



Measurement of $Z/\gamma^* + b$ -jet Production Cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF detector

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

July 2012

supervised by
Dr. Verónica Sorin
(IFAE)

Tutor:
Prof. Mario Martínez Pérez
(ICREA/IFAE and UAB)

Contents

1	Introduction	1
2	QCD Theory	3
2.1	The QCD Lagrangian	3
2.1.1	Lattice QCD and perturbative approximation	5
2.1.2	Renormalization and running coupling constant	6
2.2	A typical hadron collision and the factorization theorem	7
2.3	The Initial State: PDFs and their evolution	9
2.4	The final state evolution and hard scattering	17
2.5	The hadronization	19
2.5.1	The cluster model	20
2.5.2	The string and Lund Model	20
2.5.3	Underlying event (UE)	21
2.5.4	Multiple Interactions	23
2.6	The predictive tools MC	23
2.6.1	LO event generators	24
2.6.2	NLO predictions	25
2.6.3	NNLO predictions	25
2.6.4	Parton Shower	25
2.6.5	MEPS	28
2.7	The Jet production	29
2.7.1	The Jet Definition	30
2.7.2	Cone Algorithm	32
2.7.3	Clustering Algorithm	32
3	Z + b Theoretical predictions	35
3.1	Introduction: the 4FNS and the 5FNS scheme	35
3.2	NLO Associated production of a Z boson and a Single Heavy quark Jet	36
3.3	Z+ 2 jets with one b-quark jet	37
3.4	Z + two high p_T b-quark jet	38
4	The Experimental Environment	41
4.1	The Fermilab Tevatron collider	41
4.1.1	Proton and Antiproton production	43
4.1.2	Injection and collisions	44
4.1.3	Tevatron performance	44
4.2	The CDFII Detector	44
4.2.1	The CDFII Coordinate system	46
4.2.2	The tracking system	47

Contents

4.2.3	Central Outer tracker	51
4.3	Track reconstruction	51
4.4	Time of Flight detector	54
4.5	Calorimeters	55
4.5.1	Central region: CEM, CHA, WHA	57
4.5.2	Forward region: PEM and PHA	57
4.6	Muons systems	58
4.7	The Cerenkov Luminosity counters	58
4.8	Trigger and data acquisition system	59
4.8.1	Level 1	60
4.8.2	Level-2	61
4.8.3	Level 3	62
4.8.4	Operation and data quality	62
5	Physics Objects Reconstruction	65
5.1	Primary Vertex	65
5.1.1	z primary vertex	65
5.1.2	3-D Primary Vertex	65
5.2	Electrons	66
5.3	Muons	68
5.4	Jets	69
5.4.1	The CDF MidPoint Jet clustering Algorithm	69
5.4.2	Jet Corrections	70
5.4.3	b-tagging : bottom jet identification	75
6	The $Z/\gamma^* + b$-jet cross section measurement	79
6.1	Measurement definition	79
6.2	Definition of the dataset	80
6.3	Monte Carlo Samples	81
6.3.1	Alpgen+Pythia $Z + jets$ MC	81
6.3.2	Pythia MC	81
6.4	Event Selection	82
6.4.1	Muon identification and $Z/\gamma^* \rightarrow \mu^+\mu^-$ reconstruction	82
6.4.2	Electron identification and $Z/\gamma^* \rightarrow e^+e^-$ reconstruction	87
6.4.3	Jet selection	89
6.4.4	Event selection Summary	93
6.5	Background Modeling	93
6.5.1	Diboson and $t\bar{t}$	94
6.5.2	QCD and $W + jets$ backgrounds	94
6.6	Pretag Sample: Data - Monte Carlo Comparison	96
6.7	Composition of b-tagged sample	96
6.7.1	Fitting procedure	98
6.7.2	Bias Checks	102
6.7.3	Background Expectation	103
6.8	Acceptance	106
6.9	Systematics	106

6.9.1	Vertex Mass shape systematics	107
6.9.2	Bootstrap	109
6.9.3	Acceptance systematics	111
6.9.4	Background subtraction systematics	111
7	$Z/\gamma^* + b$-jet Differential Cross Sections	113
7.1	Measurement definition	113
7.2	Sample composition	113
7.3	Unfolding	114
7.4	Systematic Uncertainties	117
8	Results	123
8.1	Theory predictions	123
8.1.1	pQCD calculation	123
8.1.2	Non pQCD corrections	123
8.1.3	Theoretical Uncertainties	124
8.2	Comparison with theoretical predictions	125
8.2.1	Integrated $Z/\gamma^* + b$ -jet production cross section	126
8.2.2	Differential cross section	126
8.3	Comparison with LO ME+PS predictions	126
9	Conclusions	131

1 Introduction

Processes at hadron colliders, such as the production of jets, are described by the Quantum Chromodynamics theory (QCD). Precise descriptions of processes involving jets in association with a vector boson have nowadays large relevance as they represent irreducible background to other Standard Model (SM) processes and searches for new physics.

The experimental study and understanding of the b -jet production in association with a Z boson are crucial for many reasons. For one side, it is the most important background for a light Higgs boson [1] decaying into a bottom-antibottom quark pair and produced in the ZH mode. This is one of the most promising channels for the Higgs search at Tevatron in particular since the latest results [2] have excluded the high mass region ($M_H \geq 127 \text{ GeV}/c^2$). For another side the signature of b -jets and a Z boson is also background to new physics searches, such as supersymmetry, where a large coupling of the Higgs boson to bottom quarks is allowed [3].

The production cross section measurement of b -jets in events with a Z boson has already been performed at hadron colliders, at the Tevatron by CDF [4] and D0 experiments [5] and are now pursued at the LHC by ATLAS [6] and CMS [7]. In particular the CDF measurement was performed with only 2 fb^{-1} and was limited by the statistical uncertainty.

This PhD thesis presents a new measurement of the $Z/\gamma^* + b$ -jet production cross section using the complete dataset collected by CDF during the Run II.

Z/γ^* bosons are selected in the electron and muon decay modes and are required to have $66 \leq M_Z \leq 116 \text{ GeV}/c^2$ while jets, reconstructed with the MidPoint algorithm, have to be central ($|Y| \leq 1.5$) with $p_T \geq 20 \text{ GeV}/c$. The per jet cross section is measured with respect to the Z/γ^* inclusive and the $Z/\gamma^* + \text{jets}$ cross sections. Results are compared to leading order (LO) event generator plus parton shower and next-to-leading order (NLO) predictions corrected for non perturbative effects such as hadronization and underlying event. Differential distributions as a function of jet transverse moment and jet rapidity are also presented together with the comparison to NLO pQCD predictions for different renormalization and factorization scales and various PDF sets.

Chapter 2 and Chapter 3 are dedicated to explain the main features of the theory of Quantum Chromodynamics, to provide a description of the predictive Monte Carlo tools used in experimental context to simulate signal and background and to review the challenges of the calculation of the $Z + b$ -jet production processes. The Tevatron collider and the CDF experiment are described in Chapter 4. The procedure followed at CDF for the reconstruction of physics objects is treated with particular attention to the b -jets identification technique in Chapter 5. The analysis strategy for the integrated cross section is well discussed in Chapter 6, while the methodology for the differential cross section measurements as a function of jet p_T and jet rapidity are presented in Chapter 7. Finally, results are reported in Chapter 8 including the comparison to different theoretical predictions. Chapter 9 is devoted to the

summary and the conclusions.

2 QCD Theory

In this chapter the theory of Quantum Chromodynamics in a phenomenological point of view is presented. Firstly the corresponding lagrangian is introduced, and later a detailed description of the techniques, used to compute it, is given. In the last part of the chapter is explained how these methods are implemented in simulation Monte Carlo programs to perform cross section prediction. An important QCD signature is the production of collimated jets of hadrons and since the aim of this thesis is a jet production cross section also the theoretical issues concerning the definition of a jet is treated widely.

2.1 The QCD Lagrangian

Quantum ChromoDynamics (QCD) [10] is the Standard Model (SM) theory that describes the strong interaction, which is one of the four fundamental forces in Nature. The strong interaction is responsible for binding together quarks and gluons to form hadrons, among which the proton and the neutron are the most well-known examples.

According to QCD hadrons are made up of quarks of different flavors. Gluons are the bosons that mediate the strong interaction. Each flavor comes with a color and gluons carry one color and one anti-color index. The coupling strength of gluons and quarks interaction is g . Two important features of QCD are color confinement and asymptotic freedom. Color confinement implies that gluons and quarks cannot be observed as free particles at large distances and they are confined in bound states (hadrons) (Figure 2.1). Asymptotic freedom means that strong interactions become large at low energy and smaller at high energy.

In quantum field theory QCD is expressed by the Lagrangian density for the strong interaction:

$$L = \bar{\psi}_q^i (i\gamma^\mu) (D_\mu)_{ij} \psi_q^j - m_q \bar{\psi}_q^i \psi_q^j - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} \quad (2.1)$$

where ψ_q^i represents a quark field with color index i , γ^μ is a Dirac matrix with μ being a Lorentz vector index, m_q the mass of the quark, $F_{\mu\nu}^a$ is the gluon field strength tensor for a gluon with color index a (in the adjoint representation, $a \in [1, \dots, 8]$).

D_μ is the covariant derivative in QCD

$$(D_\mu)_{ij} = \delta_{ij} \partial_\mu - i g_s t_{ij}^a A_\mu^a \quad (2.2)$$

with g_s the strong coupling ($g_s^2 = 4\pi\alpha_s$), A_μ^a is the gluon field with color index a and t_{ij}^a are the generators of the $SU(3)$ symmetry¹.

