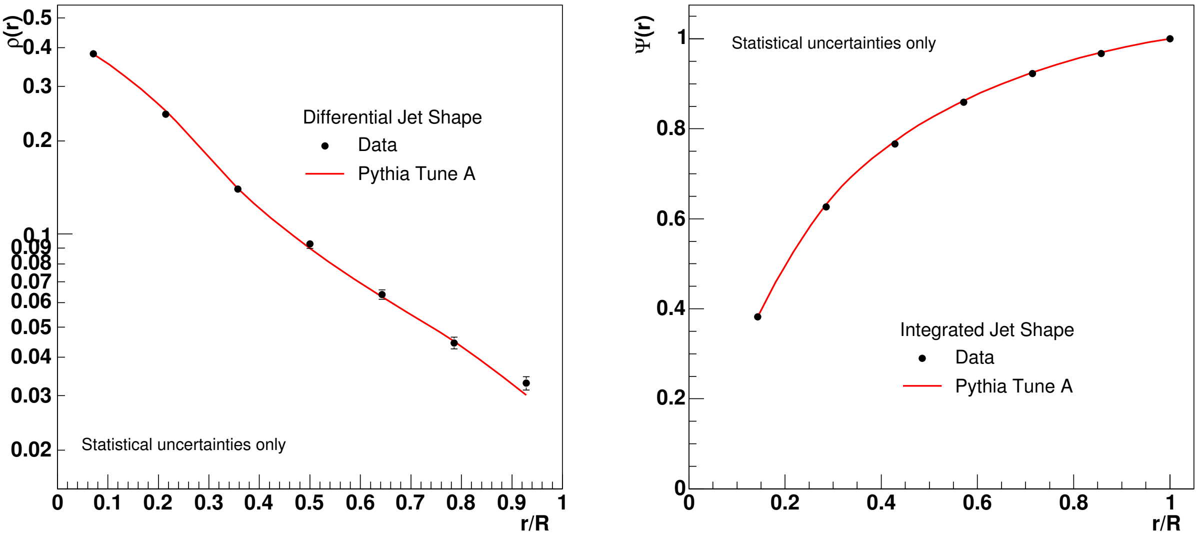
(a) (b)

Figure 5.1: Illustrations of the (a) differential and (b) integrated jet shapes.

The jet shapes are only measured in events with exactly one reconstructed vertex, to avoid the events with multiple $p\bar{p}$ interactions that would dazzle the effects from the non-perturbative contributions. The same event selection defined in Section 4.3.3 is applied in this case. The measurement is restricted to jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 30$ GeV/c and $|y^{\mathrm{jet}}| < 0.7$. The jet shapes measured with tracks are useful to cross check the calorimeter measurements and their difference will be used to estimated some of the systematic uncertainties. The jet shapes as measured in the detector are shown in Fig. 5.2 compared to the prediction from PYTHIA-TUNE A.

The measurement is unfolded bin-by-bin to the hadron level using PYTHIA-TUNE A Monte Carlo. The unfolding procedure is the same described in Section 4.6. The unfolding factors are determined as follows:

$$\mathcal{U}_{\rho}(r) = \frac{\rho_{\mathrm{HAD}}(r)}{\rho_{\mathrm{CAL}}(r)}\ ,\quad \mathcal{U}_{\Psi}(r) = \frac{\Psi_{\mathrm{HAD}}(r)}{\Psi_{\mathrm{CAL}}(r)} \tag{5.3}$$



(a) (b)

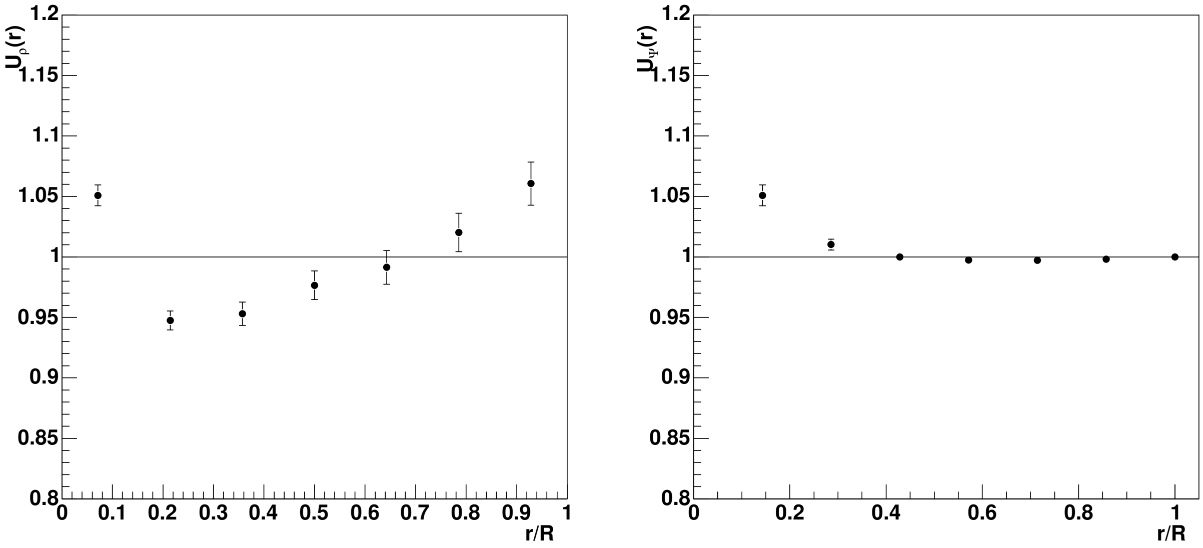
Figure 5.2: (a) Differential and (b) integrated jet shapes measured at the detector level in data and PYTHIA-TUNE A Monte Carlo for jets with $|y^{\rm jet}| < 0.7$ in $Z/\gamma^*(\to e^+e^-) +$ jets events with exactly one reconstructed vertex.

where HAD refers to the jet shape measurements at hadron level and CAL to the calorimeter measurements. At the hadron level, the shapes of the jets are reconstructed using Eqs. 5.1 and 5.2, and the list of stable particles belonging to the jet. Figure 5.3 shows the unfolding factors for the differential and integrated jet shapes.

5.1.1 Systematic Uncertainties on the Jet Shapes

The sources of systematic uncertainty on the jet shapes taken into account are:

- **Jet energy**: the energy of the jet has been varied up to ±3%, depending on $p_{\rm T}^{\rm jet}$ and $y^{\rm jet}$, to account for the uncertainty on the determination of the energy of the jet (Section 3.3.2.1). Given that the jet shapes are defined as momentum ratios, the effect on the measured jet shapes is only about 1%-2%, dominated by the uncertainty on the absolute jet energy scale determination. Other sources of systematic uncertainties related with the jet energy corrections are negligible.
- **Efficiency**: the ratios of the jet shape measurements as determined using calorimeter towers and tracks, $\zeta_\rho = \rho_{\rm CAL}/\rho_{\rm TRK}$ and $\zeta_\Psi = \Psi_{\rm CAL}/\Psi_{\rm TRK}$ in data and in Monte Carlo are compared. The double ratios $\zeta_\rho^{\rm data}/\zeta_\rho^{\rm MC}$ and $\zeta_\Psi^{\rm data}/\zeta_\Psi^{\rm MC}$ leave possible remaining



(a) (b)

Figure 5.3: Unfolding factors for the (a) differential and (b) integrated jet shapes obtained with PYTHIA-TUNE A Monte Carlo.

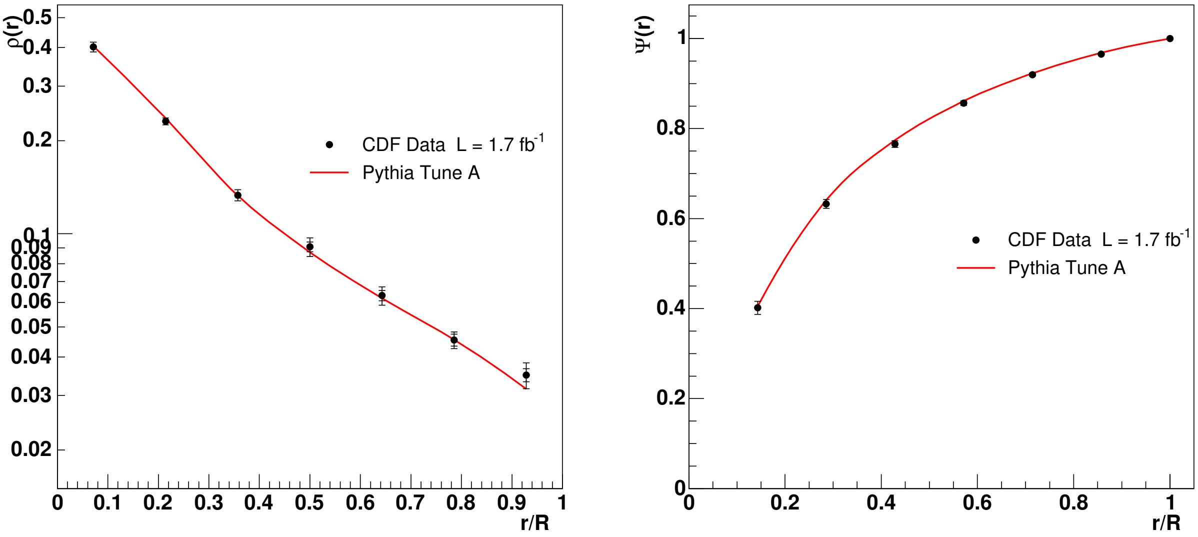
imperfections in the description of the inactive material in front of the calorimeter, and the calorimeter response to low energy particles in the Monte Carlo. The deviation from the unity of the double ratios are thus used as systematic uncertainties. The effect on the jet shapes varies from 3% in the center of the jet to 8% in the tail.

- **Reweighting of the Monte Carlo**: no substantial effect was observed due to the used of a reweighted Monte Carlo sample to unfold the jet shapes

The different uncertainties are added in quadrature and the total uncertainty on the jet shapes varies between 4% and 10%. The contributions to the total uncertainties are summarized in Table XIV. Figure 5.4 shows the final results fully corrected to the hadron level.

5.1.2 Underlying Event in Monte Carlo Generators

Soft isotropic underlying event contributions are more visible in the tail of the jets. This corresponds to $r/R \lesssim 1$ in the differential jet shape distribution, where the transverse momentum from the jet is comparable to the amount of momentum coming from underlying event activity. Therefore, higher underlying event activity leads to jets with higher tails and wider jet shape distributions. This effect can be observed in Fig. 5.5 where the jet shapes from PYTHIA are compared to the data. When the underlying event activity is switched off in



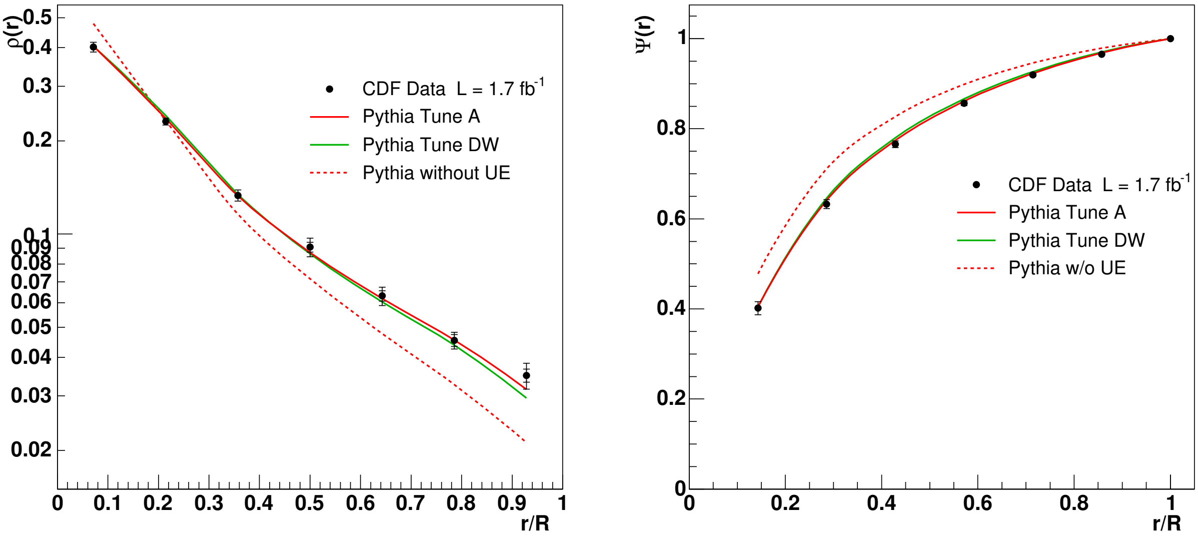
(a) (b)

Figure 5.4: (a) Differential and (b) integrated jet shapes measurement for jets with $p_T^{\text{jet}} > 30$ GeV/c and $|y^{\text{jet}}| < 0.7$ in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ events. First tick in the error bars correspond to the statistical uncertainty, the second tick mark correspond to the total (statistical $\oplus$ systematic) uncertainty. Results are compared to the predictions from Pythia-Tune A.

Systematic Uncertainties	
Absolute jet energy scale	1%-2%
η dependence	0.5%
Multiple particle interactions	$< 0.01\%$
Reweighting of the Monte Carlo	$< 0.01\%$
Efficiency	3%-8%
Total	4%-10%

Table XIV: Systematic uncertainties on the jet shapes measurement.

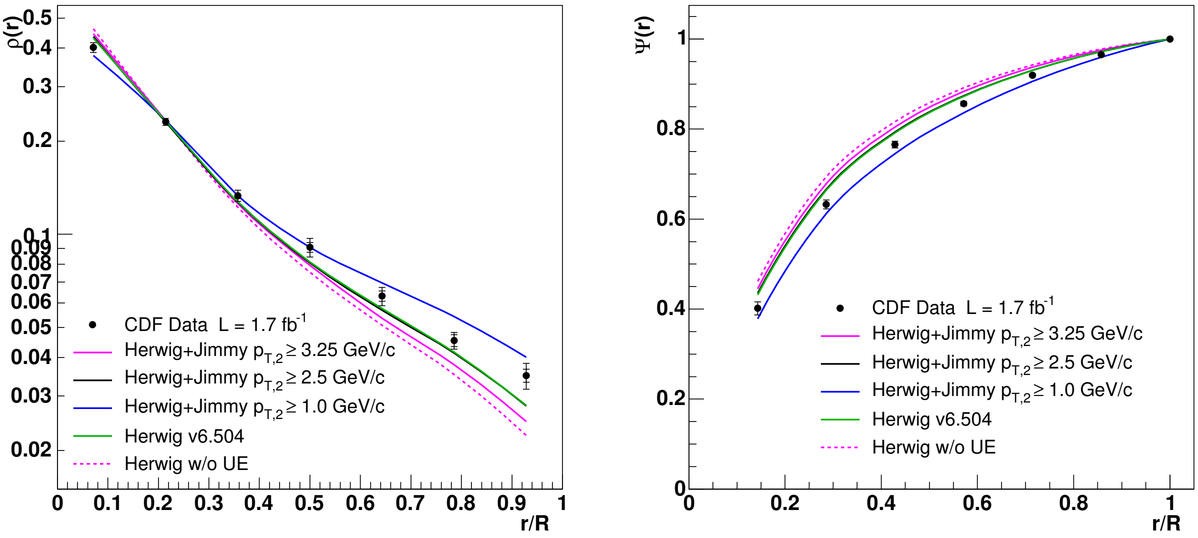
Pythia, the jets get significantly narrower. If the underlying event model is governed by the Tune A or Tune DW settings (described in detail in Section 2.8.1) the jet shapes are in very good agreement with the data. Tune A settings enhance the initial-state gluon radiation and the soft interactions between the remnants of the proton and the antiproton compared to that in the default Pythia. Tune A was determined using Run I data. Tune DW is similar to Tune A, but changes the $k_\perp$ distribution of the partons inside the hadrons, which produces minor changes in the jet shapes (Fig. 5.5) it introduces a significant change in the rate of production of jets at the hadron level.



(a) (b)

Figure 5.5: (a) Differential and (b) integrated jet shapes for jets with $p_T^{\text{jet}} > 30$ GeV/c and $|y^{\text{jet}}| < 0.7$ in $Z/\gamma^*(\to e^+e^-)+$jets events compared to the predictions from Pythia.

Previous Run II results in dijet events [57] showed that the default underlying event model in Herwig 6.4 failed to reproduce the jet shapes in the data and produced too narrow jets. Figure 5.6 shows the jet shapes in Herwig 6.504 using the default underlying event settings, where the jets are narrower than in the data. Nevertheless, Herwig also allows the use of external generators like Jimmy for multiple parton scattering. The amount of soft activity generated by Jimmy is mainly determined by a p_T threshold of the secondary *parton − parton* interactions $p_{T,2}$ (Section 2.8.2.1). Lowering the threshold increases the amount of underlying event generated and produces wider jets. This is the correlation observed in Fig. 5.6 between $p_{T,2}$ and the jet shapes in Herwig 6.510 using Jimmy 1.4. More studies are being carried out to determine a set of Jimmy parameters able to describe the data in the whole kinematic range.