¹ they are proportional to the hermitian and traceless Gell-Mann matrices ($t_{ij}^a = \frac{1}{2}\lambda_{ij}^a$)

2 QCD Theory

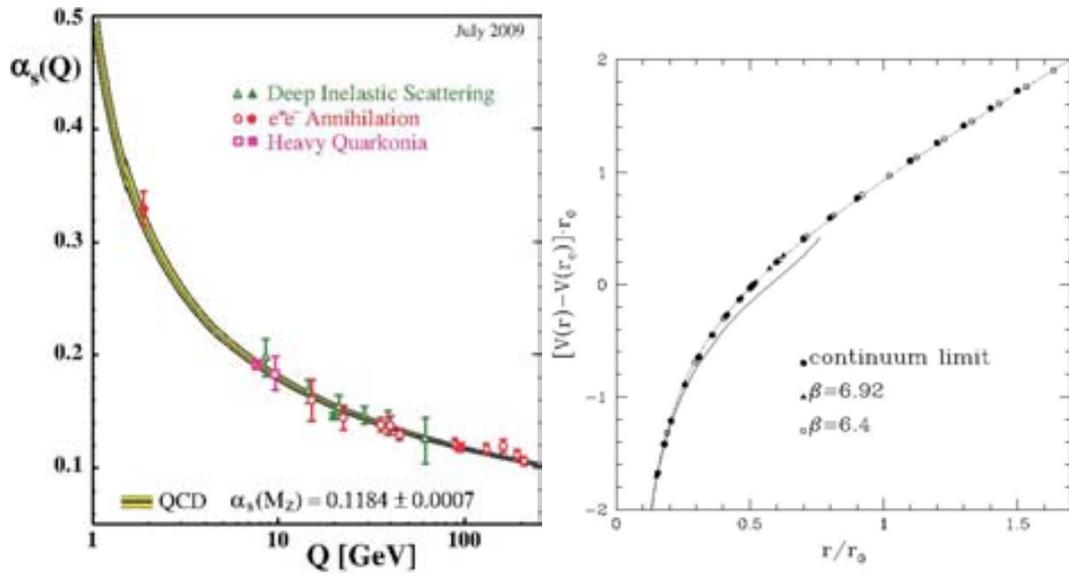


Figure 2.1: Figure that represents the asymptotic freedom and the confinement. Figures taken from [8] and [9]

In fact, QCD is a theory based on the gauge group $SU(3)$. Quarks are in the fundamental representation of the $SU(3)$ color group while gluons are in the adjoint representation. It is worth noting that the gluon field acts on quark color, taking away one color and replacing it with another, as it is shown in color flow diagram in Figure 2.2.

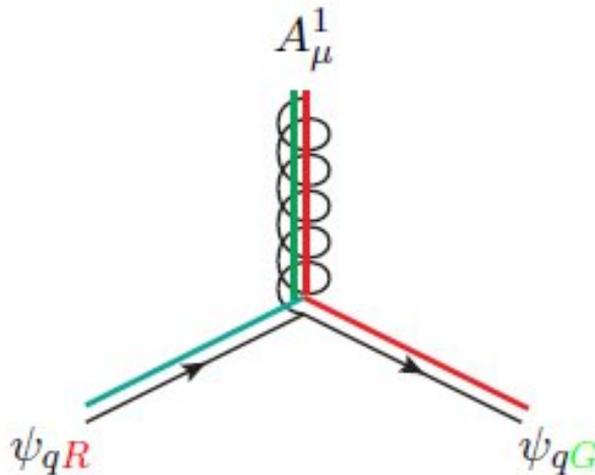


Figure 2.2: Color flow of a qqg vertex in QCD.

Moreover $F_{\mu\nu}^a$ is defined as:

$$F_{\mu\nu}^a = [\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - gf^{abc}A_\mu^b A_\nu^c] \quad (2.3)$$

where the third non-abelian term is due to $SU(3)$ being non-abelian. Self interaction is given by the corresponding term in $F_{\mu\nu}^a$ in the action. This term was introduced in order to be gauge invariant under local $SU(3)$ transformations and several experimental confirmations² of it have been obtained and it is a logical generalization of $F_{\mu\nu}$ in QED to $SU(N)$ group.

2.1.1 Lattice QCD and perturbative approximation

There are two main techniques used to solve QCD and to calculate its predictions: lattice QCD [14] and perturbative QCD [13].

Lattice QCD (Figure 2.3) is formulated on a grid of points in discrete space time, introducing a cut-off at the order of $1/a$, where a is the lattice spacing, which regularizes the theory. As a result lattice QCD is mathematically well-defined. Fields representing quarks are defined at lattice sites while gluons fields are defined on the link connecting close sites. The observables are determined using numerical simulation done with Monte Carlo algorithms.

This had a great success in predicting the hadron mass and it is a very precise method even though it presents an inconvenience: the huge computational time needed to extract the solution. Therefore, it is not suited for treating the complexity of high multiplicity events such as those produced at hadron colliders.

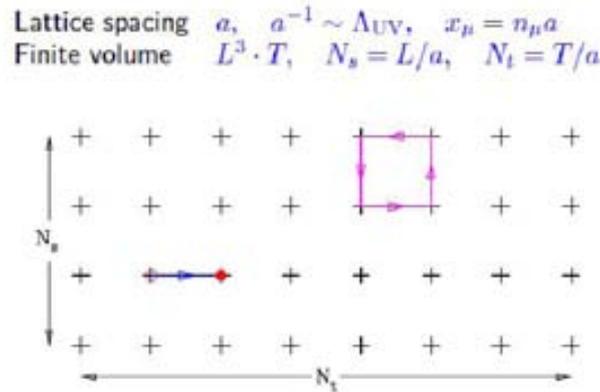


Figure 2.3: Example of Lattice scheme.

Perturbative QCD (pQCD) is the most used method but its validity is limited to high energy scales where the strong coupling constant is small. It is based on the order by order expansion in the coupling constant $\alpha_s = \frac{g_s^2}{4\pi}$. For example in this regime a cross section of one process is expressed by

$$\sigma = \sigma_0 + \sigma_1 \alpha_s + \sigma_2 \alpha_s^2 + .. \quad (2.4)$$

²Jet multiplicities and so on

2 QCD Theory

where the σ_i are the cross sections at different perturbative orders. They are evaluated with the help of Feynman diagrams (Figure 2.4) where each QCD vertex contributes as α_s . The first order diagrams consist of tree diagrams, second order loop diagrams and so on. One might calculate the first terms only since the others should be small.



Figure 2.4: Feynman diagram at LO and NLO for Z+jet process.

2.1.2 Renormalization and running coupling constant

Let us consider a perturbative expansion of one physical observable at one determined energy scale Q . At second perturbative order, an ultraviolet singularity appears in large loop momenta. To handle this divergence a dimensional regularization procedure is done. This introduces an arbitrary renormalization scale μ that represents the point at which the subtractions are performed to remove the ultraviolet divergence.

μ is an arbitrary parameter and is not present in the QCD Lagrangian. For these reasons a physical observable σ should not depend on μ and using an effective way the dependence on μ can be absorbed in the renormalized coupling constant $\alpha_s(\mu)$.

Mathematically this may be expressed by

$$\mu^2 \frac{\partial}{\partial \mu^2} \sigma(Q^2/\mu^2, \alpha_s) = \left[\mu^2 \frac{\partial}{\partial \mu^2} + \mu^2 \frac{\partial \alpha_s}{\partial \mu^2} \frac{\partial}{\partial \alpha_s} \right] \sigma = 0 \quad (2.5)$$

Defining $t = \ln \frac{Q^2}{\mu^2}$ and $\beta(\alpha_s) = \mu^2 \frac{\partial \alpha_s}{\partial \mu^2}$, the previous equation becomes

$$\left[-\frac{\partial}{\partial t} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} \right] \sigma = 0 \quad (2.6)$$

Solving this equation and later differentiating the solutions one finds the following differential equations

$$\frac{\partial \alpha_s(Q^2)}{\partial t} = \beta(\alpha_s(Q^2)), \quad \frac{\partial \alpha_s(Q^2)}{\partial \alpha_s} = \frac{\beta(\alpha_s(Q^2))}{\beta(\alpha_s)}$$

where $\alpha_s(Q^2)$ is a new function called running coupling constant. In this way we have put the scale dependence in the running coupling constant $\alpha_s(Q^2)$.

Using the renormalization group equation, α_s may be expressed by the following formula:

$$t \frac{\partial \alpha_s}{\partial t} = \beta(\alpha_s(Q^2)) \quad \beta(\alpha_s) = -b\alpha_s^2(1 + b_1\alpha_s + \dots)$$

2.2 A typical hadron collision and the factorization theorem

where

$$b = \frac{11C_A - 2n_f}{12\pi} \quad b_1 = \frac{153 - 19n_f}{24\pi^2}$$

n_f is the number of light quarks and C_A is the color factor $C_A = 3$.

β is the derivative of α_s with respect to the energy scale, thus when β is negative, α becomes small for large energy scales. This is the meaning of asymptotic freedom [15], the fact that the coupling becomes weaker at high momentum scales. In this region quarks and gluons are treated as free particles which do not interact, and the perturbative expansion of QCD is valid. On the other side, when the scale is small the interaction becomes strong, this leads to confinement of quarks and gluons, they are constrained to form colorless clusters called hadrons.

If $\alpha_s(Q^2)$ and $\alpha_s(\mu^2)$ are in the perturbative region the higher terms of the perturbative expansion can be neglected; thus a simple solution for $\alpha_s(Q^2)$ is:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + b\alpha_s(\mu^2) \log \frac{Q^2}{\mu^2}} = \frac{1}{b \log \frac{Q^2}{\Lambda^2}} \quad (2.7)$$

Perturbative theory expresses how the coupling constant varies as a function of the energy scale, but experimental measurements are needed to determine it.