(a) (b)

Figure 5.6: (a) Differential and (b) integrated jet shapes for jets with $p_T^{\text{jet}} > 30$ GeV/c and $|y^{\text{jet}}| < 0.7$ in $Z/\gamma^*(\to e^+e^-) + $jets events compared to the predictions from HERWIG and HERWIG interfaced with JIMMY.

5.2 Energy Flow in $Z/\gamma^*(\to e^+e^-) + $jets Events

The energy flow in the event is measured in the following way: event-by-event, the origin $\phi = 0$ is set in the direction of the Z/γ^* boson, and the energy profile in the transverse plane is measured versus $|\phi|$ (Fig. 5.7):

$$\left\langle \frac{dp_T}{d|\phi|} \right\rangle = \frac{1}{N_{\text{events}}} \sum \frac{dp_T}{d|\phi|} \tag{5.4}$$

Only towers in the central calorimeter ($|y^{\text{tower}}| < 0.7$) are used. Towers from the energy clusters of the electrons used to reconstruct the Z/γ^* boson are not included in the profile. As in the jet shape study, only events with one reconstructed vertex are selected to avoid the influence of multiple $p\bar{p}$ interactions.

Figure 5.8 shows the energy flow distribution. We distinguish two main regions: at $|\phi| \sim \pi$, back-to-back with the Z/γ^* boson, there are the calorimeter towers from the leading jet. Between the boson and the jet, $0 < |\phi| \lesssim 1.5$, the main contribution to the energy of

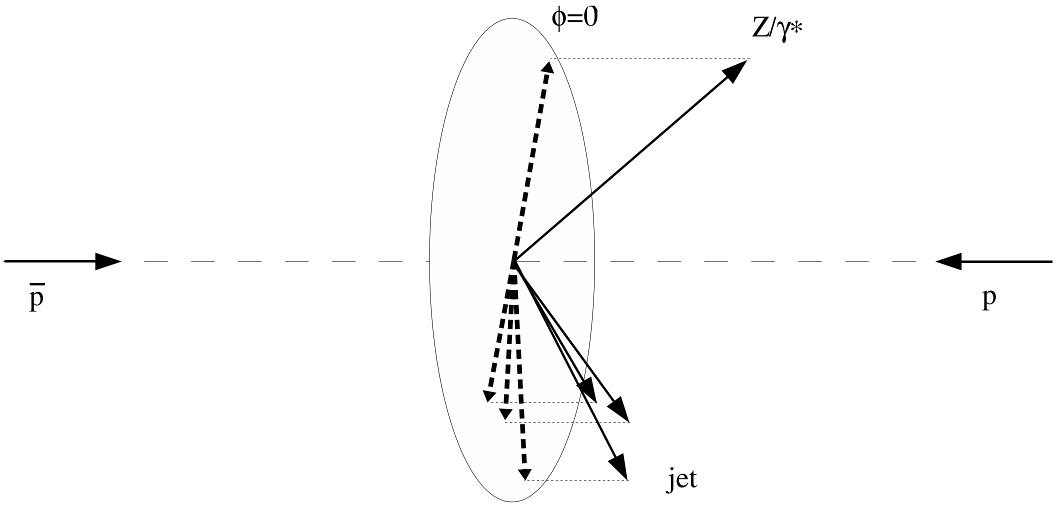
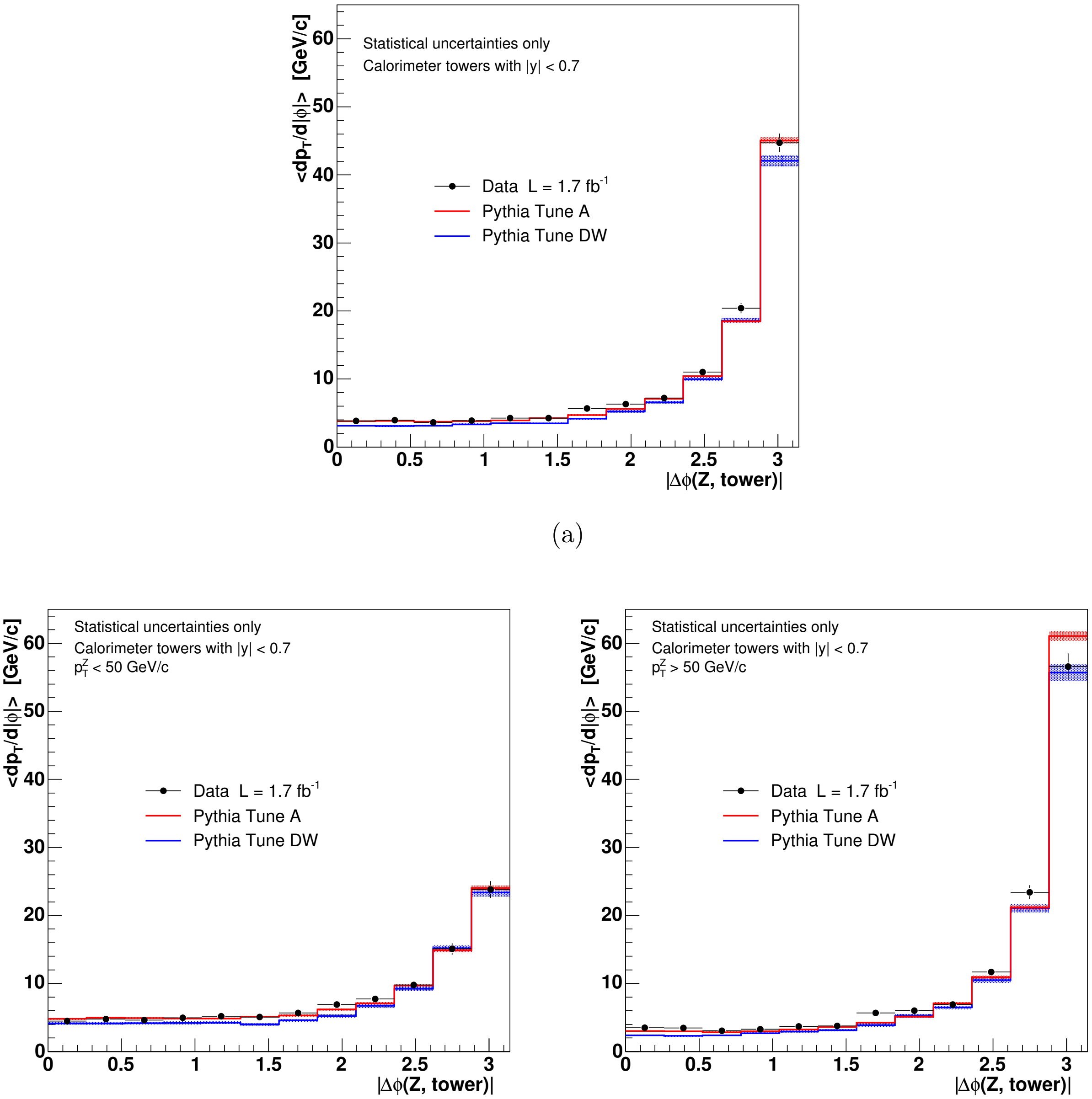


Figure 5.7: Diagram of the energy flow definition in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$.

the calorimeter towers comes from the underlying event and, to a lesser extent, from the presence of additional jets in the event. The measurement is repeated in two regions of p_T^Z: $p_T^Z < 50$ GeV/c and $p_T^Z > 50$ GeV/c. It is clear that while the energy flow in the "jet" region increases due to the higher p_T^{jet} needed to balance p_T^Z, the region of interest dominated by the underlying event keeps the same level around $2 - 4$ GeV/c. That indicates some level of decoupling between the jet from the hard interaction and the underlying event. Figure 5.8 also shows that PYTHIA-TUNE A and TUNE DW accurately describe the energy flow measured in the data.



(a)

(b)

(c)

Figure 5.8: (a) Energy flow of the central calorimeter towers in $Z/\gamma^*(\to e^+e^-) + $jets events compared to the predictions from Pythia-Tune A and Tune DW. (b) Energy flow in events with $p_T^Z < 50$ GeV/c and (c) events with $p_T^Z > 50$ GeV/c. Only statistical uncertainties are included in the plots.

Chapter 6

Results

In this chapter, the measured inclusive jet differential cross sections in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ events are compared to NLO pQCD predictions. The theoretical predictions are corrected for non-perturbative contributions. The measurements are also compared to the predictions from several LO plus parton shower Monte Carlo programs.

6.1 NLO pQCD Calculation

The NLO pQCD predictions are obtained with the MCFM program [16] that provides NLO predictions for up to two jets in the final state and LO predictions for up to 3 jets in the final state (Section 2.6). Calculations are performed using CTEQ6.1M PDFs [58]. The renormalization and factorization scales, μ_{R} and μ_{F} respectively, are set equal and given by $\mu_0 = \sqrt{(M_Z)^2 + (p_{\text{T}}^Z)^2}$. Jets are search for using a Midpoint algorithm with a cone size of $R = 0.7$ and $R_{\text{sep}} = 1.3$ (Section 2.9.1) where, compared to the original implementation, the clustering of the partons has been modified to accommodate the E-scheme. The next sections describe the non-perturbative contributions to the cross sections and the uncertainties considered in the theoretical predictions.

6.1.1 Non-pQCD Contributions

Fixed-order pQCD predictions are compared to the data that have been corrected to the hadron level. The perturbative nature of the parton calculation implies that non-perturbative processes like the underlying event or the fragmentation of the partons into hadrons are not included. The contributions due to non-pQCD effects are estimated from Monte Carlo. We used PYTHIATUNE A that has been shown to accurately describe the jet shapes and energy flow in the data (see Chapter 5), and the $p_{\text{T}}^{\text{jet}}$ distribution (see Section 4.6).

The underlying event activity adds extra energy inside the cone of the jet from soft interactions between the $p\overline{p}$ remnants. On the other hand, the hadronization of the partons introduces some transverse momentum to the hadrons with respect to the original parton direction and some of them end up out of the cone of the jet. This leads to negative contribution to the jet energy. The harder is the jet, the more collimated are the hadrons leading to smaller underlying event and hadronization corrections at higher $p_{\mathrm{T}}^{\mathrm{jet}}$.

A global non-perturbative correction factor C_{HAD} is determined by comparing the cross sections in Pythia-Tune A with and without underlying event, and before and after the hadronization of the partons carried out using string fragmentation:

$$C_{\mathrm{HAD}}(\alpha) = \frac{\left(\frac{d\sigma}{d\alpha}\right)^{\mathrm{UE}}_{\mathrm{HAD}}}{\left(\frac{d\sigma}{d\alpha}\right)^{\mathrm{no\ UE}}_{\mathrm{PAR}}} \qquad \alpha \equiv p_{\mathrm{T}}^{\mathrm{jet}}, y^{\mathrm{jet}}, N_{\mathrm{jets}} \tag{6.1}$$

where HAD refers to hadron level cross sections, PAR to parton level cross sections before hadronization, and UE and no UE refer to cross sections with and without generating underlying event activity, respectively. The corrections due to underlying event and due to hadronization, a priori, do not factorize, but the relative contribution from each one can be approximately estimated in the following way:

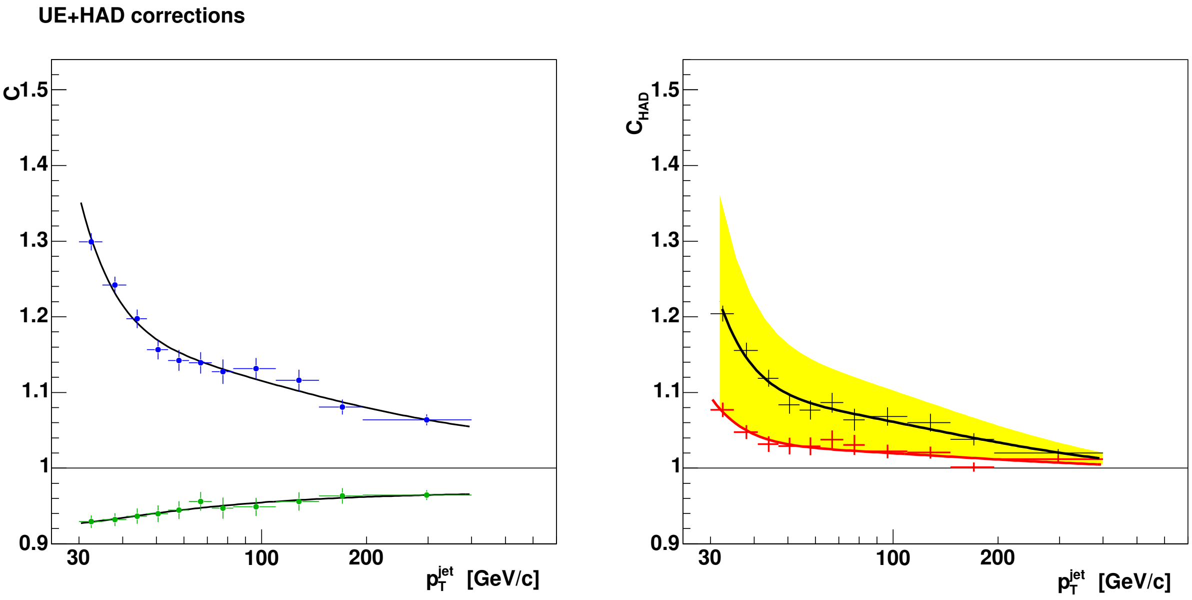
$$C_{\mathrm{HAD}}(\alpha) = C_{\mathrm{UE}} \times C_{\mathrm{hadronization}} = \frac{\left(\frac{d\sigma}{d\alpha}\right)^{\mathrm{UE}}_{\mathrm{HAD}}}{\left(\frac{d\sigma}{d\alpha}\right)^{\mathrm{no\ UE}}_{\mathrm{HAD}}} \times \frac{\left(\frac{d\sigma}{d\alpha}\right)^{\mathrm{no\ UE}}_{\mathrm{HAD}}}{\left(\frac{d\sigma}{d\alpha}\right)^{\mathrm{no\ UE}}_{\mathrm{PAR}}} \tag{6.2}$$

Figure 6.1 shows C_{UE}, $C_{\mathrm{hadronization}}$ and the global C_{HAD} as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for the case of $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet cross section. Underlying event corrections largely dominate over the hadronization corrections mainly in the low $p_{\mathrm{T}}^{\mathrm{jet}}$ region ($p_{\mathrm{T}}^{\mathrm{jet}} < 100$ GeV/c). C_{UE} is large at low $p_{\mathrm{T}}^{\mathrm{jet}}$ (30%) and decreases rapidly. $C_{\mathrm{hadronization}}$ corrects the cross section in the opposite direction ($\sim 8\% - 4\%$) and it also decreases with $p_{\mathrm{T}}^{\mathrm{jet}}$.

The C_{HAD} corrections are found to be constant with $|y^{\mathrm{jet}}|$ (Fig. 6.2), around 1.13 for the $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet cross section and 1.19 for the $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 2$ jets cross section. The C_{HAD} corrections increase at higher jet multiplicities as the $p_{\mathrm{T}}^{\mathrm{jet}}$ distributions for the second and third jets are more sensitive to the extra contributions from the underlying event. This is also observed in the $p_{\mathrm{T}}^{\mathrm{jet}}$ and N_{jets} distributions. The values of C_{HAD} used to correct the NLO cross sections are reported in Tables XVI to XIX as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$, $|y^{\mathrm{jet}}|$ and N_{jets}.

6.1.2 PDF Uncertainty on the Theoretical Predictions

The uncertainties due to the PDFs are computed with the Hessian method, as described in Section 2.6.2. The + and − deviations along the 20 eigenvectors of CTEQ6.1M PDFs



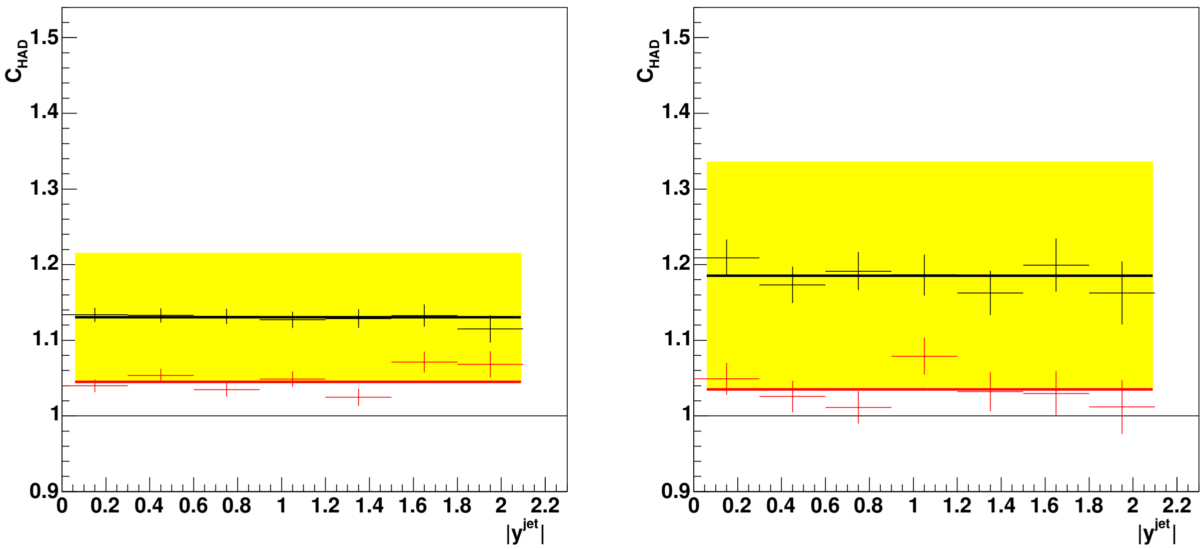
(a) (b)

Figure 6.1: (a) C_{UE} (blue) and $C_{\mathrm{hadronization}}$ (green) corrections for $Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet cross section. (b) Global C_{HAD} correction factors determined using PYTHIATUNE A (black) and using TUNE DW (red). The nominal correction is TUNE A and the yellow band indicates the assigned systematic uncertainty that is the symmetrized difference between TUNE A and TUNE DW corrections.

are taken. Asymmetric uncertainties are obtained by summing in quadrature the maximal deviation in each direction associated to each of the 20 eigenvectors (Eq. 2.26). The total PDF uncertainty varies from 2% at low $p_{\mathrm{T}}^{\mathrm{jet}}$ to 10% at high $p_{\mathrm{T}}^{\mathrm{jet}}$. As a function of $|y^{\mathrm{jet}}|$, the uncertainties are rather constant around 3%. The uncertainties on the theoretical predictions due to the PDFs are shown in the Data/Theory ratios in Figs. 6.3 to 6.9.

6.1.3 Dependence on Renormalization and Factorization Scale

Changing the renormalization and factorization scales by a factor of 2 from its nominal value, $\mu_0/2$ and $2\mu_0$, makes the cross section prediction to vary between 10% and 15%. The changes in the theoretical predictions due to the variation of the scale are shown in the Data/Theory ratios in Figs. 6.3 to 6.9 and they are summarized in Table XV.



(a) (b)

Figure 6.2: $C_{\text{HAD}}(|y^{\text{jet}}|)$ corrections for (a) $Z/\gamma^*(\to e^+e^-)+\geq 1$ jet and (b) $Z/\gamma^*(\to e^+e^-)+\geq 2$ jets cross sections. The correction factors determined using PYTHIATUNE A (black) and using TUNE DW (red). The yellow band indicates the assigned systematic uncertainty that is the symmetrized difference between TUNE A, the nominal factors, and TUNE DW.

	Uncertainty/Variation
Non-pQCD corrections	2-4%
PDF uncertainties	2-10%
Renorm. and fact. scale variation	10-15%

Table XV: Uncertainties and dependencies of the NLO pQCD cross sections.

6.1.4 Uncertainty on the non-pQCD Corrections

The $C_{\rm HAD}$ correction is model dependent because it is estimated using a given Monte Carlo generator. PYTHIA-TUNE DW, which has a different model for the underlying event, is employed to estimated the uncertainty on $C_{\rm HAD}$. Chapter 5 shows the jet shapes and the energy flow measured in the data that are well reproduced by PYTHIA-TUNE DW. Therefore,PYTHIA-TUNE DW is also a good tool to obtain realistic corrections to compare to the nominal $C_{\rm HAD}$. Figure 6.1(b) shows the difference between the non-pQCD corrections using both underlying event models as a function of $p_{\rm T}^{\rm jet}$. The differences are larger at low $p_{\rm T}^{\rm jet}$ (5%) where the underlying event has a larger effect. The uncertainty is taken as the difference between the two corrections. The same method is used to determine the uncertainties on $C_{\rm HAD}$ as a function of $y^{\rm jet}$ (Fig. 6.2) and $N_{\rm jets}$.

Is important to notice that PYTHIA-TUNE DW uses the same string hadronization model than PYTHIA-TUNE A. Hence, it is not testing any dependence on the hadronization model. Using HERWIG to estimate the uncertainty on the hadronization model could be a possible procedure, but it has been shown in Chapter 5 that HERWIG failed to reproduce the jet shapes in the data, and it would lead to unrealistic $C_{\rm HAD}$ corrections and uncertainties. However, the non-pQCD correction is largely dominated by the underlying event contribution (Fig. 6.1), and any possible dependence on the hadronization model is covered by the uncertainty already assigned to $C_{\rm HAD}$. Furthermore, uncertainties on $C_{\rm HAD}$ estimated using HERWIG in inclusive jet measurements [18] are very similar to the ones obtained here.

6.2 Comparison with theoretical predictions

Figures 6.3 and 6.4 report the inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+\geq N_{\rm jets}$ production for $N_{\rm jets} \geq 1$ and $N_{\rm jets} \geq 2$, respectively, as a function of $p_{\rm T}^{\rm jet}$ for jets with $p_{\rm T}^{\rm jet} > 30$ GeV/c and $|y^{\rm jet}| < 2.1$. The measurements are compared to the NLO pQCD predictions corrected with $C_{\rm HAD}$, also shown at the bottom of the figures. Both figures show a good agreement between data and theory within their uncertainties. To quantify the agreement a χ^2 test is performed according to the formula:

$$\chi^2 = \sum_{j=1}^{\rm bins} \frac{\left[\sigma_j^{\rm D} - \sigma_j^{\rm th}(\bar{s})\right]^2}{\left[\delta\sigma_j^{\rm D}\right]^2 + \left[\delta\sigma_j^{\rm th}(\bar{s})\right]^2} + \sum_{i=1}^{\rm syst.} [s_i]^2 \tag{6.3}$$

where $\sigma_j^{\rm D}$ is the measured cross section in bin j of the $p_{\rm T}^{\rm jet}$distribution, $\sigma_j^{\rm th}(\bar{s})$ is the corresponding theoretical prediction, and $\bar{s}$ denotes the vector of standard deviations s_i for the different independent sources of systematic uncertainties. The sum in j runs over the bins in the distribution and the sum in i, over the 10 sources of systematic uncertainty, listed in

Table XIII, plus the uncertainty on the C_{HAD} correction and the 5.8% uncertainty on the luminosity measurement (Section 3.2.6). Every source of uncertainty is treated as independent from each other and fully correlated across bins in $p_{\mathrm{T}}^{\mathrm{jet}}$. These correlations are taken into account in $\sigma_j^{\mathrm{th}}(\bar{s})$ where variations of s_i for a given source i affect coherently all the $\sigma_j^{\mathrm{th}}(\bar{s})$ values. The χ^2 formula is minimized with respect to $\bar{s}$. After the minimization, $\bar{s}$ represents the configuration of systematic uncertainties that leads to the best agreement between the theoretical prediction and the measured cross section. The χ^2 test gives a probability of 99% for $N_{\mathrm{jets}} \geq 1$ and 22% for $N_{\mathrm{jets}} \geq 2$, supporting the good agreement observed.

The differential jet cross sections as a function of $|y^{\mathrm{jet}}|$ for $N_{\mathrm{jets}} \geq 1$ and $N_{\mathrm{jets}} \geq 2$ are presented in Figs. 6.5 and 6.6 together with the C_{HAD} corrections applied in each case. The good agreement between the data and the theoretical predictions is also observed in this case. As a complementary result, the differential cross section is also measured as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ of the first leading jet in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet events, and as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ of the second leading jet in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 2$ jets (Figs. 6.7 and 6.8).

The total event cross section as a function of the jet multiplicity N_{jets} is shown in Fig. 6.9 compared to NLO pQCD predictions for $N_{\mathrm{jets}} \geq 1, 2$ and to LO pQCD predictions for $N_{\mathrm{jets}} \geq 1, 2, 3$. The NLO pQCD predictions are in good agreement with the data, with χ^2 probabilities of 83% for $N_{\mathrm{jets}} \geq 1$ and 99% for $N_{\mathrm{jets}} \geq 2$. The LO pQCD predictions underestimate the measured cross section by a factor about 1.4. The ratio Data/LO is approximately independent of N_{jets}, and seems to be compatible with a constant NLO/LO k-factor.

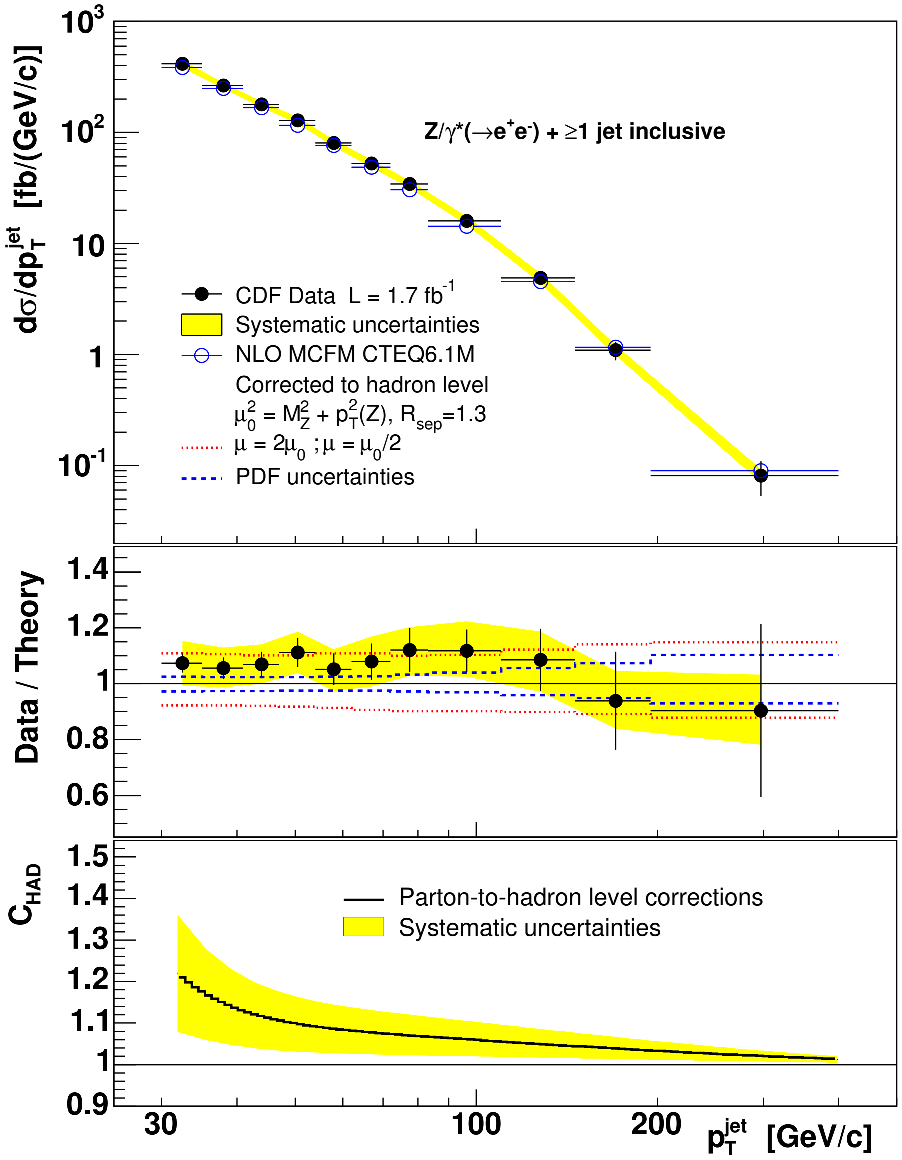


Figure 6.3: (top) Measured inclusive jet differential cross section as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+\geq 1$ jet production (black dots) compared to the NLO pQCD prediction (open circles). The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. (middle) Data/theory ratio with the PDF uncertainty (blue dashed lines) and the variation of the theoretical predictions with the scale μ (red dotted lines). (bottom) C_{HAD} corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty.

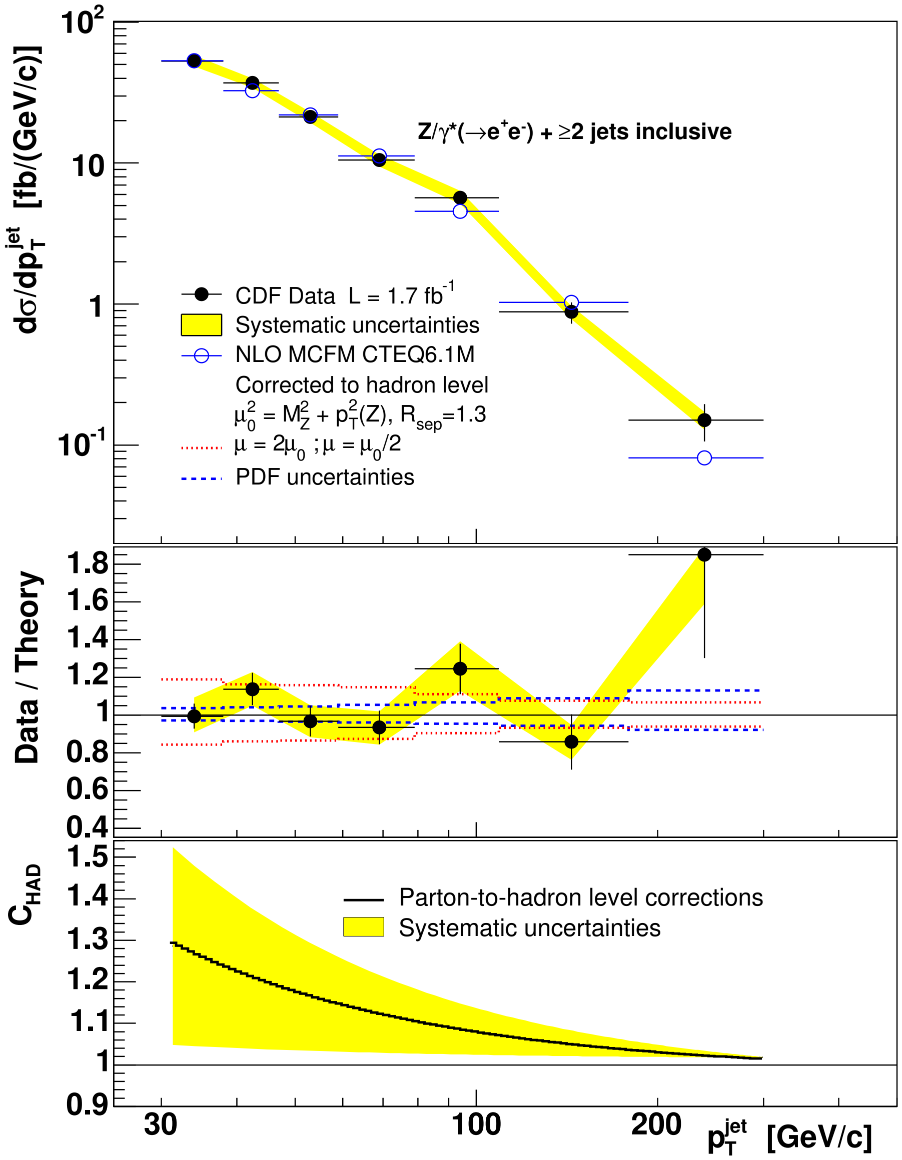


Figure 6.4: (top) Measured inclusive jet differential cross section as a function of $p_T^{\rm jet}$ in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production (black dots) compared to the NLO pQCD prediction (open circles). The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. (middle) Data/theory ratio with the PDF uncertainty (blue dashed lines) and the variation of the theoretical predictions with the scale μ (red dotted lines). (bottom) $C_{\rm HAD}$ corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty.