The parameter Λ was historically introduced as a reference scale. It is dimensional parameter defined as:

$$\log \frac{Q^2}{\Lambda^2} = - \int_{\alpha_s(Q^2)}^{\infty} \frac{\partial x}{\partial \beta(x)} \quad (2.8)$$

and it represents the scale at which the coupling would diverge and also the order of magnitude of the energy where the perturbative theory is valid, the actual value of Λ is ~ 200 MeV. Its precise value depends on the perturbative order at which it is evaluated and on the number of active flavors. For energy scales $Q \gg \Lambda$ the perturbative approximation is valid since there is $\alpha_s \ll 1$, while for energies $\sim \Lambda$ the interaction between quarks becomes very strong and the perturbative QCD is not longer applicable.

In the last years a lot of experimental measurements were performed for several processes and in different energy region to determine the α_s . The actual results, in Figure 2.5, were obtained with important theoretical and experimental improvements, and show a really nice agreement between predictions and experimental measurements. They lead to a world average estimation at the reference scale of Z^0 boson mass of $\alpha_s(M_{Z^0}) = 0.1184 \pm 0.0007$.

2.2 A typical hadron collision and the factorization theorem

It is fundamental to understand the phenomenological and theoretical aspects that are present in a typical hadron collider collision, in order to be able to perform some predictions for the experimental measurements.

In particular we focus our attention on hadron collisions where proton-anti-proton pairs collide at large center-of-mass energies and undergo very inelastic interactions with large momentum

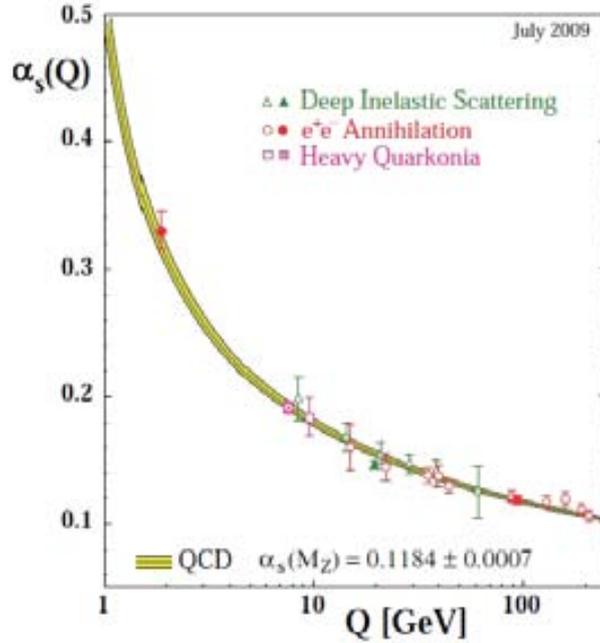


Figure 2.5: The QCD coupling measured at different scales Q and different experiments. The band is obtained by running the world average within its uncertainty.

transfers q between interacting partons. The final state of this interaction is characterized by a large multiplicity of hadrons associated with the evolution of the fragments that have interacted.

The fundamental concept that allows us to calculate the predictions of these physical phenomena is the factorization [16]. This permits us to separate independent stages of the overall process, each one with its particular dynamics and solution techniques. In particular, we can decouple the complex structure of the proton and the final state hadron formation from the elementary structure of the perturbative hadron interaction of the partons. The delimitation of these two phases is determined by a factorization scale that is explained in details in the following section.

In fact in hadron collisions three main phases are recognized: the initial state of the proton, the hard scattering and its evolution, and the final state characterized by hadronization. While the hard scattering could be described with perturbative methods, the initial and final states are evaluated using phenomenological models extracted from experimental data.

The complete picture of a hadron collision is shown in Figure 2.6, where the different parts are divided by a circle.

According to the factorization theorem the differential cross section as a function of a generic hadronic observable X is expressed by:

$$\frac{d\sigma}{dX} = \sum_{j,k} \int d\hat{X} f_j(x_1, Q) f_k(x_2, Q) \frac{d\hat{\sigma}_{j,k}}{d\hat{X}} F(\hat{X} \rightarrow X; Q) \quad (2.9)$$

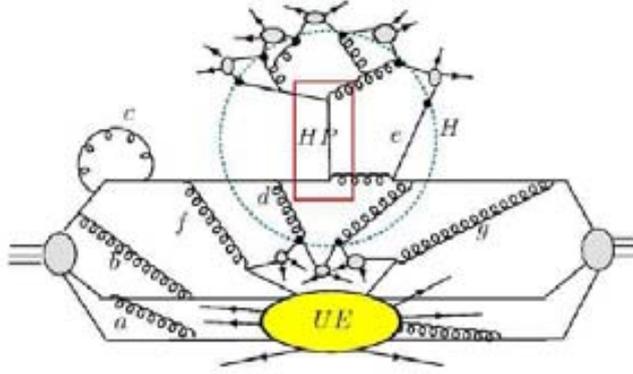


Figure 2.6: The three main phases of a hadron collision: at the beginning the proton (anti-proton) is made of quarks with a continuous exchange of gluons at high virtuality (well-described by PDF) and in the collision one from each proton interacts in a hard-scattering process (HP) forming other partons at high momentum transfer. These start to radiate gluons until they reach low energy scale, where the strong interaction becomes very strong and constrains the quarks to form colorless clusters, (hadrons). Nearby partons merge into colorless clusters that then decay phenomenologically into physical hadrons. Partons that are not involved in the hard scatter could interact later and are called Underlying events (UE).

where the sum is over the j, k parton types inside the proton, the function $f_j(x, Q)$ (PDF) parameterizes the number density of parton type j with the momentum fraction x in a proton at a scale Q ; \hat{X} is the parton level kinematic variable; $\hat{\sigma}_{j,k}$ is the parton cross section and $F(\hat{X} \rightarrow X; Q)$ is the transition function that parameterizes the hadronization. Even if not formally proven³, such factorization scheme is the core of predictive power in the theoretical description of hadron-hadron collisions.

2.3 The Initial State: PDFs and their evolution

Over the last years the knowledge of PDFs and the proton structure has been widely developed with deep inelastic scattering experiments (DIS) [17]. The DIS, represented schematically in Figure 2.7, is a lepton-proton scattering in which the photon exchanged between lepton and the proton has a large virtuality Q .

The idea is that by measuring all the kinematical variables of the outgoing lepton one can study the structure of the proton in terms of the probe characteristics. DIS experiments, such as ZEUS and HERA, confirmed the parton structure and the existence of gluons, achieving the formulation of the so-called Parton Model. This is characterized by a proton formed

³In fact, factorization was proven only for a certain specific processes such as electron-positron annihilation, DIS, Drell-Yan process, single-particle inclusive production in hadron collisions.

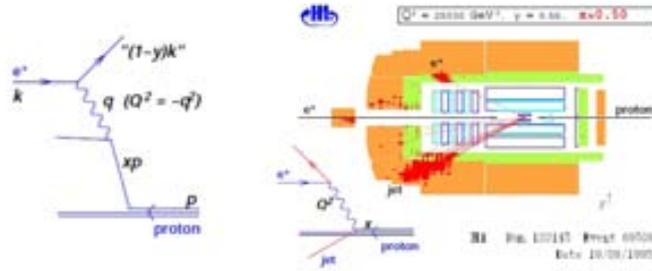


Figure 2.7: Deep Inelastic Scattering scheme and HERA experimental layout.

by point-like partons (valence quarks, sea quarks and gluons) that carry a fraction of the proton momentum. In this naive model the PDFs do not depend on the scale where they are evaluated. Therefore, they only depend on the momentum fraction carried. This is known as Bjorken scale [19]. Partons inside the proton are seen as point-like particles and they can be treated as free particles.

This naive parton model had a great success due to the nice theoretical predictions of a several categories of high-energy cross sections performed using only known partonic cross sections and parton distributions that would show an universal behavior independently on the particular scattering process.

This model was surpassed by QCD since it was not able to account for some experimental results, such as the violation of Bjorken scale for small and large value of x . It can be considered a 0 order approximation of QCD.

In fact, as the probing scale is increased, the observed parton is resolved into several, softly interacting particles (Figure 2.9): the increase in number of constituent partons turns in a decrease of the momentum carried by each of them. This implies an increase in the parton densities at low momentum fraction values, and a decrease of the densities at high momentum fractions. Such scale dependence, known as the breaking of the Bjorken scaling, was experimentally observed.