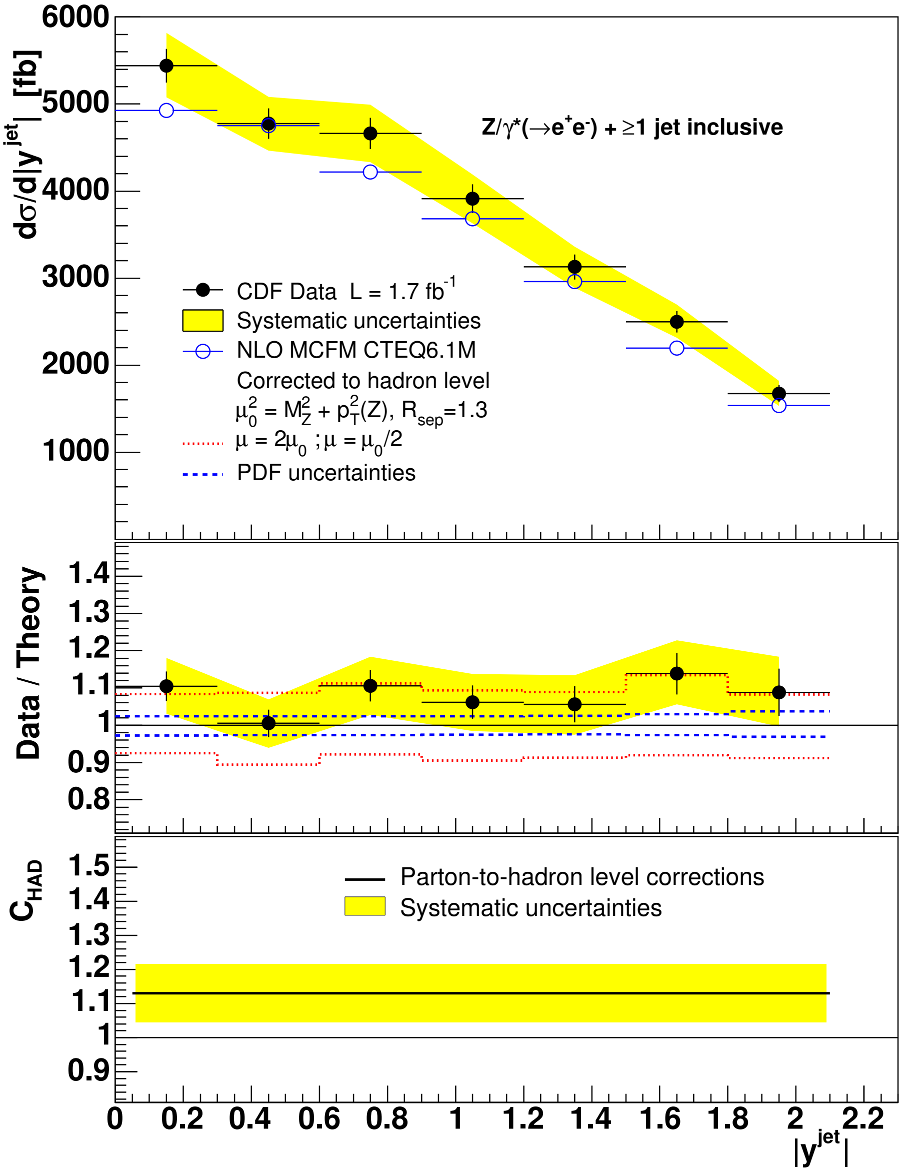


Figure 6.5: (top) Measured inclusive jet differential cross section as a function of $|y^{\rm jet}|$ in $Z/\gamma^*(\to e^+e^-)+\geq 1$ jet production (black dots) compared to the NLO pQCD prediction (open circles). The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. (middle) Data/theory ratio with the PDF uncertainty (blue dashed lines) and the variation of the theoretical predictions with the scale μ (red dotted lines). (bottom) $C_{\rm HAD}$ corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty.

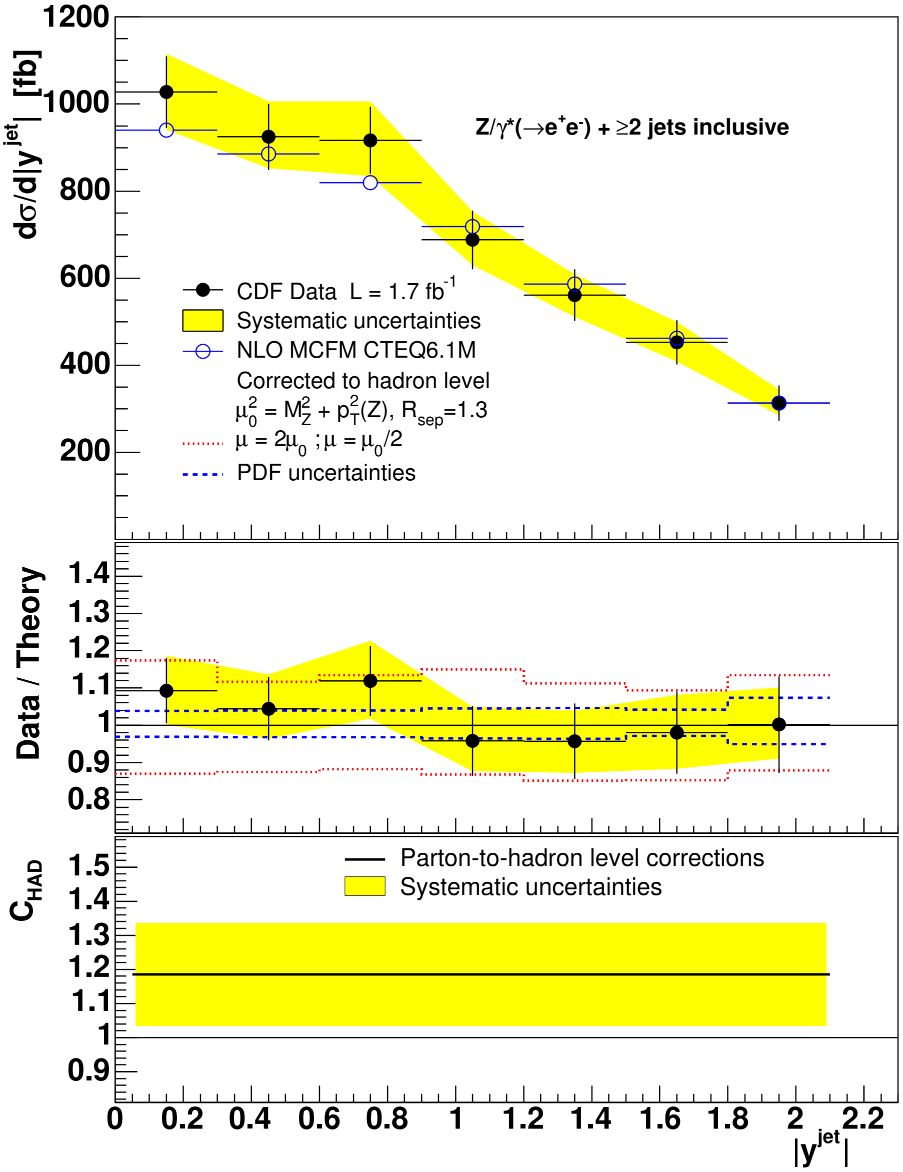


Figure 6.6: (top) Measured inclusive jet differential cross section as a function of $|y^{\mathrm{jet}}|$ in $Z/\gamma^*(\to e^+e^-)+ \geq 2$ jets production (black dots) compared to the NLO pQCD prediction (open circles). The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. (middle) Data/theory ratio with the PDF uncertainty (blue dashed lines) and the variation of the theoretical predictions with the scale μ (red dotted lines). (bottom) C_{HAD} corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty.

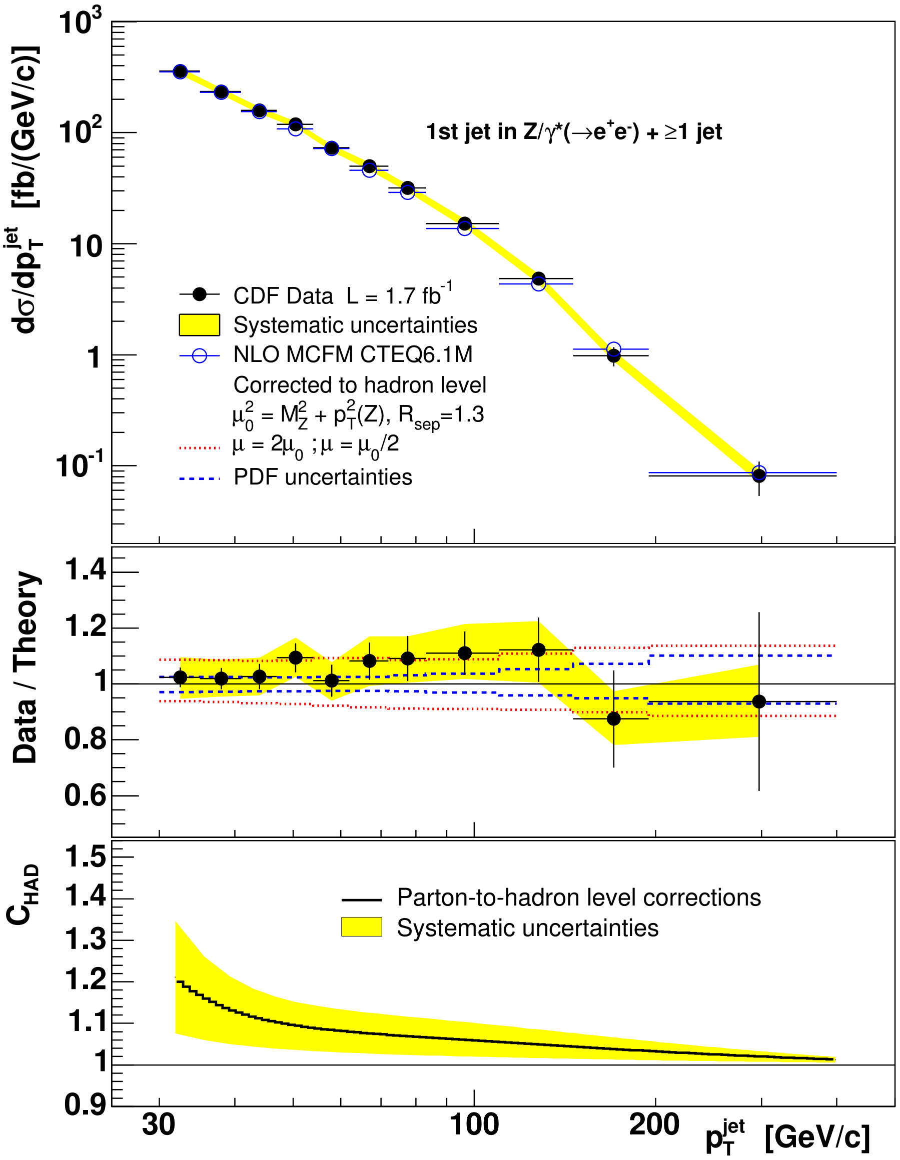


Figure 6.7: (top) Measured differential cross section as a function of the $p_T^{\rm jet}$ of the first leading jet in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet production (black dots) compared to the NLO pQCD prediction (open circles). The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. (middle) Data/theory ratio with the PDF uncertainty (blue dashed lines) and the variation of the theoretical predictions with the scale μ (red dotted lines). (bottom) $C_{\rm HAD}$ corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty.

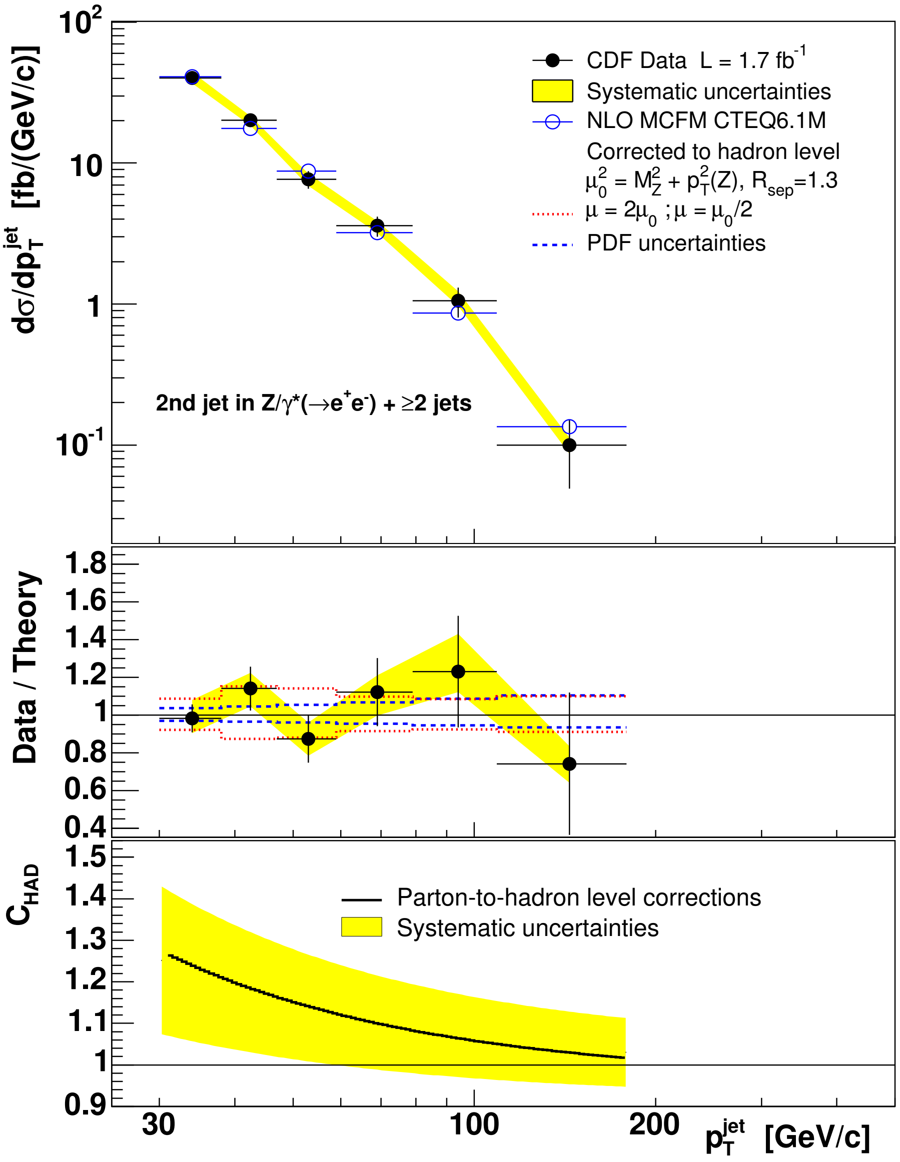


Figure 6.8: (top) Measured differential cross section as a function of the p_T^{jet} of the second leading jet in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 2$ jets production (black dots) compared to the NLO pQCD prediction (open circles). The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. (middle) Data/theory ratio with the PDF uncertainty (blue dashed lines) and the variation of the theoretical predictions with the scale μ (red dotted lines). (bottom) C_{HAD} corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty.

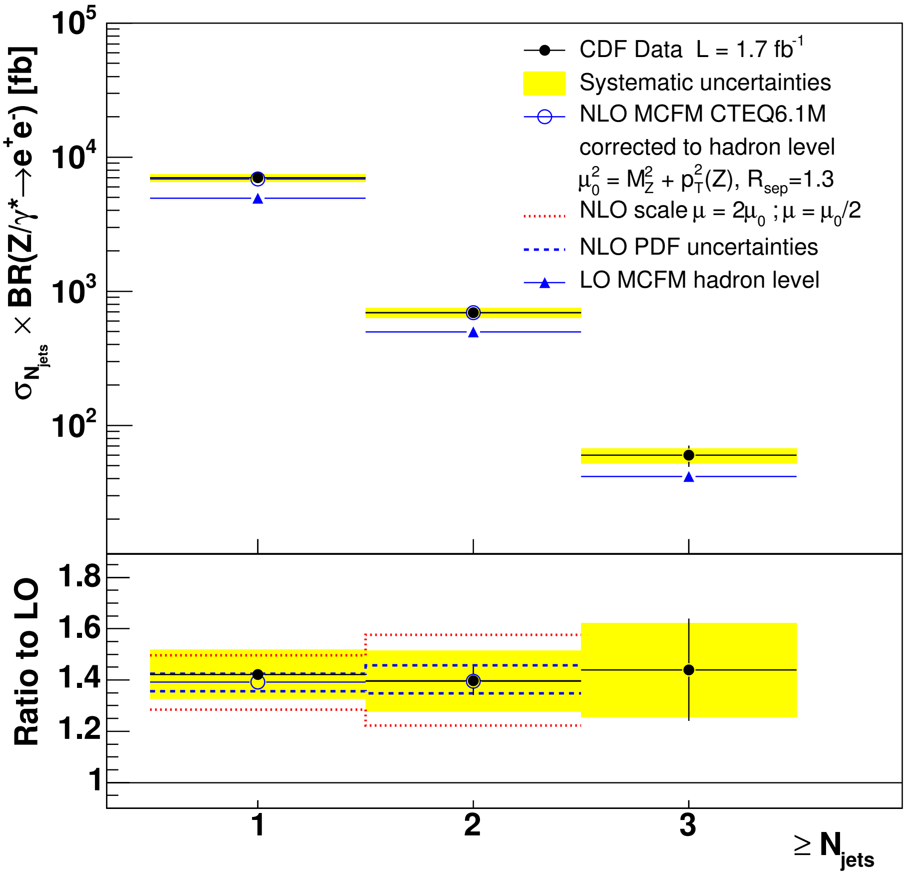


Figure 6.9: (top) Measured total event cross section in $Z/\gamma^*(\to e^+e^-) + \text{jets}$ events (black dots) as a function of inclusive jet multiplicity N_{jets} compared to LO (triangles) and NLO (open circles) pQCD predictions. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. (bottom) Data/LO (black dots) and data/LO (open circles) ratios with the NLO PDF uncertainty (blue dashed lines) and the variation of the NLO predictions with the scale μ (red dotted lines).