Let's consider the QCD point of view: inside the proton soft gluon exchanges between quarks continuously occur. Considering a DIS process as in Figure 2.10 with a photon of energy $Q(q)$ interacting with a quark inside a proton, one sees that gluons emitted from quarks at a scale μ greater than q are re-absorbed, since their lifetime is very short. Therefore the quark remains unchanged after the re-absorption. On the other hand a gluon emitted at scale less than Q has a lifetime longer than the time it takes for the quark to interact with the photon, and by the time it tries to reconnect to its parent quark, the quark has been kicked away by the photon. Since the gluon has taken away some of the quark momentum, the momentum fraction x of the quark as it enters the interaction with the photon is different from the momentum it had before. Therefore its density is affected. Nevertheless we could show that the quark state measurement is independent from the nature of the probe particles and depends only on the fraction of momentum and the energy of the probe. This is known as the universality of the PDFs, by which we can extract the PDF from an experiment and

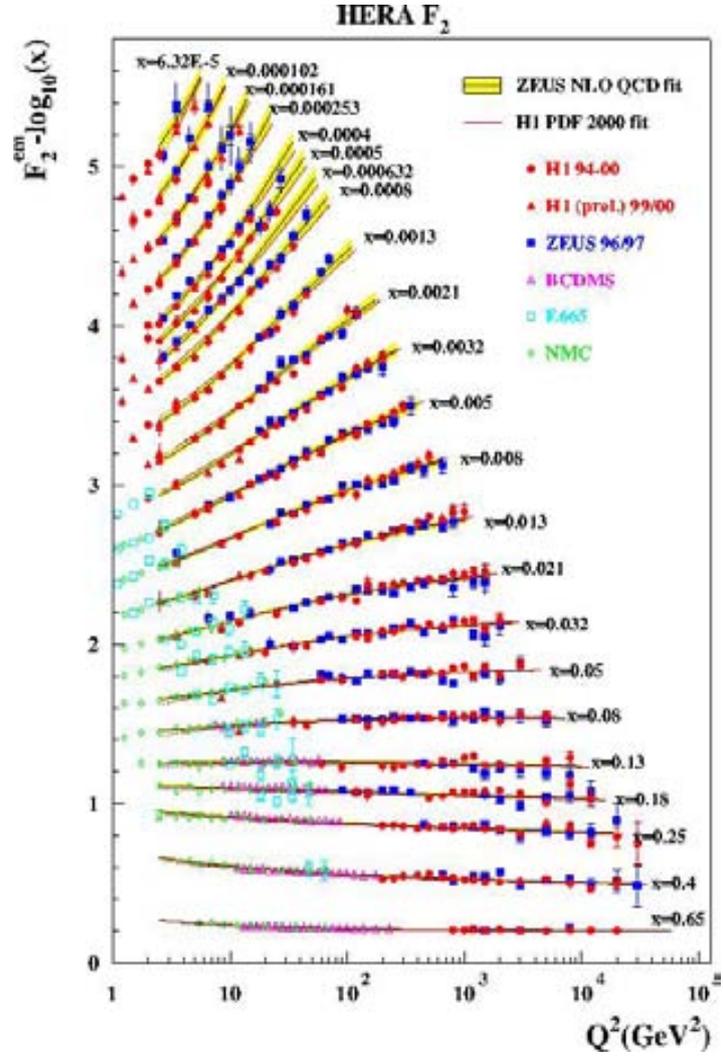


Figure 2.8: Experimental results for F_2 as a function of Q^2 for many different x values, compared to the results of a global fit by the ZEUS collaboration. From [11]

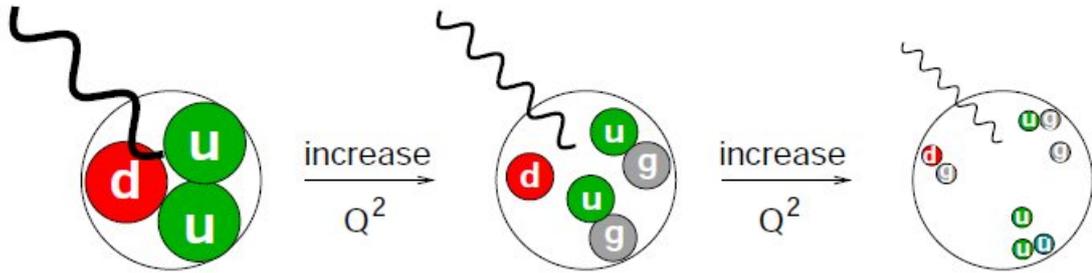


Figure 2.9: With ever shorter wavelength photon probes one resolves more and more structure inside the proton. From [11]

apply the results to other processes/experiments.

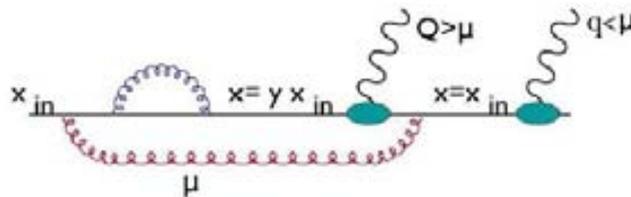


Figure 2.10: Soft gluon emission inside the proton and an deep inelastic scattering scheme with a photon.

For these reasons it is important to study quantitatively the DIS process in which a quark emits a gluon

$$\gamma^* + q \rightarrow g + q$$

When evaluating this process, a logarithmic collinear infrared divergence appears. This divergence is fixed introducing a small cut-off μ , called factorization scale. The main idea in using this cut-off is that any emission that occurs with energy greater than μ is absorbed (factorized) into the PDF itself. Thus the PDFs become a function of μ . As a result, a renormalized quark distribution function $q(x, \mu^2)$ (where the infrared divergence is absorbed) is redefined.

In the same way an analogous parton distribution function $g(x, \mu^2)$ is evaluated for the gluon considering $\gamma^* + g \rightarrow \bar{q} + q$.

$q(x, \mu^2)$ and $g(x, \mu^2)$ cannot be evaluated with perturbative theory since some contributions are in the long-distance region of strong interaction where pQCD loses its validity. However they can be determined experimentally from structure functions data $F(x, Q^2)$ at any particular scale, since $F(x, Q^2) = x \sum e^2 q(x, Q^2)$. This is the meaning of the factorization

theorem that allows us to separate or factorize the long and short distance contributions of any physical cross section involving large momentum transfer.

As already stated perturbative QCD does not predict the form of the PDFs but can describe their evolution with the variation of the scale μ^2 . In fact the structure function of a proton should be independent from the choice of μ due to its arbitrariness. So differentiating, as we did for the renormalization coupling constant, the following relations are obtained:

$$t \frac{\partial}{\partial t} q(x, t) = \int_x^1 \frac{dy}{y} P\left(\frac{x}{y}\right) q(y, t) \quad (2.10)$$

known as the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equation [18], where the convention $t = \mu^2$ is used.

This equation is the analogous of the β function equation and it describes the evolution of the quark and gluon parton distribution function as a function of the momentum fraction x carried by the parton and of the scale t . This is expressed using the so-called splitting functions P . They are gluon radiations ($q \rightarrow qg$), gluon splittings ($g \rightarrow gg$) and quark pair productions ($g \rightarrow q\bar{q}$) and can be calculated with perturbative QCD.

More precisely, the evolution equation for the quark field is given by:

$$t \frac{dq(x, t)}{dt} = \frac{\alpha_s(t)}{2\pi} \int_x^1 \frac{dy}{y} \left[q(y, t) P_{qq}\left(\frac{x}{y}, \alpha_s(t)\right) + g(y, t) P_{qg}\left(\frac{x}{y}, \alpha_s(t)\right) \right] \quad (2.11)$$

and for the gluon field:

$$t \frac{dg(x, t)}{dt} = \frac{\alpha_s(t)}{2\pi} \int_x^1 \frac{dy}{y} \left[g(y, t) P_{gg}\left(\frac{x}{y}, \alpha_s(t)\right) + \sum_{q, \bar{q}} q(y, t) P_{gq}\left(\frac{x}{y}, \alpha_s(t)\right) \right]$$

where we can express the splitting functions at LO calculations:

$$\begin{aligned} P_{qq} &= \frac{4}{3} \frac{1+x^2}{1-x} \\ P_{gg} &= 2C_A \left[\frac{1-x}{x} + \frac{x}{1-x} + x(1-x) \right] \\ P_{qg} &= \frac{1}{2} [x^2 + (1-x)^2] \end{aligned}$$

The Leading-Order DGLAP splitting functions P_{ij} have an appealing physical interpretation as the probabilities that a parton splits collinearly into a parton j plus something else carrying a momentum fraction x of the original parton.

Solving the DGLAP equations at leading order we can easily understand the behavior of parton densities. An increase in the scale turns in a decrease of the densities at large x and in a increase at small x . In other terms, when the scale increase there is an increase in the phase space for gluon emission by the quarks with a consequent reduction in quark momenta (Figure 2.11).

In literature and on the common LHAPDF interface ⁴ different PDF distributions/releases are available. The most used at LHC and TeVatron are CTEQ⁵ [20], MSTW⁶ [21] and the

⁴<http://projects.hepforge.org/lhapdf/>

⁵Coordinate Theoretical-Experimental project on Qcd

⁶Martine-Stirling-Thorne-Watt

2 QCD Theory

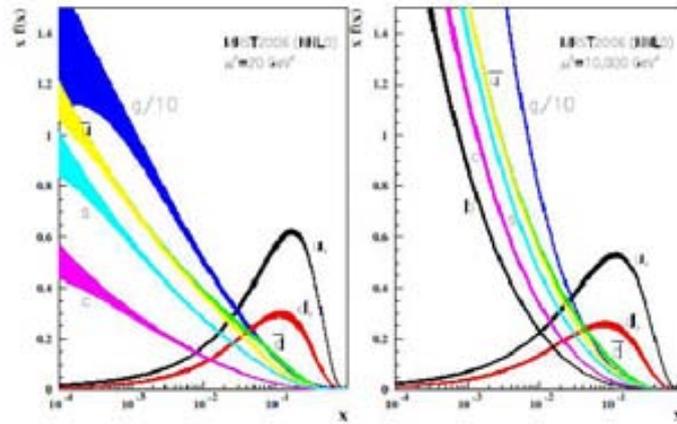


Figure 2.11: Example of proton PDFs measured at $Q^2 = 20 \text{ GeV}^2$ and at $Q^2 = 10000 \text{ GeV}^2$ in a DIS experiment. The contribution coming from gluon increase with Q^2 .