$p_{\mathrm{T}}^{\mathrm{jet}}$ [GeV/c]	$\frac{d\sigma}{dp_{\mathrm{T}}^{\mathrm{jet}}} \pm$ (stat.) $\pm$ (syst.) $\pm$ (lum.) [fb/(GeV/c)]	$C_{\mathrm{HAD}} \pm$ (stat.) $\pm$ (syst.) parton → hadron
$Z/\gamma^{*}(\rightarrow e^{+}e^{-})+\geq 1$ jet		
30 - 35	$413.3 \pm 13.3^{+30.4}_{-31.3} \pm 24.0$	$1.209 \pm 0.010 \pm 0.134$
35 - 41	$263.3 \pm 9.4^{+18.3}_{-17.4} \pm 15.3$	$1.146 \pm 0.010 \pm 0.096$
41 - 47	$178.3 \pm 7.5^{+12.0}_{-11.6} \pm 10.3$	$1.114 \pm 0.011 \pm 0.077$
47 - 54	$128.5 \pm 5.9^{+8.7}_{-8.4} \pm 7.5$	$1.097 \pm 0.012 \pm 0.066$
54 - 62	$80.5 \pm 4.3^{+5.5}_{-6.0} \pm 4.7$	$1.086 \pm 0.013 \pm 0.059$
62 - 72	$52.5 \pm 3.2^{+4.4}_{-4.3} \pm 3.0$	$1.078 \pm 0.013 \pm 0.053$
72 - 83	$34.2 \pm 2.4^{+2.5}_{-2.8} \pm 2.0$	$1.072 \pm 0.015 \pm 0.049$
83 - 110	$16.0 \pm 1.1^{+1.5}_{-1.3} \pm 0.9$	$1.063 \pm 0.012 \pm 0.043$
110 - 146	$4.9 \pm 0.5^{+0.5}_{-0.5} \pm 0.3$	$1.051 \pm 0.012 \pm 0.035$
146 - 195	$1.1 \pm 0.2^{+0.1}_{-0.1} \pm 0.06$	$1.040 \pm 0.008 \pm 0.027$
195 - 400	$0.08 \pm 0.03^{+0.01}_{-0.01} \pm 0.005$	$1.021 \pm 0.005 \pm 0.013$
$Z/\gamma^{*}(\rightarrow e^{+}e^{-})+\geq 2$ jets		
30 - 38	$52.9 \pm 3.5^{+5.3}_{-4.6} \pm 3.1$	$1.262 \pm 0.022 \pm 0.217$
38 - 47	$37.0 \pm 2.8^{+2.9}_{-2.8} \pm 2.1$	$1.207 \pm 0.024 \pm 0.169$
47 - 59	$21.2 \pm 1.8^{+1.9}_{-1.9} \pm 1.2$	$1.164 \pm 0.025 \pm 0.130$
59 - 79	$10.5 \pm 1.0^{+0.9}_{-1.0} \pm 0.6$	$1.123 \pm 0.024 \pm 0.093$
79 - 109	$5.7 \pm 0.6^{+0.7}_{-0.5} \pm 0.3$	$1.087 \pm 0.026 \pm 0.062$
109 - 179	$0.88 \pm 0.15^{+0.09}_{-0.10} \pm 0.05$	$1.052 \pm 0.020 \pm 0.030$
179 - 300	$0.15 \pm 0.04^{+0.02}_{-0.02} \pm 0.009$	$1.026 \pm 0.010 \pm 0.008$

Table XVI: Measured inclusive jet differential cross section in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+\geq N_{\mathrm{jets}}$ production as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ with $N_{\mathrm{jets}} \geq 1$ and $N_{\mathrm{jets}} \geq 2$. The systematic uncertainties are fully correlated across $p_{\mathrm{T}}^{\mathrm{jet}}$ bins. The parton-to-hadron correction factors $C_{\mathrm{HAD}}(p_{\mathrm{T}}^{\mathrm{jet}}, N_{\mathrm{jets}})$ are applied to the pQCD predictions.

$p_{\mathrm{T}}^{\mathrm{jet}}$	$\frac{d\sigma}{dp_{\mathrm{T}}^{\mathrm{jet}}} \pm$ (stat.) $\pm$ (syst.) $\pm$ (lum.)	$C_{\mathrm{HAD}} \pm$ (stat.) $\pm$ (syst.)
[GeV/c]	[fb/(GeV/c)]	parton → hadron
First leading jet in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+ \geq 1$ jet		
30 - 35	$360.4 \pm 12.3^{+25.2}_{-26.7} \pm 20.9$	$1.200 \pm 0.011 \pm 0.128$
35 - 41	$234.7 \pm 8.8^{+15.8}_{-14.6} \pm 13.6$	$1.140 \pm 0.011 \pm 0.087$
41 - 47	$159.2 \pm 7.0^{+10.4}_{-10.4} \pm 9.2$	$1.110 \pm 0.012 \pm 0.067$
47 - 54	$118.7 \pm 5.6^{+7.8}_{-7.5} \pm 6.9$	$1.094 \pm 0.012 \pm 0.058$
54 - 62	$72.9 \pm 4.1^{+4.8}_{-5.2} \pm 4.2$	$1.084 \pm 0.013 \pm 0.053$
62 - 72	$49.8 \pm 3.1^{+4.1}_{-4.1} \pm 2.9$	$1.076 \pm 0.013 \pm 0.049$
72 - 83	$31.7 \pm 2.3^{+2.3}_{-2.5} \pm 1.8$	$1.070 \pm 0.015 \pm 0.046$
83 - 110	$15.2 \pm 1.1^{+1.4}_{-1.3} \pm 0.9$	$1.062 \pm 0.012 \pm 0.041$
110 - 146	$4.9 \pm 0.5^{+0.4}_{-0.5} \pm 0.3$	$1.051 \pm 0.012 \pm 0.034$
146 - 195	$0.98 \pm 0.20^{+0.11}_{-0.11} \pm 0.06$	$1.040 \pm 0.008 \pm 0.027$
195 - 400	$0.08 \pm 0.03^{+0.01}_{-0.01} \pm 0.005$	$1.020 \pm 0.005 \pm 0.013$
Second leading jet in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+ \geq 2$ jets		
30 - 38	$40.1 \pm 3.1^{+4.0}_{-3.2} \pm 2.3$	$1.222 \pm 0.024 \pm 0.165$
38 - 47	$20.1 \pm 2.1^{+1.5}_{-1.7} \pm 1.2$	$1.174 \pm 0.030 \pm 0.145$
47 - 59	$7.7 \pm 1.1^{+0.7}_{-0.8} \pm 0.4$	$1.137 \pm 0.035 \pm 0.129$
59 - 79	$3.6 \pm 0.6^{+0.3}_{-0.4} \pm 0.2$	$1.102 \pm 0.040 \pm 0.113$
79 - 109	$1.06 \pm 0.25^{+0.17}_{-0.09} \pm 0.06$	$1.071 \pm 0.048 \pm 0.100$
109 - 179	$0.10 \pm 0.05^{+0.01}_{-0.01} \pm 0.009$	$1.041 \pm 0.057 \pm 0.087$

Table XVII: Measured inclusive jet differential cross section in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+ \geq N_{\mathrm{jets}}$ production as a function of the $p_{\mathrm{T}}^{\mathrm{jet}}$ of the N^{th} leading jet with $N_{\mathrm{jets}} \geq 1$ and $N_{\mathrm{jets}} \geq 2$. The systematic uncertainties are fully correlated across $p_{\mathrm{T}}^{\mathrm{jet}}$ bins. The parton-to-hadron correction factors $C_{\mathrm{HAD}}(p_{\mathrm{T}}^{\mathrm{jet}}, N_{\mathrm{jets}})$ are applied to the pQCD predictions.

$\lvert y^{\rm jet}\rvert$	$\frac{d\sigma}{d\lvert y^{\rm jet}\rvert} \pm$ (stat.) $\pm$ (syst.) $\pm$ (lum.) [fb]	$C_{\rm HAD} \pm$ (stat.) $\pm$ (syst.) parton → hadron
	$Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet	
0.0 - 0.3	$5442 \pm 195^{+375}_{-364} \pm 316$	$1.130 \pm 0.004 \pm 0.085$
0.3 - 0.6	$4776 \pm 177^{+306}_{-313} \pm 277$	$1.130 \pm 0.004 \pm 0.085$
0.6 - 0.9	$4664 \pm 178^{+331}_{-330} \pm 271$	$1.130 \pm 0.004 \pm 0.085$
0.9 - 1.2	$3913 \pm 166^{+280}_{-284} \pm 227$	$1.130 \pm 0.004 \pm 0.085$
1.2 - 1.5	$3129 \pm 144^{+231}_{-242} \pm 181$	$1.130 \pm 0.004 \pm 0.085$
1.5 - 1.8	$2499 \pm 122^{+195}_{-182} \pm 145$	$1.130 \pm 0.004 \pm 0.085$
1.8 - 2.1	$1672 \pm 98^{+147}_{-141} \pm 97$	$1.130 \pm 0.004 \pm 0.085$
	$Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets	
0.0 - 0.3	$1027 \pm 82^{+88}_{-86} \pm 60$	$1.185 \pm 0.011 \pm 0.150$
0.3 - 0.6	$925 \pm 76^{+81}_{-72} \pm 54$	$1.185 \pm 0.011 \pm 0.150$
0.6 - 0.9	$917 \pm 77^{+89}_{-83} \pm 53$	$1.185 \pm 0.011 \pm 0.150$
0.9 - 1.2	$688 \pm 67^{+66}_{-58} \pm 40$	$1.185 \pm 0.011 \pm 0.150$
1.2 - 1.5	$561 \pm 59^{+48}_{-50} \pm 33$	$1.185 \pm 0.011 \pm 0.150$
1.5 - 1.8	$452 \pm 51^{+47}_{-44} \pm 26$	$1.185 \pm 0.011 \pm 0.150$
1.8 - 2.1	$313 \pm 40^{+31}_{-28} \pm 18$	$1.185 \pm 0.011 \pm 0.150$

Table XVIII: Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+\geq N_{\rm jets}$ production as a function of $\lvert y^{\rm jet}\rvert$ with $N_{\rm jets} \geq 1$ and $N_{\rm jets} \geq 2$. The systematic uncertainties are fully correlated across $y^{\rm jet}$ bins. The parton-to-hadron correction factors $C_{\rm HAD}(y^{\rm jet}, N_{\rm jets})$ are applied to the pQCD predictions.

$N_{\rm jets}$	$\sigma \pm$ (stat.) $\pm$ (syst.) $\pm$ (lum.) [fb]	$C_{\rm HAD} \pm$ (stat.) $\pm$ (syst.) parton → hadron
≥ 1	$7003 \pm 146^{+483}_{-470} \pm 406$	$1.138 \pm 0.005 \pm 0.072$
≥ 2	$695 \pm 37^{+59}_{-60} \pm 40$	$1.192 \pm 0.016 \pm 0.109$
≥ 3	$60 \pm 11^{+8}_{-8} \pm 3.5$	$1.367 \pm 0.069 \pm 0.325$

Table XIX: Measured total event cross section in inclusive $Z/\gamma^*(\rightarrow e^+e^-)+\geq N_{\rm jets}$ production for $N_{\rm jets} \geq 1$, $N_{\rm jets} \geq 2$ and $N_{\rm jets} \geq 3$. The systematic uncertainties are fully correlated across $N_{\rm jets}$ bins. The parton-to-hadron correction factors $C_{\rm HAD}(N_{\rm jets})$ are applied to the pQCD predictions.

6.3 Comparison to LO Matrix Elements + Parton Shower Predictions

The measured cross sections and several additional observables have been compared to LO plus parton shower Monte Carlo predictions (see Section 2.8).

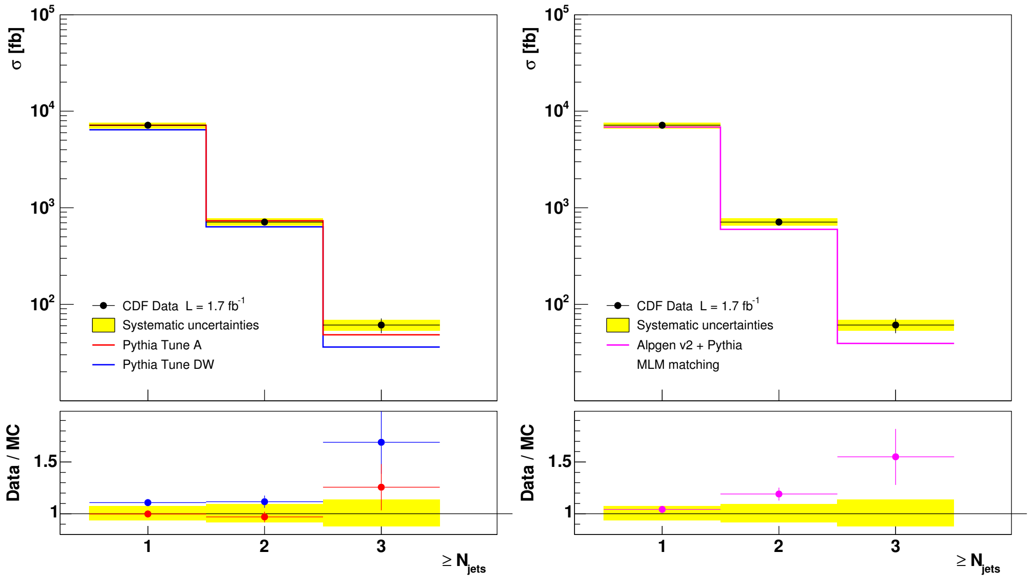
6.3.1 Jet Multiplicity

In PYTHIA $p\overline{p} \rightarrow Z/\gamma^{*}+p+X$ processes, the leading parton is generated using the 2-to-2 exact LO matrix elements, followed by a parton shower for which the first radiation is generated with an angular distribution that mimics the matrix elements. This should provide a rather good description of the p_{T} of the first two jets in the event. Figure 6.10(a) shows the total event cross section in $Z/\gamma^{*} \rightarrow e^{+}e^{-}$ events as a function of the inclusive jet multiplicity compared to PYTHIA predictions. The ratio Data/MC for events with $N_{\mathrm{jets}} \geq 1$ and $N_{\mathrm{jets}} \geq 2$ is constant, which indicates a good description of the relative jet population by PYTHIA. The case with three jets in the final state suffers from the lack of statistics but some deviation starts to be observed. While TUNE A also describes the absolute cross section, TUNE DW differs by approximately 20%. As was already pointed out, PYTHIA-TUNE A describes the absolute rate of jets.

In the ALPGEN sample considered in this thesis, exclusive $Z/\gamma^{*} + n$ *partons* final states has been produced for $n = 0, 1, 2, 3, 4$ using LO matrix elements, interfaced with PYTHIA for parton shower evolution. The PDFs used are CTEQ5L and PYTHIA is also employed for the hadronization process. The multi-parton samples are MLM-matched and then combined using their relative cross sections provided by ALPGEN, and the final sample is normalized to the inclusive Drell-Yan cross section measured at CDF [55]. A special set of parameters, denoted as PYTHIA-TUNE BW [26], is employed to model the underlying event. TUNE BW was specially developed for the MLM matching and modifies the TUNE A in order to avoid double counting of soft radiation in the $Z/\gamma^{*} + 0\ p$ final state. Figure 6.10 shows the total cross section as a function of the jet multiplicity compared to the ALPGEN + PYTHIA prediction. The ratio Data/MC shows a clear deviation, where ALPGEN underestimates the jet rates at large N_{jets}, and fails to describe the shape of the cross section measured in the data. At the time of writing this thesis, further studies are being carried out in order to further tune the parton shower parameters in ALPGEN.

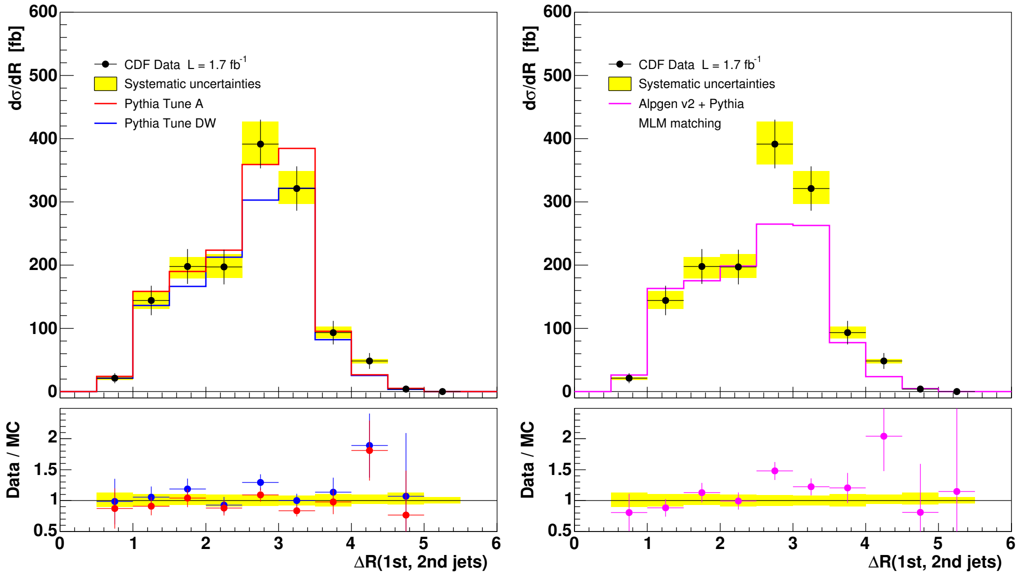
6.3.2 Jet Topologies in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+ \geq 2$ jets Events

Figures 6.11, 6.12 and 6.13 present the differential cross sections in $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+ \geq 2$ jets events as a function of ΔR_{jj}, $\Delta\phi_{\mathrm{jj}}$ and M_{jj} of the two leading jets in the event. The measurements are compared to the predictions from Pythia and Alpgen+Pythia. All the comparisons are clearly limited by the statistics but, in general, it can be concluded that Pythia-Tune A is the sample that better reproduces the shapes measured in the data. Also Tune DW, with a similar shape than Tune A, describes well the data. On the other hand, Alpgen+Pythia do not reproduce the ΔR_{jj} distribution for $\Delta R_{\mathrm{jj}} \sim \pi$.