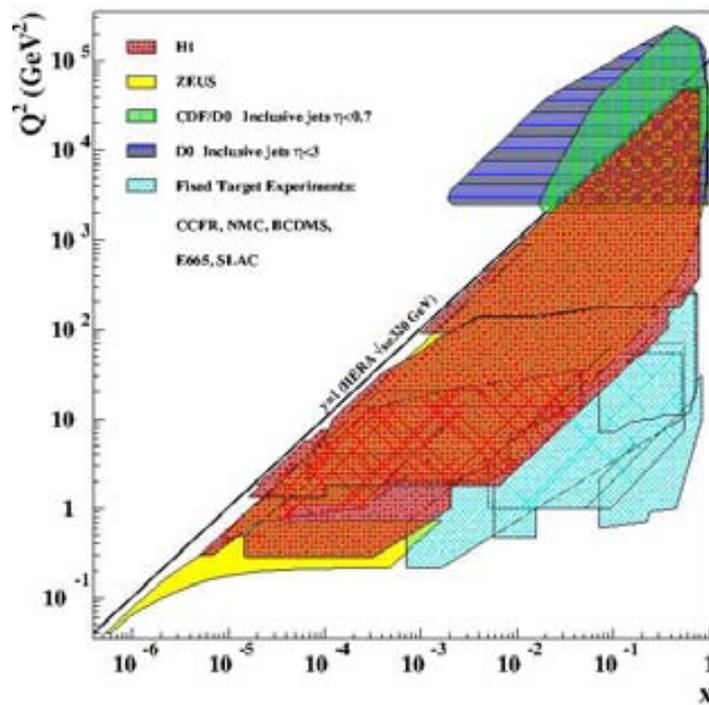


Figure 2.12: Kinematic regions and data sets typically used in PDF fits.

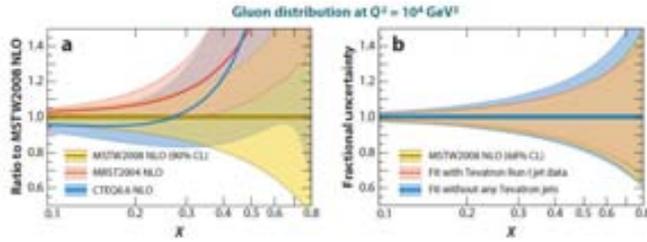


Figure 2.13: PDF gluon distributions comparison between different distributions.

recent NNPDF⁷ [22]. All of these are obtained through global fits to experimental data. The data of different kinematic range (Figure 2.12) used are results of deep inelastic scattering, Drell-Yan and jet data from Tevatron, fixed target experiments as well as HERA. These three PDF releases differ from each other in many aspects: the input data, the value of α_s , the treatments of heavy quarks, the value of heavy quark masses, the parameterization of PDFs, the implementation of them and the way to treat and to include the experimental uncertainties. The main features are summarized in the table 2.1.

	CTEQ	MSTW	NNPDF
parameters	20	20	259
$\alpha_s(M_Z)$	0.118	0.120	0.119

Table 2.1: Comparison between the main important PDF distribution

As one can see the PDFs are affected by an uncertainty due to the need to combine large number of datasets from different experiments (Figure 2.12) and different theoretical inputs. A way to estimate this uncertainty is based on the Hessian formalism [23]. As told before the extraction of PDFs is based on global fits to data, done through a minimization of an effective global χ^2 in the space of the free parameters. The method to calculate the uncertainties consists in considering the variation of the χ^2 around the minimum neighborhood. We chose a variation of χ^2 around a minimum in a region determined by an arbitrary parameter T , called tolerance parameter, in order that

$$\Delta\chi^2 \leq T^2$$

Usually T is chosen to be equal to 10 or 15 and it is tuned in relation to the quality of the agreement with experimental data.

Expanding it quadratically one obtains:

$$\Delta\chi^2 = \chi^2 - \chi_0^2 \sim \frac{1}{2} \sum_i \sum_j H_{ij} (a_i - a_i^0) (a_j - a_j^0) \quad (2.12)$$

⁷Neural Network PDF

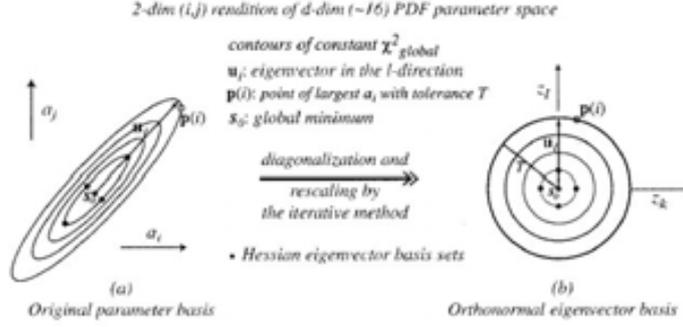


Figure 2.14: Illustration of the Hessian formalism. An iterative procedure diagonalized the Hessian matrix and re-scales the eigenvectors to adapt the step size to their natural scales. [23]

where a_i are the different free PDF parameters and H_{ij} is the Hessian matrix, i.e. the matrix of second derivatives⁸ of the corresponding PDF.

The Hessian matrix has a complete set of orthonormal eigenvectors and the displacements from the minimum are conveniently expressed in terms of those. Therefore for each eigenvector we can find the displacements around the minimum in the direction along the vector, a_i^+ and a_i^- for the i -th eigenvector that represent the up and down uncertainty of the parameter a_i^0 .

The whole picture is shown in Figure 2.14.

The PDF uncertainties for gluon and up quark distribution coming from this procedure are in Figure 2.15. In the u quark distribution the uncertainty is much smaller than in the gluon. This reflects the large amount of experimental data sensible to u quark included in the analysis.

The uncertainties in the pQCD cross section prediction due to the PDFs are determined in the following way:

$$\delta\sigma^+ = \sqrt{\sum_i (\max(\sigma(a_i^+) - \sigma(a_0), \sigma(a_i^-) - \sigma(a_0), 0))^2}$$

$$\delta\sigma^- = \sqrt{\sum_i (\min(\sigma(a_i^+) - \sigma(a_0), \sigma(a_i^-) - \sigma(a_0), 0))^2}$$

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

⁸This follows directly as consequence of $f(x) = f(x_0) + J\Delta x + \frac{1}{2}\Delta x^T H \Delta x$ thus since we are expanding near a minimum, the Jacobian matrix is 0

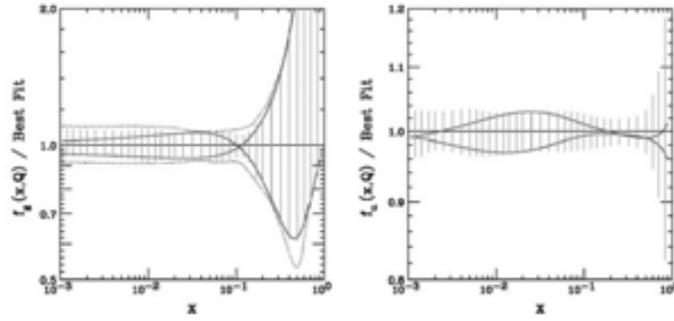


Figure 2.15: Comparison between gluon and u-quark PDF. As told the gluon uncertainty is bigger in particular for x greater than 0.35. [23]

2.4 The final state evolution and hard scattering

So far we have presented the evolution of quarks and gluons in the initial state. In what follows the state after hard interaction will be described. The hard interaction could be evaluated with Matrix Element (ME) methods and perturbative QCD using Feynman rules. Here a phenomenological picture of the final states and their evolution will be given. This is known as parton shower, that describes well the evolution in moment transfer of the partons from the high scales of the hard process to the lowest one associated with the confinement. For simplicity it is better to consider the cleanest and most precise process of hadrons production, that is $e^+e^- \rightarrow Z/\gamma^* \rightarrow q\bar{q}$, i.e. the annihilation of an electron and a positron in a virtual photon or a Z which decays immediately in a $q\bar{q}$ pair (Figure 2.16). This is useful for our scope because QCD is only involved in the final states and it origins a point like source of quarks pairs, easing the description of the process. In any case the conclusions are general.

The $q\bar{q}$ formed in the hard interaction could be considered as free particles since at high

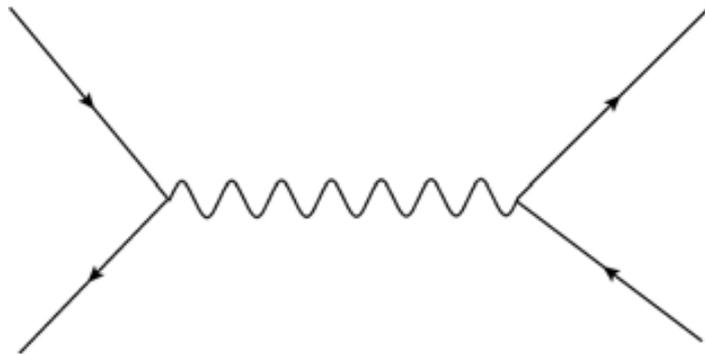


Figure 2.16: $e^+e^- \rightarrow Z/\gamma^* \rightarrow q\bar{q}$ LO diagram

energy α_s is small. Also the gluons radiated from them could be treated perturbatively, at least until they reach energies of ~ 1 GeV. At this scale the coupling becomes quite strong and hadronization begins to play an important role, clustering partons in a colorless bunch.

2 QCD Theory

The clearest experimental evidence is the jet formation coming from the evolution of the partons scattered.

High hadron multiplicity is another feature of hadron collisions. This is a direct consequence of gluon emission, due to the fact that gluons emitted from quarks can emit other gluons in the same direction of the emitting quark.

In order to explain this process more precisely, it is interesting to illustrate the emission of soft gluons from the struck quark (showering). Hence a soft and collinear emission of gluons is considered. Soft means that the emitted gluon is at very low energy when compared to the quark and collinear means that it is emitted at a angle close to the forward direction.

Thus the full differential cross section for $q\bar{q}$ production plus soft gluon emission can be written as the $q\bar{q}$ matrix element and the phase space multiplied for a soft gluon emission probability dS :

$$|M_{q\bar{q}g}|^2 d\Phi_{q\bar{q}g} \sim |M_{q\bar{q}}|^2 d\Phi_{q\bar{q}} dS \quad (2.13)$$

and the probability of soft gluon emission from the $q\bar{q}$ system is expressed by:

$$dS = \frac{2\alpha_s C_F}{\pi} \frac{dE}{E} \frac{d\theta}{\sin\theta} \frac{d\phi}{2\pi} \quad (2.14)$$

where ϕ is the azimuthal angle of the gluon with respect to the quark, and θ is the polar angle.

This result shows the presence of two non integrable divergences: one is called infrared divergence and takes place when $E \rightarrow 0$, while the other is the collinear divergence when $\theta \rightarrow 0$, i.e. when gluon becomes collinear with the quark direction. These two divergences are a general property of QCD and appear when there is a gluon emission from a quark. They are not physical, they simply indicate a breakdown of the perturbative approach.

Two more important aspects of the gluon emission after hard scattering are the angular and color ordering. The angular ordering consists in the continuous reduction of the angle of gluon radiation, so the gluon emission results, in its evolution, more and more collinear with the quark. On the other hand, color ordering (that could be better illustrated in the Figure 2.17) forces the $q\bar{q}$ pairs that are in the color singlet to be close in phase space achieving a sort of confinement. Both processes have interesting consequences for the hadronization because they prepare the confinement and the clustering in colorless hadrons.

The fact that color always flows directly from the emitting parton to the emitted one, the collimation of the jet and the softening of the radiation emitted at later stages ensure that partons form a color-singlet cluster close in the phase space. As a consequence, hadronization occurs locally inside the jet and only partons nearby are involved. This was also formulated as Local Parton Hadron Duality [24] which affirms that the transition from partons to hadrons is local in the phase space. Therefore, the hadrons direction and their kinematic are closely related to the original partons. This principle allows us to extrapolate parton information from a jet measurement.

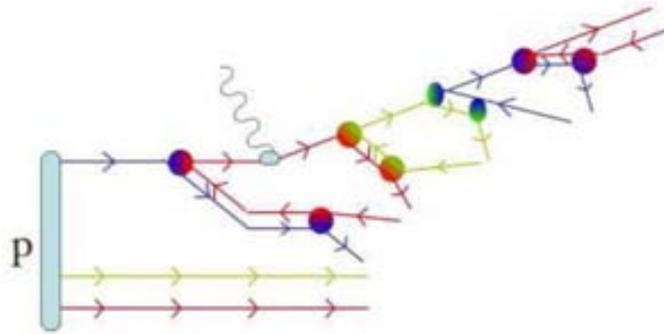


Figure 2.17: Color flow in a DIS process. Figure taken from [12]

2.5 The hadronization

The perturbative theory is valid until the partons reach energies of 1 GeV (infrared cutoff). That is the energy where the strong coupling constant becomes quite strong and the gluons emitted have enough energy to create pions, the lightest particles. The transition from partons to hadrons is described by a non perturbative model that is implemented in the MC generators. The two models used nowadays in high energy physics are Cluster Model [25] implemented in Herwing and Lund Model [26] implemented in Pythia. These models are universal, since the hadronization is independent of the hard scattering energy and final states. In fact hadronization starts only when partons reach the infrared cutoff. A more energetic hard scattering implies only that the showering before the hadronization results larger.

The simplest process to form a cluster occurs when a quark from one gluon combines with an anti-quark carrying the opposite color, forming a color single state. The quarks kinetic energy slowly converts into potential energy between the quarks forming resonances. These resonances are usually unstable and could decay in stable particles.

The measurement of hadron multiplicities from Z decays is used to tune the phenomenological parameters of the model. These parameters can be used to describe hadronization at different energies and in different high-energy hadron production colliders. This consistence is tested looking at the invariant mass distribution of clusters of quarks after non perturbative gluon splitting at different energy for hard scattering (Figure 2.18).

There are two important concepts in hadronization: the local parton to hadron duality and the low scale effective α_s that permits to extend the use of perturbative QCD also to low scales. These two principles are used in the models implemented in MC event generators. The main difference between the string model (Lund) and the cluster model is that the latter transforms the partonic system directly into hadrons, while the former employs an intermediate stage of cluster objects with a typical mass scale of few GeV.

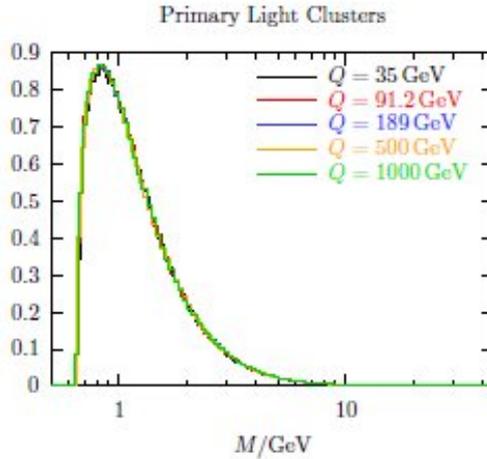


Figure 2.18: Invariant mass distribution of cluster of color singlet quarks after non-perturbative gluon splitting.

2.5.1 The cluster model

In the cluster model the perturbative QCD uses the parton evolution until low energy (beyond the infrared cutoff) and it is based on its preconfinement properties due to angular and color ordering. At this point the gluons are forced to decay in a $q\bar{q}$ pair and all quarks are clustered in color singlet with a mass of a few GeV. These clusters characterized by mass and flavor quantum numbers are treated as resonances that decay into two hadrons proportionally to their phase space. Heavy clusters could decay into smaller clusters, that afterwards decay into hadrons.

This is a simple model that successfully describes the characteristics of hadron distributions in jet fragmentation but it forces perturbative QCD beyond its limits of validity.

2.5.2 The string and Lund Model

In this model the QCD potential between two quarks at low scales is parameterized by field lines seen to be compressed to a tube-like region by self interactions among soft gluons. The potential (Figure 2.19) is given by:

$$V(r) = \frac{4\pi}{3r} + kr \quad (2.15)$$

where r is the distance between the two quarks, and k is the constant of the QCD potential ($k \approx 1\text{GeV}/fb$ from hadron spectroscopy). The first term represents the Coulomb potential that has effects only at low distances, so it is only fundamental in the internal organization inside hadrons and not in their formation.

In this context string represents the field line connecting two quarks and the force field is linearly increasing with the distance according the QCD potential. The hadronization

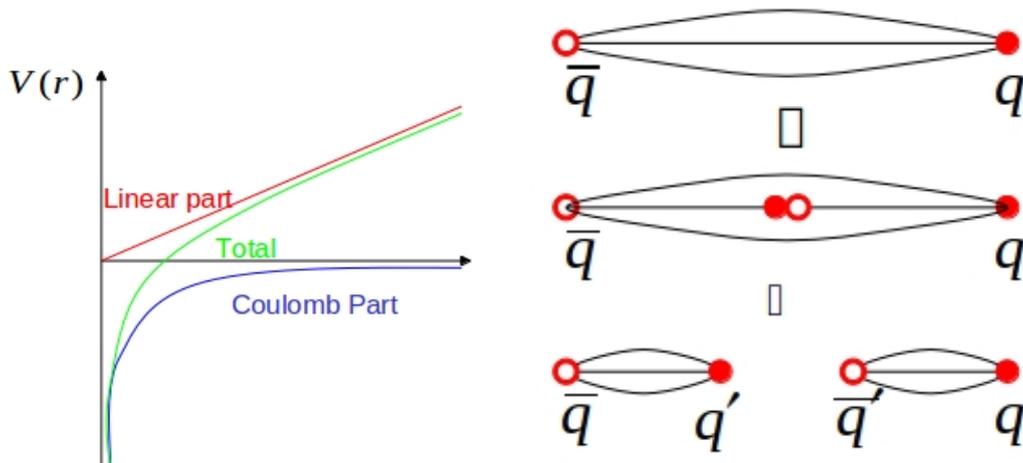


Figure 2.19: QCD potential and string model scheme.

dynamics is described by the dynamics of the string. Therefore as the two original quarks separate from each other the potential energy grows linearly until it reaches a level beyond which the string breaks forming another $q' \bar{q}'$. Thus two new strings are obtained, both with determined energy, mass and quantum numbers. If the invariant mass of the $q\bar{q}'$ or $q'\bar{q}$ pairs is sufficient the strings can break again, otherwise the process stops.

In particular in the Lund string model (Figure 2.20) is assumed that the process continues until all hadrons are on-shell. The hadrons produced retain a fragment of the original quark momentum.

The string model is collinear and infrared safe, i.e. the emissions of a collinear and/or soft gluon does not perturbate the fragmentation of a string as approaching the small angle/energy limit, but it is not able to take into account for collective phenomena that could eventually happen in the high-energy hadronic collisions since it considers only independent strings.

2.5.3 Underlying event (UE)

Another phenomenon that we have to consider is the presence of underlying events that are originated by the secondary interactions with hadrons remnants that do not participate in the hard interaction and contribute to the final state. These processes are non perturbative and are modeled as hadronization. Underlying event can come from two different sources: the hadronization of partonic constituents that have undergone multiple parton interactions (MPI) and the activity along the beam direction, concerned the hadronization of the beam remnants. UE models are tested with measurements sensitive to it like jet shapes and event profile.