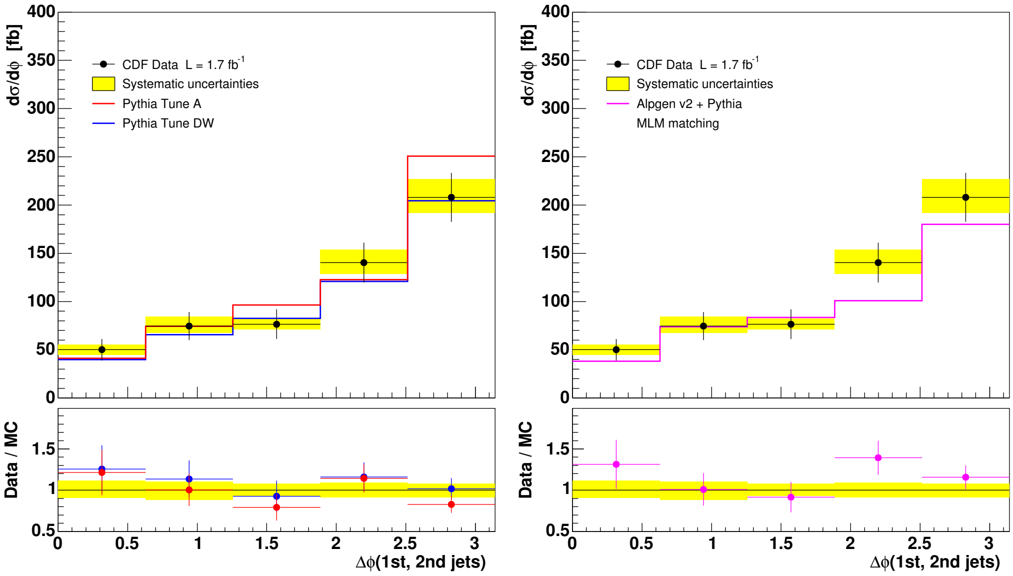
(a) (b)

Figure 6.10: Measured total event cross section in inclusive $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ (black dots) as a function of jet multiplicity N_{jets} compared to (a) Pythia predictions (lines) using the Tune A and Tune DW underlying event settings, and to (b) alpgen v2 predictions (line) matched to the parton shower from Pythia using the MLM matching technique.



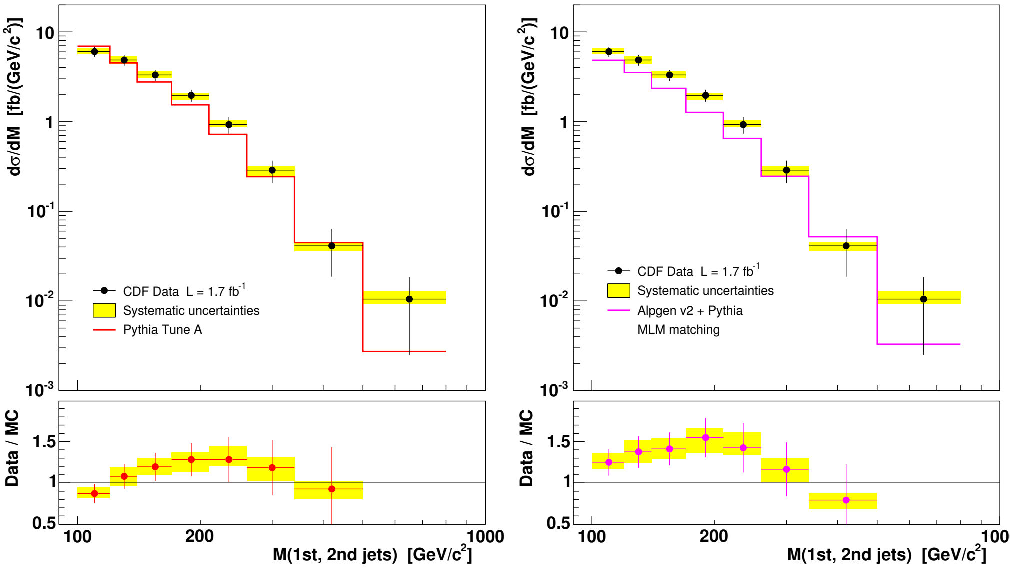
(a) (b)

Figure 6.11: Measured differential cross section as a function of ΔR_{jj} in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production (black dots) compared to (a) Pythia predictions (lines) using the Tune A and Tune DW underlying event settings, and to (b) Alpgen v2 predictions (line) matched to the parton shower from Pythia using the MLM matching technique.



(a) (b)

Figure 6.12: Measured differential cross section as a function of $\Delta\phi_{\mathrm{jj}}$ in $Z/\gamma^{*}(\rightarrow e^{+}e^{-}) + \geq 2$ jets production (black dots) compared to (a) Pythia predictions (lines) using the Tune A and Tune DW underlying event settings, and to (b) Alpgen v2 predictions (line) matched to the parton shower from Pythia using the MLM matching technique.



(a) (b)

Figure 6.13: Measured differential cross section as a function of M_{jj} in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production (black dots) compared to (a) Pythia predictions (lines), and to (b) Alpgen v2 predictions (line) matched to the parton shower from Pythia using the MLM matching technique.

Chapter 7

Summary and Conclusions

The measurement of the inclusive production of collimated jets of hadrons in association with a Z/γ^* boson in $p\overline{p}$ collisions provides a stringent test of pQCD, and is sensitive to the presence of new particles decaying into Z/γ^*+jets final states. At the leading order in pQCD, $Z/\gamma^* +$ jets events are driven by the processes $gq \rightarrow Z/\gamma^* + q$ and $q\bar{q} \rightarrow Z/\gamma^* + g$, while higher orders contributions, including additional parton radiation, produce multiple jets in the final state. NLO pQCD predictions for $Z/\gamma^* +$ jets production are only available for jet multiplicities up to $N_{\text{jets}} = 2$. The understanding of $Z/\gamma^* +$ jets final states from data is therefore crucial since they also constitute important irreducible backgrounds in searches for new physics. Previous results from Run I at the Tevatron have been compared to LO plus parton shower Monte Carlo predictions affected by large scale uncertainties.

This thesis reports new measurements of the inclusive jet cross sections in $Z/\gamma^* \rightarrow e^+e^-$ production using 1.7 fb^{-1} of data collected by the CDF experiment in Run II. The final results refer to hadron level jets with $p_{\text{T}}^{\text{jet}} > 30$ GeV/c and $|y^{\text{jet}}| < 2.1$, in a limited and well-defined kinematic range for the Z/γ^* decay products: $E_{\text{T}}^{\text{e}} > 25$ GeV, $|\eta^{e_1}| < 1.0$, $|\eta^{e_2}| < 1.0$ or $1.2 < |\eta^{e_2}| < 2.8$, $66 < M_{ee} < 116$ GeV/c^2 and $\Delta R_{e-\text{jet}} < 0.7$. The data are compared to NLO pQCD predictions, which are computed using the MCFM program with CTEQ6.1M PDFs, with the renormalization and factorization scales set to $\mu^2 = (M_Z)^2 + (p_{\text{T}}^Z)^2$, and using a Midpoint algorithm with $R = 0.7$ and $R_{\text{sep}} = 1.3$ to reconstruct jets at the parton level. The theoretical predictions include parton-to-hadron correction factors that approximately account for non-perturbative contributions from the underlying event and fragmentation into hadrons. The measured cross sections are well described by NLO pQCD predictions including non-perturbative corrections.

This analysis constitutes a baseline for future studies at the LHC, where $Z/\gamma^* +$jets final states will be used to understand Standard Model processes as well as to look for physics beyond the Standard Model.

Appendix A

Data Quality Monitoring

CDF has implemented a series of systems to monitor the quality of the data recorded by the detector and used in the physics analysis. Two main parts can be distinguished in the data quality monitoring: *online*, where the proper working of the different parts of the detector and the data taking systems are monitored; and *offline*, where the production of the final datasets and calibrations are checked. The final output of the complete Data Quality Monitoring (DQM) system is a series of lists of 'good runs' in which the required parts of the detector have been checked to work properly and the data were correctly stored and processed.

The following sections focus mainly in the description of the DQMon monitor for which I made major contributions.

A.1 Online DQM System

The online DQM system is integrated in the CDF online monitoring [59]. It is intended as an early warning system for malfunctioning subsystems or subdetectors as data is being taken. It uses a copy of a small fraction of the events that passed the Level 3 requirements, that the Consumer Server Logger sends to 10 different programs (*consumers*) running in parallel. The consumers fill diagnostic histograms and perform tests on them to identify possible problems in the detector or the data acquisition systems. Some of the consumers relevant for the tasks of the DQMon monitor (see Section A.1.1) are:

- **YMon**. During a physics run, YMon receives minimum bias events and fills histograms of occupancies (% hits per channel per event), average energy per channel and energy distributions.

- **BeamMon**. The purpose of BeamMon is to locate the beam position and determine some properties of the beam spot.

- **LumMon**. It is a consumer that monitors the performance of the CLC counters and the online luminosity measurements obtained with them. LumMon can also measure the position of the interaction point with the time difference between the East and West modules of the CLC.

- **SVXMon**. It is the Silicon monitoring consumer for the online and offline diagnostics of the Silicon Vertex detector. For every Silicon chip, SVXMon creates histograms of occupancies, average pulse heights, distribution shapes, etc.

- **PhysMon**. After some processing of the data, PhysMon measures some higher level objects like electrons, muons, photons, Z, W, J/Ψ and trigger efficiencies.

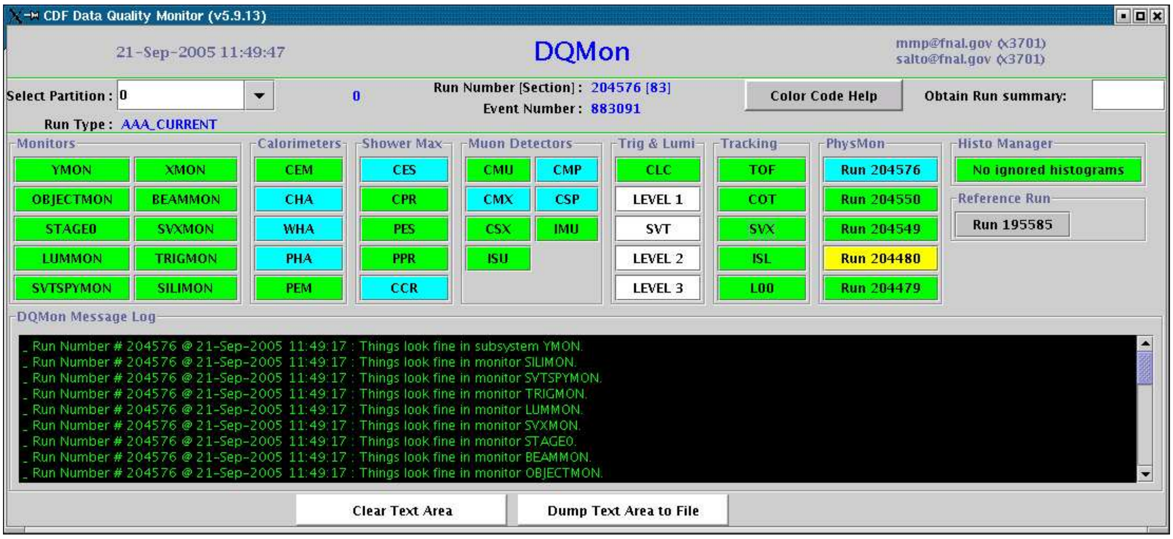
One person of the shift crew in the control room, the Consumer Operator (CO), is in charge to check the output histograms of the consumers and to bring to the attention of the Scientific Coordinator (SciCo) any potential problem. The CO has a list of items to check in every histogram, elaborated by the experts on every subsystem. Thirty minutes after a run has started, then every 2 hours, and at the end of every run, the CO goes through the check list and determines which subsystems can be set good or bad for that run. Most of the checks are based on YMon histograms: occupancy plots from the CLC, COT, TOF, calorimeters, showermax detectors and muon systems allow to find dead and hot channels. Other checks consist on comparing performance plots with templates or reference plots.

At the end of every run, the SciCo uses the results of the check list to fill the *online good run bits* in a database, where every bit represents the status of a given subsystem during that run.

A.1.1 DQMon Monitor

The Data Quality Monitor (DQMon) is an automated system developed to continuously perform the CO checklist tests on the consumer outputs. DQMon is intended to be complementary to the task of the CO. The monitor is continuously running in the control room and can issue visual (popup windows) and voice alarms to alert the shift crew in case a problem is found. As an automated system, DQMon is very fast and efficient in tasks like counting dead and hot channels, calculating histogram averages, and is a valuable tool to spot possible problems fast.

Figure A.1 shows the display of DQMon running in the control room of the detector. It shows the status of the consumer programs, the results of the checks on the detector subsystems, and the PhysMon results on the last five runs. The color code used in the buttons is common in all the monitors in the CDF control room: green: ok, red: problem or bad, blue: not enough statistics to perform the checks, yellow: warning or close to a problem, and white: no information. The leftmost buttons show the status of the consumer programs. DQMon checks periodically if they are updating their outputs.


CDF Data Quality Monitor (v5.9.13)
DQMon
Color Code Help
Obtain Run summary:
Monitors
YMON
XMON
OBJECTMON
BEAMMON
STAGE0
SVXMON
LUMMON
TRIGMON
SVTSPYMON
SILIMON
Calorimeters
CEM
CHA
WHA
PHA
PEM
Shower Max
CES
CPR
PES
PPR
CCR
Muon Detectors
CMU
CMP
CMX
CSP
CSX
IMU
ISU
Trig & Lumi
CLC
LEVEL 1
SVT
LEVEL 2
LEVEL 3
Tracking
TOF
COT
SVX
ISL
L00
PhysMon
Run 204576
Run 204550
Run 204549
Run 204480
Run 204479
Histo Manager
No ignored histograms
Reference Run
Run 195585
DQMon Message Log
Clear Text Area
Dump Text Area to File

21-Sep-2005 11:49:47

mmp@fnal.gov (x3701)
salto@fnal.gov (x3701)

Select Partition : 0 0

Run Number [Section] : 204576 [83]
Event Number : 883091

Run Type : AAA_CURRENT

```
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in subsystem YMON.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor SILIMON.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor SVTSPYMON.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor TRIGMON.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor LUMMON.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor SVXMON.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor STAGE0.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor BEAMMON.
_ Run Number # 204576 @ 21-Sep-2005 11:49:17 : Things look fine in monitor OBJECTMON.
```

Figure A.1: DQMon panel displays the status of some monitors and the results of the checks performed in the data from several subsystems.

The buttons next to the consumers buttons show the results of the check list. The histograms in YMon, LumMon and BeamMon are grouped according to the subsystem they belong, and every buttons represents one subsystem. The results of the check list performed in the histograms are displayed in the color of the buttons. If any histogram does not pass the checks, the buttons of the corresponding subsystem becomes red. Clicking on a button allows to see the results of every individual histogram, making easier to investigate the source of the problem. The monitoring of entire subsystems (buttons) can be disabled if necessary, but it is also possible to disable individual histograms while the button keeps monitoring the rest of the histograms in that subsystem.

The results of the last five runs analyzed by PhysMon is displayed in the rightmost column of buttons in DQMon using the same color code. Clicking on a given run allows to see further details on the checks performed by PhysMon.

DQMon outputs a first diagnosis for the online bits of every run according to the results of the checks. These bits are only an educated suggestion as the ultimate decision is always taken by the SciCo.

A.2 Offline DQM System

The online DQM system is complemented by the offline monitoring system, that can perform more robust checks with full statistics. A number of observables is chosen to monitor the behavior of different detector subsystems: calorimeters, tracking and muon systems. The offline monitoring also includes checks of high level objects like electrons, photons, jets and J/Ψ. As the offline checks are performed after reprocessing the data (Production+Validation) it also monitors the production process itself.

The result are the final run bits where the status of every subsystem is set. Using the run bits, several good run lists are produced depending on the subdetector data required in every physics analysis.

A.2.1 Partial Run Recovery

Some long runs (some of them are longer than 24 hours) had to be discarded because some component of the detector failed to work during only a fraction of the run. To recover the good data from these runs, the offline checks are performed only in a part of the run (long enough to have enough statistics to take decisions). Then, only the affected part of the run is discarded usually saving most of the data. Depending on the good run list, up to 80 pb^{-1} has been recovered from partial runs.