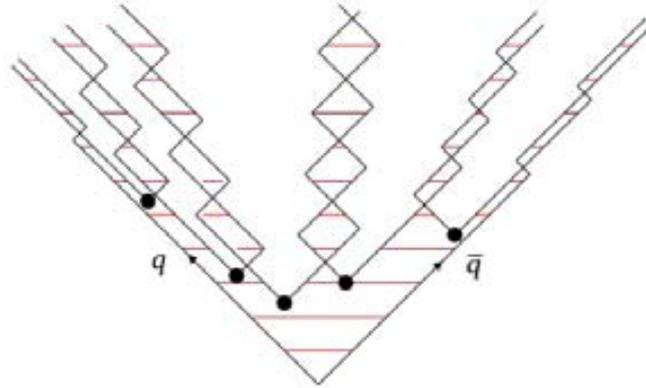


Figure 2.20: Lund Model

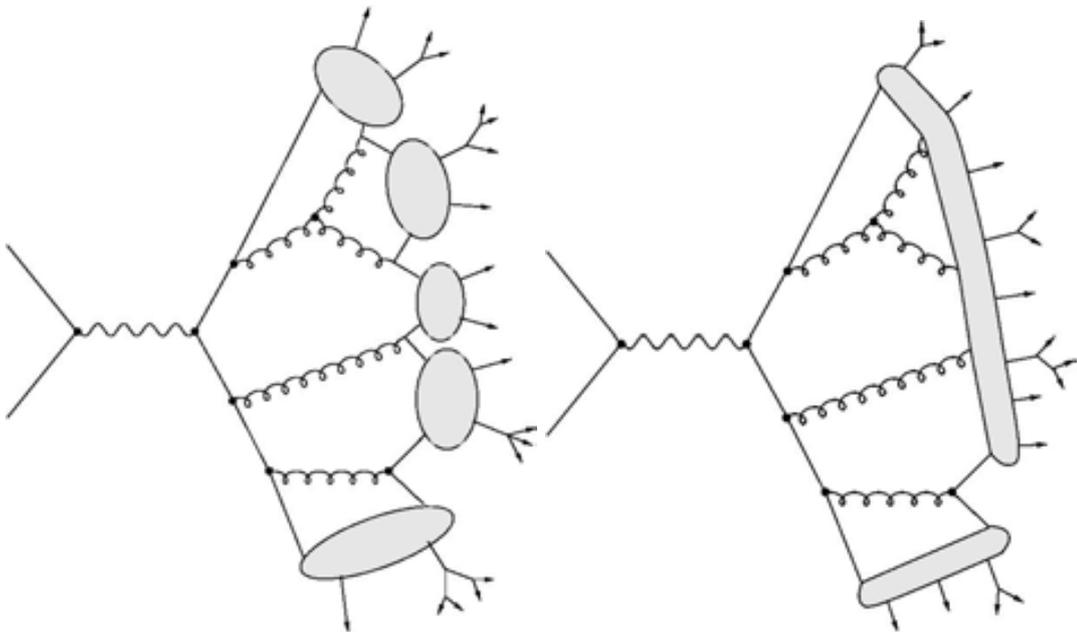


Figure 2.21: Illustration of string and cluster fragmentation.

2.5.4 Multiple Interactions

In a hadron collider, usually more than one pair of partons interact leading to the possibility of multiple interactions. This could have a non negligible effect in the final state: for example MPI increase the multiplicity and summed transverse energy. For this reason it is important to have a MC model that takes into account this effect, avoiding the misreconstruction of the event. Nowadays all MC event generators take into account MPI effects.

2.6 The predictive tools

MC event generator programs are used at hadron colliders in many parts of the analysis to understand the phenomenology of the physical process and to predict variable distributions. These are useful tools in high energy physics that can be used at various stages of an experiment, for example to optimize the detector study, its trigger design or to estimate some physical processes and test the analysis strategy.

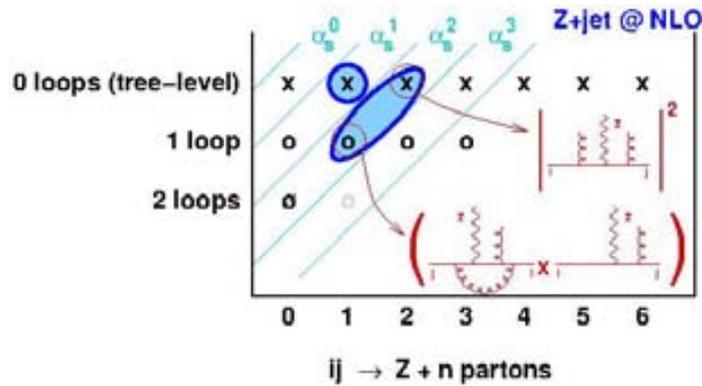


Figure 2.22: Illustration of the contributions that are known for $ij \rightarrow Z + N$ partons where i and j are arbitrary incoming partons, according to the number of outgoing partons, the number of loops are the number of powers of the coupling. An x means a squared tree-level diagram, and o represent the interference of 1-loop diagram with a tree diagram, while O represents the interference of a two-loop diagram with a tree one. The entries in the shaded ellipses are those that are relevant for NLO calculation of the cross section for the production of a Z boson with a jet. Image is taken from [11]

There are two main kinds of MC event generators: the fixed order perturbative matrix elements (ME) and the parton shower (PS). ALPGEN [27], MADGRAPH and MCFM belong to the first category, PYTHIA [33] and HERWING, that are the most common, to the latter. ME MCs are tools that generate events at parton-level evaluating the matrix element of the hard process calculating only the first terms of the perturbative expansion. In particular leading order (LO) MC stops at first (Born) term, next-to-leading (NLO) order considers also

the 1-loop corrections, and next-to-next-to-leading (NNLO) order up to 2-loop corrections.

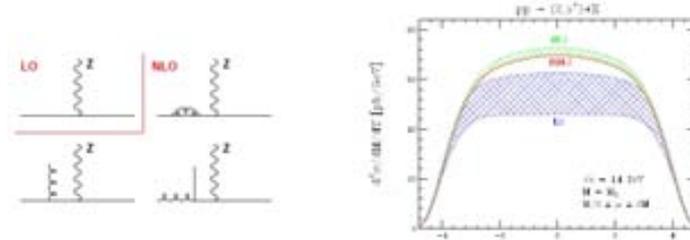


Figure 2.23: Feynman diagrams at LO and NLO for Z+jets process. In the right plot the cross section at LHC for the Z boson differential in rapidity at LO, NLO, NNLO is shown. Image taken from [11]

2.6.1 LO event generators

LO ME includes only the calculation of tree-level diagrams, in kinematic regions in which their contributions are finite. Therefore the simplest approach is to carry out a Monte Carlo integration over phase-space points with a subroutine that determines whether a given phase space point passes the cuts and it calculates the squared matrix elements and PDF factors for each partonic subprocesses.

The problem with LO predictions is that only the shape of distributions is reliable since the absolute normalization is badly described due to large contribution coming from higher order corrections. It is common to introduce a K factor when comparing results from event generators with experimental data. This K factor is the ratio between NLO cross section prediction and LO.

Many MC programs are available to make LO predictions, in particular the most used are ALPGEN and MADGRAPH. They allow to perform a prediction at high parton multiplicity, up to 6 partons in the final states. In fact they use methods designed to be particularly efficient at high multiplicities which build up amplitudes for complex processes by recursively reusing simpler ones.

ALPGEN

It is a tree-level ME generator and allows us to calculate multi-partons (up to 10) cross sections for processes in hadronic collisions. As to the parton level description it could be interfaced with parton shower programs (PYTHIA, HERWING) that simulate the full showering and hadronization of the event. It uses the ALPHA algorithm to compute tree level scattering amplitudes for large parton multiplicities in the final state. The idea of ALPHA algorithm [28] is that Matrix Element is the Legendre transformation of the effective Lagrangian. In such a way the problem can be reduced to a simpler situation, more suitable for a numerical approach (for example the number of degrees of freedom is finite). Therefore

it could be evaluated in a reasonable CPU time. The advantage of ALPHA is that its complexity increases slower than the number of Feynman diagrams when increasing the particles in the final state.

2.6.2 NLO predictions

Higher order calculations (including loop effects) are not fully automated. They consist of more than just one matrix element with a fixed number of final state particles, but they include terms with extra particles in loops and legs (Figure 2.23). This extra emission introduces infrared divergences which must cancel between the various terms. This is done technically with infrared subtraction methods.

One must note that NLO calculations are computing intensive: for some observables it is not unusual to have to devote several years of CPU time in order to get adequate numerical convergence of the Monte Carlo integration.

Example of NLO generators are MCFM [30] and BLACK-HAT.

MCFM

Literally a Monte Carlo for FeMtobarn processes, it is a NLO ME event generator that permits us to calculate Di-boson and Drell-Yan+ heavy flavor processes. MCFM uses the dipole method to cancel the infrared divergence between real and virtual one loop contribution. It was developed by Campbell et al. and it had a good success to predict TeVatron results and distributions, since one of the main features is to be less sensitive to scale dependence. Full differential distributions are available.

2.6.3 NNLO predictions

NNLO predictions suffer from the same problem of canceling divergences between real and virtual corrections that are present at NLO, with the complication that instead of having one soft and one collinear divergence there are now two of each, greatly complicating the task of figuring out counter-terms to allow experimental cuts to be implemented in four dimensions.

2.6.4 Parton Shower

The problem with ME generators is that they do not cover the regions where partons become soft and collinear and they stop the prediction at parton level. Parton Shower MC takes into account the soft radiation and it evaluates higher orders based on these two concepts:

- an iterative structure that allows simple expressions for $q \rightarrow qg$, $g \rightarrow gg$ and $g \rightarrow q\bar{q}$ branchings to be combined to build up complex multi-parton final states
- a Sudakov factor that offers a physical way to handle the cancellation between real and virtual divergences

The starting point is to factorize a complex $2 \rightarrow n$ process, where n represents a large number of partons in the final states, into a simple core process convoluted with shower as in Figure 2.24.