Appendix B

Breakdown of the Unfolding Factors

Although the Z/γ^* boson reconstruction and the jet selection cannot be completely factorized, one can still try to have an idea of what are the contributions fraction of detector effects related to the Z/γ^* reconstruction and the fraction related to the jets selection, the unfolding factors can be split in different steps (Section 4.6).

The first step (A) corresponds to the correction of going from the hadron level to the calorimeter level with the minimal set of cuts: kinematic cuts (E_{T}, p_{T} and y cuts on electrons and jets), invariant mass cut and vertex requirement:

$$\mathrm{A} = \frac{Z/\gamma^* \left[\mathrm{had}: E_{\mathrm{T}}^{e}, \eta^{e}, M_{ee}\right] + \text{particle jets}\left[p_{\mathrm{T}}^{\mathrm{jet}}, y^{\mathrm{jet}}\right]}{Z/\gamma^* \left[\mathrm{rec}: E_{\mathrm{T}}^{e}, \eta^{e}, M_{ee}, vtx\right] + \text{calorimeter jets}\left[orig.\ clus., p_{\mathrm{T,corr}}^{\mathrm{jet}}, y^{\mathrm{jet}}\right]} \tag{B.1}$$

Notice that our definition of jet only exists when we have a reconstructed Z/γ^* boson and the electron calorimeter clusters has been removed. In this first step, not all the cuts on the electrons have been applied, so we do not have a reconstructed Z/γ^* boson, and therefore, no jets. In this intermediate states of the unfolding, a collection of jets clustered in the original way (*orig. clus.*) without removing the electron calorimeter clusters is used. As a last step in the unfolding breakdown, the jets are reclustered after removing the electron calorimeter clusters (step D). Anyway, this is only required to look at the intermediate steps of the unfolding. When calculating the unfolding factors, Eq. 4.21 is used where the jets are always well defined and no *reclustering* is required.

The second step (B) shows the effect of applying the fiduciality cuts and the cuts on the tracks of the electrons. In the third step (C), the rest of the ID cuts on the electrons are applied. The last step (D) reflects the effects of removing the electron calorimeter clusters before clustering the jets. The total unfolding is A × B × C × D. Figure B.1 shows the unfolding factors of the p_T^{jet} distribution for every step. The major effect is from step A, that corrects the data to the particle level, with only kinematic requirements. Steps B and C, that add more requirements on the identification of the electrons have smaller effects. Finally, removing the electron clusters from the calorimeter before running the jet algorithms, has values below 1.0, meaning that it increases the cross section.

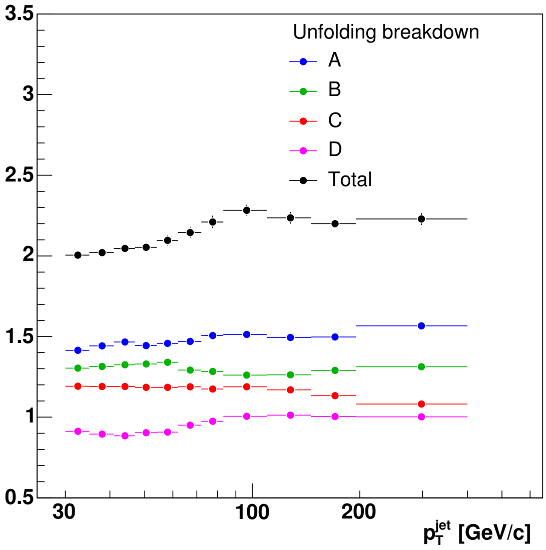


Figure B.1: Breakdown of the unfolding factors of the p_T^{jet} distribution for events with a Z/γ^* boson and at least one jet.

References

[1] D.J. Gross and F. Wilczec, “*Asymptotically Free Gauge Theories*”, Phys. Rev. D **8**, 3633 (1973)

[2] F. Abe *et al.* (CDF Collaboration), “*Properties of Jets in Z Boson Events from 1.8* TeV $p\overline{p}$ *Collisions*”, Phys. Rev. Lett. **77**, 448 (1996) [`arXiv:hep-ex/9603003`]

[3] F. Halzen and A.D. Martin, “*Quarks and Leptons: An Introductory Course in Modern Particle Physics*”, Wiley Text Books (1984)

[4] F. Englert and R. Brout, “*Broken Symmetry and the Mass of Gauge Vector Mesons*”, Phys. Rev. Lett. **13**, 321 (1964); G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, “*Global Conservation Laws and Massless Particles*”, Phys. Rev. Lett. **13**, 585 (1964)

[5] S.P. Martin, “*A Supersymmetry Primer*”, [`arXiv:hep-ph/9709356`] (2006)

[6] H.E. Haber, “*Recent Directions in Particle Physics*”, Proceedings of 1992 Theoretical Advanced Study Institute in Elementary Particle Physics, World Scientific (Singapore, 1993)

[7] H.P. Nilles, “*Supersymmetry, Supergravity and Particle Physics*”, Phys. Rept. **110**, 1 (1984)

[8] H. Fritzsch and M. Gell-Mann, “*Proceedings of the XVIth International Conference on High energy Physics*”, Batavia, IL, Vol. 2, p. 135 (1972); H. Fritzsch, M. Gell-Mann and H. Leutwyler, “*Advantages of the Color Octet Gluon Picture*”, Phys. Lett. B **47**, 365 (1973)

[9] R.K. Ellis, W.J. Stirling and B.R. Webber, “*QCD and Collider Physics*”, Cambridge University Press (1996)

[10] Y.M. Yao *et al.* (Particle Data Group), “*Review of Particle Physics*”, J. Phys. G **33**, 1 (2006) and 2007 partial update for edition 2008 (http://pdg.lbl.gov/)

[11] J.D. Bjorken, "*Asymptotic Sum Rules at Infinite Momentum*", Phys. Rev. **179**, 1547 (1969); E.D. Bloom *et al.* (SLAC-MIT Collaboration), "*High-Energy Inelastic e-p Scattering at 6° and 10°*", Phys. Rev. Lett. **23**, 930 (1969)

[12] H. Abramowicz *et al.* (CDHS Collaboration), "*Neitrino and Antineutrino Charged-Current Inclusive Scattering in Iron in the Energy Range* $20 < E_\nu < 300$ GeV",, Z. Phys. C **17**, 283 (1983)

[13] R. Brandelik *et al.* (TASSO Collaboration), "*Evidence for Planar Events in* e^+e^- *Annihilation at High-Energies*", Phys. Lett. B **86**, 243 (1979)

[14] V.N. Gribov and L.N. Lipatov, "e^+e^- *Pair Annihilation and Deep Inelastic ep Scattering in Perturbation Theory*", Sov. J. Nucl. Phys. **15**, 438 (1972); G. Altarelli and G. Parisi, "*Asymptotic Freedom in Parton Language*", Nucl. Phys. B **126**, 298 (1977); Y.L. Dokshitzer, "*Calculation of the Structure Functions for Deep Inelastic Scattering and* e^+e^- *Annihilation by Perturbation Theory in Quantum Chromodynamics*", Sov. Phys. JETP **46**, 641 (1977)

[15] Z. Nagy and D.E. Soper, "*General Subtraction Method for Numerical Calculation of One-loop QCD Matrix Elements*", J. High Energy Phys. **09**, 055 (2003); R.K. Ellis, D.A. Ross and A.E. Terrano, "*Calculation of Event-Shape Parameters in* e^+e^- *Annihilation*", Phys. Rev. Lett. **45**, 1226 (1980)

[16] J. Campbell and R.K. Ellis, "*Next-to-Leading Order Corrections to* $W + 2$ jets *and* $Z + 2$ jets *Production at Hadron Colliders*", Phys. Rev. D **65**, 113007 (2002), [`arXiv:hep-ph/0202176`]

[17] J. Pumplin, *et al.*, "*Uncertainties of Predictions from Parton Distribution Functions II. The Hessian Method*", Phys. Rev. D **65**, 014013 (2001), [`arXiv:hep-ph/0101032`]

[18] A Abulencia *et al.* (CDF Collaboration), "*Measurement of the Inclusive Jet Cross Section Using the* k_{T} *Algorithm in* $p\overline{p}$ *Collisions at* $\sqrt{s} = 1.96$ TeV *with the CDF II Detector*", Phys. Rev. D **75**, 092006 (2007), [`arXiv:hep-ex/0701051`]

[19] V.V. Sudakov, "*Vertex Parts at Very High Energies in Quantum Electrodynamics*", Sov. Phys. JETP **3**, 65 (1956)

[20] B. Melle and P. Nason, "*The Fragmentation Function for Heavy Quarks in QCD*", Nucl. Phys. B **361**, 626 (1991)

[21] B. Andersson, G. Gustafson, g. Ingelman and T. Sjöstrand, "*Parton Fragmentation and String Dynamics*", Phys. Rept. **97**, 31 (1983); T. Sjöstrand, "*Jet Fragmentation of Nearby Partons*", Nucl. Phys. B **248**, 469 (1984)

[22] B.R. Webber, "*Simulation of QCD Jets Including Soft Gluon Interference*", Nucl. Phys. B **238**, 492 (1984)

[23] D. Amati and G. Veneziano, "*Preconfinement as a Property of Perturbative QCD*", Phys. Lett. B **83**, 87 (1979)

[24] T. Sjöstrand *et al.*, "*High-energy-Physics Event Generation with* PYTHIA *6.1*", Comput. Phys. Commun. **135**, 238 (2001), [`arXiv:hep-ph/0010017`]

[25] T. Affolder *et al.* (CDF Collaboration), "*Charged Jet Evolution and the Underlying Event in Proton-Antiproton Collisions at 1.8* TeV", Phys. Rev. D **65**, 092002 (2002)

[26] R. Field, "*CDF Run II Monte-Carlo Tunes*", FERMILAB-PUB-06-408-E (2006)

[27] G Corcella *et al.*, "HERWIG *6.5 Release Note*", J. High Energy Phys. **01**, 010 (2001), [`arXiv:hep-ph/0210213`] (2002)

[28] J.M. Butterworth, J.R. Forshaw and M.H. Seymour, "*Multiparton Interactions in Photoproduction at HERA*", Z. Phys. C **72**, 637 (1996), [`arXiv:hep-ph/9601371`]

[29] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A.D. Polosa, "ALPGEN*, a Generator for Hard Multiparton Processes in Hadronic Collisions*", J. High Energy Phys. **0307**, 001 (2003), [`arXiv:hep-ph/0206293`]

[30] F. Caravaglios, M.L. Mangano, M. Moretti and R. Pittau, "*A New Approach to Multi-jet Calculations in Hadron Collisions*", Nucl. Phys. B **539**, 215 (1999), [`arXiv:hep-ph/9807570`]

[31] M.L. Mangano, "*The So-Called* MLM *Matching*", Fermilab ME/MC Tunning Workshop (October 2002)

[32] D0 Collaboration, "*Z + Jet Production in the D0 Experiment: a Comparison Between Data and the* PYTHIA *and* SHERPA *Monte Carlos*", D0 Public Note 5066-CONF (2006)

[33] S. Catani, F. Krauss, R. Kuhn, B.R. Webber, "*QCD Matrix Elements + Parton Showers*", J. High Energy Phys. **0111**, 063 (2001), [`arXiv:hep-ph/0109231`]

[34] S. Catania, Yu.L. Dokshitzerc, M. Olssond, G. Turnocka and B.R. Webbera, "*New Clustering Algorithm for Multijet Cross Sections in* e^+e^- *Annihilation*", Phys. Lett. B **269**, 432 (1991)

[35] J.E. Huth *et al.*, "*Toward a Standardization of Jet Definitions.*", in Snowmass Summer Study, p. 134, World Scientific (1992), FERMILAB-CONF-90-249-E

[36] G.C. Blazey *et al.*, "*Run II Jet Physics*", [`arXiv:hep-ex/0005012`] (2000) ; S.D. Ellis, J. Huston and M. Toennesmann, "*On Building Better Cone Jet Algorithms*", [`arXiv:hep-ph/0111434`] (2001)

[37] S.D. Ellis and D.E. Soper, "*Successive Combination Jet Algorithm for Hadron Collisions*", Phys. Rev. D **48**, 3160 (1993)

[38] CDF Collaboration, "*CDF Run II: Technical Design Report*", FERMILAB-PUB-96-390-E (1996)

[39] A. Sill, "*CDF Run II Silicon Tracking Projects*", Nucl. Instrum. Meth. A **447**, 1 (2000)

[40] T. Affolder *et al.*, "*Central Outer Tracker*", Nucl. Instrum. Meth. A **526**, 249 (2004)

[41] D. Acosta *et al.*, "*A Time-Of-Flight Detector in CDF-II.*", Nucl. Instrum. Meth. A **518**, 605 (2004)

[42] L. Balka *et al.*, "*The CDF Central Electromagnetic Calorimeter*", Nucl. Instrum. Meth. A **267**, 272 (1988), FERMILAB-PUB-87-172-E

[43] S. Bertolucci *et al.*, "*The CDF Central and Endwall Hadron Calorimeter*", Nucl. Instrum. Meth. A **267**, 301 (1988), FERMILAB-PUB-87-174-E

[44] M. Albrow *et al.*, "*The CDF Plug Upgrade Electromagnetic Calorimeter: Test Beam Results*", Nucl. Instrum. Meth. A **480**, 524 (2002), FERMILAB-PUB-01-045-E

[45] S. Lusin, "*Calibration and Testing of the CDF II Endplug Calorimeter*", FERMILAB-CONF-98-003-E (1997)

[46] G. Apollinari *et al.*, "*Shower Maximum Detector for the CDF Plug Upgrade Calorimeter*", Nucl. Instrum. Meth. A **412**, 512 (1998)

[47] F. Abe *et al.*, "*Precision Measurement of the Prompt Photon Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*", Phys. Rev. Lett. **73**, 2662 (1994)

[48] M. Albrow *et al.*, "*A Preshower Detector for the CDF Plug Upgrade: Test Beam Results*", Nucl. Instrum. Meth. A **431**, 104 (1999)

[49] C.M. Ginsburg, "*CDF Run 2 Muon System*", Eur. Phys. J. C **33**, S1002 (2004), FERMILAB-CONF-03-386-E

[50] I. Vila, "*Performance and First Physics Results of the SVT Trigger at CDF II*", [`arXiv:hep-ph/0307165`] (2003)

[51] D. Acosta *et al.*, "*The Performance of the CDF Luminosity Monitor*", Nucl. Instrum. Meth. A **494**, 57 (2002)

[52] S. Klimenko, J. Konigsberg and T. Liss, "*Averaging of the Inelastic Cross Sections Measured by the CDF and the E811 Experiments*", FERMILAB-FN-0741 (2003)

[53] A. Bhatti *et al.*, "*Determination of the Jet Energy Scale at the Collider Detector at Fermilab*", Nucl. Instrum. Meth. A **566**, 375 (2006), [`arXiv:hep-ex/0510047`]

[54] H.L. Lai *et al.*, "*Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions*", Eur. Phys. J. C **12**, 375 (2000), [`arXiv:hep-ph/9903282`]

[55] A Abulencia *et al.* (CDF Collaboration), "*Measurements of Inclusive W and Z Cross Sections in $p\overline{p}$ Collisions at* $\sqrt{s} = 1.96$ TeV", J. Phys. G **34**, 2457 (2007), [`arXiv:hep-ex/0508029`] ; D. Acosta *et al.* (CDF Collaboration), "*First Measurements of Inclusive W and Z Cross Sections from Run II of the Fermilab Tevatron Collider*", Phys. Rev. Lett. **94**, 091803 (2005), [`arXiv:hep-ex/0406078`]

[56] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, "*Uncertainties of Predictions from Parton Distributions II: Theoretical Errors*", Eur. Phys. J. C **35**, 325 (2004)

[57] D. Acosta *et al.* (CDF Collaboration), "*Study of Jet Shapes in Inclusive Jet Production in $p\overline{p}$ Collisions at* $\sqrt{s} = 1.96$ TeV", Phys. Rev. D **71**, 112002 (2005)

[58] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky and W.K. Tung, "*New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*", J. High Energy Phys. **0207**, 012 (2002), [`arXiv:hep-ph/0201195`]

[59] W. Wagner *et al.*, "*Online Monitoring in the CDF II Experiment*", International Europhysics Conference on High-Energy Physics (HEP 2001), FERMILAB-CONF-02-269-E (2001)

List of Figures

2.1 Value of the running strong coupling α_s as a function of the energy scale. . . 6

2.2 Schematic view of a deep inelastic scattering event: the photon (or a Z boson) interacts with a parton inside the proton. 7

2.3 Kinematic domains in x and Q^2 probed by fixed-target and collider experiments shown together with the important constrains they make on the various parton distribution functions. 8

2.4 Example of proton PDFs measured at $Q^2 = 20$ GeV^2 and $Q^2 = 10,000$ GeV^2 in a DIS experiment for gluons and quarks. The contributions from the gluons and the quarks from the sea increase with Q^2. PDFs measured in DIS experiments are used in $p\overline{p}$ colliders. 9

2.5 Three-jet event with a gluon emission. 9

2.6 Schematic hadron-hadron interaction. 10

2.7 Feynman diagrams of the lowest order processes of the splitting functions (Eq. 2.10): gluon radiation, quark pair production and gluon splitting. 10

2.8 Structure function F_2 as a function of Q^2 and x measured in DIS and fixed target experiments. The curves are NLO pQCD fits. 12

2.9 Feynman diagrams for $Z/\gamma^* + 1$ jet production at LO in a hadron collider. . . 14

2.10 Some of the Feynman diagrams for $Z/\gamma^* + 2$ jets production at LO in a hadron collider. 14

2.11 Some of the Feynman diagrams contributing to virtual NLO corrections for $Z/\gamma^* + 1$ jet production in a hadron collider. 14

2.12 Scale dependence of the $Z/\gamma^*(\rightarrow e^+e^-) +$ jets cross section predictions with the factorization and renormalization scales set equal and given by μ. 16

2.13 Uncertainties on up quark (left) and gluon (right) PDFs. The green bands represent the global uncertainty at $Q = 10$ GeV. 17

2.14 Parton shower with schematic string (left) and cluster (right) hadronization models. 19

2.15 Branching history construction using the $k_\perp$ algorithm. 23

2.16 Illustration of infrared sensitivity in jet clustering. The presence of soft radiation between the two jets may cause the merging of the jets that would not occur in the absence of the soft radiation. 25

2.17 Illustration of collinear sensitivity in jet reconstruction. The configuration in the fails to produce a seed because its energy is split among several detector towers. 25

3.1 Chain of accelerators at Fermilab (not to scale) with the Tevatron and the CDF II detector. 28

3.2 Tevatron Run II peak instantaneous luminosities since 2001 for every store and averaged every 20 stores. 29

3.3 Tevatron Collider Run II delivered integrated luminosity and integrated luminosity recorded by CDF. 30

3.4 Isometric view of the CDF Run II detector with its main subsystems: the Silicon Vertex Detector, the Central Outer Tracker, the superconducting solenoid, the electromagnetic calorimeter, the hadronic calorimeter and the muon chambers. 30

3.5 Longitudinal view of the CDF Run II detector, showing the coverage in pseudorapidity of the tracking systems. 31

3.6 Layout of three COT supercells. 32

3.7 Elevation view of the CDF Run II detector calorimeters. 34

3.8 General diagram of CDF Run II trigger and data acquisition systems. The figure corresponds to the original design of the system conceived to run at 132 ns clock cycles that was the expected bunch spacing at the Tevatron. . . 37

3.9 Diagram of Level 1 and Level 2 trigger systems and data flow between them. 38

3.10 Schematic of a single CDF calorimeter tower. θ_{EM} is the angle from the interaction point to the position of the maximum of a typical electromagnetic shower developed in the tower. θ_{HAD} is the angle from the interaction point to the position of the maximum of a typical hadronic shower from a pion developed in the tower. 42

3.11 Dijet balance $\beta_{\mathrm{dijet}} = p_{\mathrm{T}}{}^{\mathrm{probe}}/p_{\mathrm{T}}{}^{\mathrm{trigger}}$ as a function of η^{jet} for data, PYTHIA and HERWIG MC samples in different $p_{\mathrm{T}}^{\mathrm{jet}}$ bins. The dijet balance shows the non-homogeneous response of the calorimeter and the gaps at $\eta = 0$ and $|\eta| = 1.1$. 44

3.12 Mean E_{T} in a random cone of $R = 0.7$ as a function of the number of reconstructed vertices. 45

3.13 Difference between the p_{T} of the particle jet and the p_{T} of the jet reconstructed in the calorimeter for different bins in $p_{\mathrm{T}}^{\mathrm{jet}}$. 46

4.1 ΔR between the two electrons as function of the p_{T} of the leading jet. 54

4.2 Electron ID efficiency as a function of the p_{T} of the jet in events with at least one jet, in the data and in the Monte Carlo for tight, loose and plug electrons. Scale factor Data/MC for tight, loose and plug electrons. The dashed lines correspond to an uncertainty of 5% for the tight electrons, 3% for the loose electrons and 7% for the plug electrons. 57

4.3 Z/γ^* boson ID efficiency in CC and CF cases, as a function of the p_{T} of the jet in events with at least one jet, in the data and in the Monte Carlo. Scale factor Data/MC for CC and CF events. The dashed lines correspond to an uncertainty of 5% on both cases. 58

4.4 Z/γ^* invariant mass for the inclusive selection. 60

4.5 Detector level cross section not corrected for inefficiencies and without background subtraction, using only the first 368 pb^{-1} of data, for jets clustered with the electron towers and jet clustered after removing the electron towers from the calorimeter. The maximum difference is 10% at low $p_{\mathrm{T}}^{\mathrm{jet}}$. 61

4.6 Number of $Z/\gamma^*(\rightarrow e^+e^-) + \geq N_{\mathrm{jets}}$ events in different instantaneous luminosity ranges, as a function of the jet multiplicity. 62

4.7 Number of jets per $Z/\gamma^* \rightarrow e^+e^-$ event as a function of instantaneous luminosity. A fit to a constant gives $\chi^2/\mathrm{ndf} = 68.1/67$. 63

4.8 Proportion of $Z/\gamma^* \rightarrow e^+e^- + 2$ jets and $Z/\gamma^* \rightarrow e^+e^- + 3$ jets events over $Z/\gamma^* \rightarrow e^+e^- + 1$ jet events as a function of instantaneous luminosity. The fit to a constant gives $\chi^2/\mathrm{ndf} = 23.1/35$. 64

4.9 Fake rates f_e^{jet} for the different types of electron selection cuts. Scale factor ξ between $E_{\mathrm{T}}^{\mathrm{jet}}$ and the E_{T} of the electron that it fakes, fitted with a Gaussian distribution. 65

4.10 Data and signal plus background estimation in the Z mass window and the side bands, in $Z/\gamma^*(\rightarrow e^+e^-) + \geq 1$ jet and $Z/\gamma^*(\rightarrow e^+e^-) + \geq 2$ jets events. . 68

4.11 Detector level cross section for the CC selection in data and PYTHIA-TUNE A Monte Carlo and their ratio Data/MC before reweighting in $\hat{p}_{\mathrm{T}}$. The curve represents the fit to a second order polynomial from which the weights are determined. The dashed lines correspond to the errors in the fit. The same cross sections after the reweighting in $\hat{p}_{\mathrm{T}}$ of PYTHIA Monte Carlo events. . . 71

4.12 Ratio of the forward electron E_{T} Data/MC(PYTHIA) before and after the reweighting in $E_{\mathrm{T}}^{e,\mathrm{fwd}}$ of the Monte Carlo sample. 72

4.13 Unfolding factors of the $p_{\mathrm{T}}^{\mathrm{jet}}$ and y^{jet} distributions in $Z/\gamma^*(\to e^+e^-)+\geq 1$ jet and $Z/\gamma^*(\to e^+e^-)+\geq 2$ jets events. 73

4.14 Unfolding factors for the $p_{\mathrm{T}}^{\mathrm{jet}}$ distribution in $Z/\gamma^*(\to e^+e^-)+\geq 1$ jet events using non-reweighted PYTHIA-TUNE A, PYTHIA-TUNE A reweighted only on $\hat{p}_{\mathrm{T}}$, varying up and down the weighting function, and reweighted on both $\hat{p}_{\mathrm{T}}$ and $E_{\mathrm{T}}^{e,\mathrm{fwd}}$. The dashed lines correspond to the assigned uncertainty of 1%. . 75

4.15 Total systematic uncertainties on the $p_{\mathrm{T}}^{\mathrm{jet}}$ and y^{jet} inclusive jet differential cross sections in $Z/\gamma^*(\to e^+e^-)+\geq 1$ jet and $Z/\gamma^*(\to e^+e^-)+\geq 2$ jets production. 77

5.1 Illustrations of the differential and integrated jet shapes. 80

5.2 Differential and integrated jet shapes measured at the detector level in data and PYTHIA-TUNE A Monte Carlo for jets with $|y^{\mathrm{jet}}| < 0.7$ in $Z/\gamma^*(\to e^+e^-)+$jets events with exactly one reconstructed vertex. 81

5.3 Unfolding factors for the differential and integrated jet shapes obtained with PYTHIA-TUNE A Monte Carlo. 82

5.4 Differential and integrated jet shapes measurement for jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 30$ GeV/c and $|y^{\mathrm{jet}}| < 0.7$ in $Z/\gamma^*(\to e^+e^-)+$jets events. First tick in the error bars correspond to the statistical uncertainty, the second tick mark correspond to the total (statistical $\oplus$ systematic) uncertainty. Results are compared to the predictions from PYTHIA-TUNE A. 83

5.5 Differential and integrated jet shapes for jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 30$ GeV/c and $|y^{\mathrm{jet}}| < 0.7$ in $Z/\gamma^*(\to e^+e^-)+$jets events compared to the predictions from PYTHIA. 84

5.6 Differential and integrated jet shapes for jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 30$ GeV/c and $|y^{\mathrm{jet}}| < 0.7$ in $Z/\gamma^*(\to e^+e^-)+$jets events compared to the predictions from HERWIG and HERWIG interfaced with JIMMY. 85

5.7 Diagram of the energy flow definition in $Z/\gamma^*(\to e^+e^-)+$jets. 86

5.8 Energy flow of the central calorimeter towers in $Z/\gamma^*(\rightarrow e^+e^-) + \text{jets}$ events compared to the predictions from PYTHIA-TUNE A and TUNE DW. Energy flow in events with $p_T^Z < 50$ GeV/c and events with $p_T^Z > 50$ GeV/c. Only statistical uncertainties are included in the plots. 87

6.1 C_{UE} and $C_{\text{hadronization}}$ corrections for $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet cross section. Global C_{HAD} factors determined using PYTHIATUNE A and TUNE DW. The nominal correction is TUNE A and the yellow band indicates the assigned systematic uncertainty that is the symmetrized difference between TUNE A and TUNE DW corrections. 91

6.2 $C_{\text{HAD}}(|y^{\text{jet}}|)$ corrections for $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet and 2 jets cross sections. The correction factors determined using PYTHIATUNE A and TUNE DW. The yellow band indicates the assigned systematic uncertainty that is the symmetrized difference between TUNE A, the nominal factors, and TUNE DW. . 92

6.3 Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet production as a function of p_T^{jet} compared to the NLO pQCD prediction. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. Data/theory ratio with the PDF uncertainty and the variation of the theoretical predictions with the scale μ. C_{HAD} corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty. 95

6.4 Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 2$ jets production as a function of p_T^{jet} compared to the NLO pQCD prediction. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. Data/theory ratio with the PDF uncertainty and the variation of the theoretical predictions with the scale μ. C_{HAD} corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty. 96

6.5 Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq 1$ jet production as a function of $|y^{\text{jet}}|$ compared to the NLO pQCD prediction. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. Data/theory ratio with the PDF uncertainty and the variation of the theoretical predictions with the scale μ. C_{HAD} corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty. 97

6.6 Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production as a function of $|y^{\rm jet}|$ compared to the NLO pQCD prediction. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. Data/theory ratio with the PDF uncertainty and the variation of the theoretical predictions with the scale μ. $C_{\rm HAD}$ corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty. 98

6.7 Measured differential cross section as a function of the $p_{\rm T}^{\rm jet}$ of the first leading jet in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 1$ jet production compared to the NLO pQCD prediction. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. Data/theory ratio with the PDF uncertainty and the variation of the theoretical predictions with the scale μ. $C_{\rm HAD}$ corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty. 99

6.8 Measured differential cross section as a function of the $p_{\rm T}^{\rm jet}$ of the second leading jet in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production compared to the NLO pQCD prediction. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. Data/theory ratio with the PDF uncertainty and the variation of the theoretical predictions with the scale μ. $C_{\rm HAD}$ corrections that have been applied to the NLO pQCD predictions with their systematic uncertainty. 100

6.9 Measured total event cross section in $Z/\gamma^*(\rightarrow e^+e^-)+$ jets events as a function of inclusive jet multiplicity $N_{\rm jets}$ compared to LO and NLO pQCD predictions. The yellow bands show the total systematic uncertainty, except for the 5.8% uncertainty on the luminosity. Data/LO and data/LO ratios with the NLO PDF uncertainty and the variation of the NLO predictions with the scale μ. . 101

6.10 Measured total event cross section in inclusive $Z/\gamma^*(\rightarrow e^+e^-)+$ jets as a function of jet multiplicity $N_{\rm jets}$ compared to PYTHIA predictions using the TUNE A and TUNE DW underlying event settings, and to alpgen v2 predictions matched to the parton shower from PYTHIA using the MLM matching technique. 106

6.11 Measured differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production as a function of $\Delta R_{\rm jj}$ compared to PYTHIA predictions using the TUNE A and TUNE DW underlying event settings, and to ALPGEN v2 predictions (line) matched to the parton shower from PYTHIA using the MLM matching technique. 107

6.12 Measured differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production as a function of $\Delta\phi_{\mathrm{jj}}$ compared to PYTHIA predictions using the TUNE A and TUNE DW underlying event settings, and to ALPGEN v2 predictions matched to the parton shower from PYTHIA using the MLM matching technique. 108

6.13 Measured differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+\geq 2$ jets production as a function of M_{jj} compared to PYTHIA predictions, and to ALPGEN v2 predictions matched to the parton shower from PYTHIA using the MLM matching technique. 109

A.1 DQMon panel displays the status of some monitors and the results of the checks performed in the data from several subsystems. 115

B.1 Breakdown of the unfolding factors of the $p_{\mathrm{T}}^{\mathrm{jet}}$ distribution for events with a Z/γ^* boson and at least one jet. 118

List of Tables

I Gauge bosons of the Standard Model, mediators of the fundamental interactions. 3

II Elementary fermions of the Standard Model and their electric charge. . . . 4

III Summary of the characteristics of the CDF Run II calorimeters. 36

IV Selection applied to every level for the high-p_{T} central electron trigger. . . . 48

V Summary of the trigger efficiency values and uncertainties per electron and per event. See Section 4.7 for the determination of the systematic uncertainties. 51

VI Energy scale factors for central and forward electrons. 52

VII List of electron selection cuts. Fiduciality cuts ensure that the energy is measured in a well instrumented region of the calorimeter. 53

VIII Efficiencies in data and Monte Carlo for $Z/\gamma^{*}+\geq 1$, 2 and 3 jets, and the relative scale factor $\kappa^{\mathrm{MC}}_{\mathrm{Data}}$. 56

IX Energy scale factors for electrons faked by jets, $\xi_{e}^{\mathrm{jet}} = E_{\mathrm{T}}^{e}/E_{\mathrm{T}}^{\mathrm{jet}}$. 66

X Number of events found in 1.73 fb^{-1} of data compared to the expected background and $Z/\gamma^{*}(\rightarrow e^{+}e^{-})+\mathrm{jets}$ events. 67

XI Number of events in data and estimated signal plus background, in the Z mass window ($66-116$ GeV/c^{2}), the lower band of the spectrum ($46-66$ GeV/c^{2}) and the higher band ($116-136$ GeV/c^{2}). 68

XII Weights applied to Pythia Monte Carlo samples. The listed factors correspond to Eq. 4.23. 70

XIII Systematic uncertainties on the cross section measurements. 76

XIV Systematic uncertainties on the jet shapes measurement. 83

XV Uncertainties and dependencies of the NLO pQCD cross sections. 92

XVI Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq N_{\text{jets}}$ production as a function of $p_{\text{T}}^{\text{jet}}$ with $N_{\text{jets}} \geq 1$ and $N_{\text{jets}} \geq 2$. The systematic uncertainties are fully correlated across $p_{\text{T}}^{\text{jet}}$ bins. The parton-to-hadron correction factors $C_{\text{HAD}}(p_{\text{T}}^{\text{jet}},N_{\text{jets}})$ are applied to the pQCD predictions. . . 102

XVII Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq N_{\text{jets}}$ production as a function of the $p_{\text{T}}^{\text{jet}}$ of the N^{th} leading jet with $N_{\text{jets}} \geq 1$ and $N_{\text{jets}} \geq 2$. The systematic uncertainties are fully correlated across $p_{\text{T}}^{\text{jet}}$ bins. The parton-to-hadron correction factors $C_{\text{HAD}}(p_{\text{T}}^{\text{jet}},N_{\text{jets}})$ are applied to the pQCD predictions. 103

XVIII Measured inclusive jet differential cross section in $Z/\gamma^*(\rightarrow e^+e^-)+ \geq N_{\text{jets}}$ production as a function of $|y^{\text{jet}}|$ with $N_{\text{jets}} \geq 1$ and $N_{\text{jets}} \geq 2$. The systematic uncertainties are fully correlated across y^{jet} bins. The parton-to-hadron correction factors $C_{\text{HAD}}(y^{\text{jet}},N_{\text{jets}})$ are applied to the pQCD predictions. . . 104

XIX Measured total event cross section in inclusive $Z/\gamma^*(\rightarrow e^+e^-)+ \geq N_{\text{jets}}$ production for $N_{\text{jets}} \geq 1$, $N_{\text{jets}} \geq 2$ and $N_{\text{jets}} \geq 3$. The systematic uncertainties are fully correlated across N_{jets} bins. The parton-to-hadron correction factors $C_{\text{HAD}}(N_{\text{jets}})$ are applied to the pQCD predictions. 104