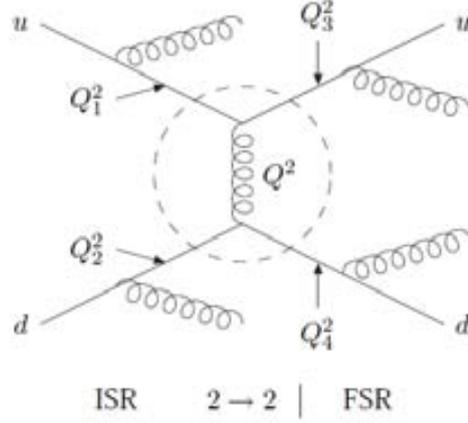


Figure 2.24: $2 \rightarrow n$ factorization PS scheme. Taken from [32]

Here there could be Initial (ISR) and Final (FSR) state radiation, where the probability to emit gluons is described by the DGLAP equation. These can blow up the probability in soft and collinear regions. Thus we introduce the Sudakov form factor [31] that expresses the probability of not emitting a gluon, thereby the DGLAP becomes:

$$dP_{q \rightarrow qg} = \frac{\alpha_s}{2\pi} \frac{dQ^2}{Q^2} P_{q \rightarrow qg}(x) dx e^{-\sum_{q,g} \int_{Q^2}^{Q_{max}^2} \frac{dQ'^2}{Q'^2} \int \frac{\alpha_s}{2\pi} P_{q \rightarrow qg}(x') dx'} \quad (2.16)$$

where the exponent is so-called the Sudakov factor. The inclusion of a Sudakov form factor ensures that the total probability for a parton to branch never exceeds unity.

The implementation of a cascade evolution (Figure 2.25) now makes sense. Starting from a simple $q\bar{q}$ system the q and \bar{q} are individually evolved downwards from the initial Q until they branch. In a branching the mother parton disappears and is replaced by two daughter partons, which in turn are evolved downwards in q and they may branch. Therefore the number of partons increases until the infrared cutoff scale is reached.

There are initial and final states radiation that share the same probabilistic framework, DGLAP equation with Sudakov form factor, where an initial parton undergoes successive branching. For ISR the branching is in terms of higher and higher space-like virtualities as the hard scattering approaches, while for the FSR the branching involves lower and lower time-like virtualities as the hard scattering recedes.

PYTHIA 6.425, MC tunings and Perugia 2011

Pythia 6.425 [33] is a Monte Carlo event generator program that calculates the hadron scattering with a LO hard scattering. Although it is optimized for $2 \rightarrow 1$ and $2 \rightarrow 2$ final states,

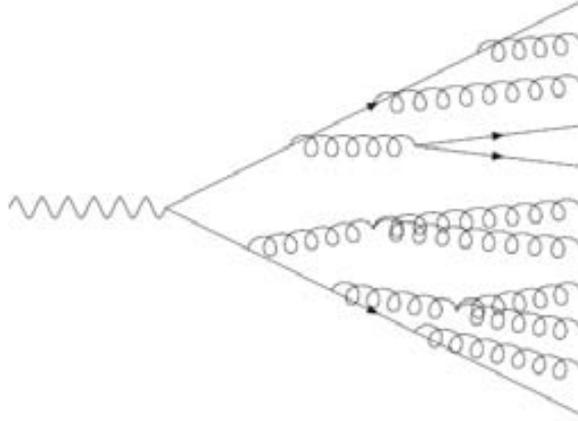


Figure 2.25: Shower Cascade. Taken from [32]

there are some processes available with three or more partons in the final state. Currently there are almost 300 different hard processes implemented. Pythia simulates initial and final state parton shower radiation that permits to have a leading logarithmic order prediction for hadron processes.

The current version differs from the previous one because it has a shower ISR and FSR based on a p_t -order sequence $1 \rightarrow 2$ parton splitting. The hadronization, as explained before, is modelled through a string Lund model.

MC tuning and Perugia2011

Monte Carlo event generators are based on various phenomenological models and have several free parameters that are a priori unknown (for instance flavor ratios, q_0). Since we want a MC that reproduces as well as possible the experimental data, we use them to constrain some free parameters. Also the value of α_s is to be optimized [35]. The problem lies on the quantity of parameters and correlations between them. For this reason, the overall task is divided in parts: parton shower, hadronization, MPI and UE. In such way the number of parameters to tune is reduced. Therefore the procedure consists of choosing some parameters, defining their interval and then evaluating the MC output changing the different parameters. We fit the result obtained to data using different datasets, different observables and we chose the nicest scenario resulting in the model context. This is usually done with PROFESSOR and RIVET MC programs.

In general data from LEP are used to tune flavor parameters (using identified particles multiplicities) and fragmentation, hadronization (with event shapes, b fragmentation measurement, multiplicities and momentum spectra). To tune the UE more than 50 distributions from CDF and D0 are used.

In particular the tuning applied in PYTHIA in this analysis is Perugia2011 [36]. Here LHC Data at 7 TeV are included and the same α_s/Λ for ISR and FSR. This permits a consistent matching with AlpGen. The set of PDF is CTEQ5L.

Other tunings applied in CDF analysis are Tune A and Tune DW. The former is based on CDF Run I data and reproduces well the underlying events, the latter that uses also Run II data, describes well the $P_{T,Z}$ distribution.

2.6.5 MEPS

As we have seen, both matrix elements (ME) and parton showers (PS) have advantages and disadvantages. Summarizing, ME allow a systematic expansion in powers of α_s , and thereby offer a controlled approach toward higher precision. Calculations can be done with several partons in the final state, as long as only Born-level results are asked for, and it is possible to select the phase space cuts for these partons precisely to the experimental needs. Loop calculations are much more difficult, on the other hand, and the mathematically correct cancellation between real and virtual emission graphs in the soft collinear regions is not physically sensible. Therefore ME cannot be used to explore the internal structure of a jet and are difficult to match to hadronization models which are supposed to take over in very soft/collinear region.

PS, on one hand, clearly are approximate and do not come with a guaranteed level of precision. The efficiency in obtaining events in a specific region of the phase space can be quite low. On the other hand, PS are universal, so for any model you only need to provide the basic hard process and then PS will turn that into reasonably realistic multi-partons topologies. The use of Sudakov form factors ensures a physically sensible behavior in the soft collinear regions and it is also here that the PS formalism is supposed to be the most reliable. It is therefore possible to obtain a good picture of the internal structure of jets and to provide a good match to hadronization models.

In conclusion: ME are good for well separated jets, PS for the structure inside the jets.

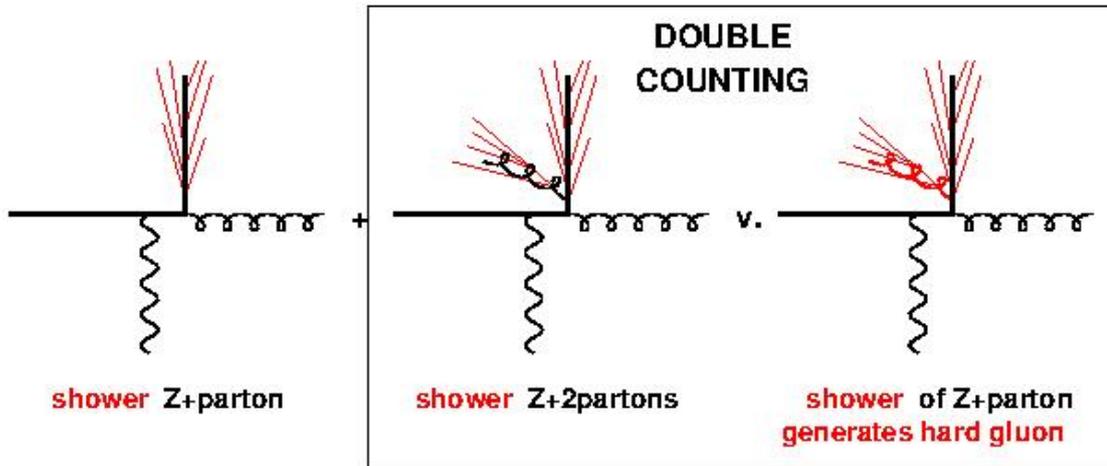


Figure 2.26: Example of double counting for MEPS MC. Image taken from [11]

Clearly the two complement each other, and a marriage is highly desirable. To do this without double-counting (Figure 2.26) is less trivial than several methods we developed, the

most used is vetoed showering done by MLM (Figure 2.27) or CKKW techniques [?].

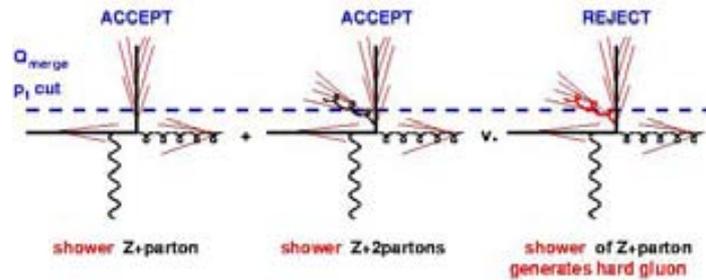


Figure 2.27: MLM mechanism. Figure taken from [11]

In Particular MLM matching proceeds as follow:

- introduce a transverse momentum cutoff Q_{ME} and an angular cutoff R_{ME} for matrix elements generation
- generate tree level hard matrix where all partons must have $p_t \geq Q_{ME}$ and be separated from another parton by an angle greater than R_{ME} . The numbers of events that one generates in the different samples are proportional to their cross section with these cuts
- for each tree level event these samples are showered with a parton shower program
- apply a jet algorithm to the shower event and identify all jets with $p_T \geq Q_{merge}$ where the merging scale is taken greater than Q_{ME}
- if each jet corresponds to one of the partons and there are no extra jets above scale Q_{merge} then accept the events
- otherwise reject the event

Attempts to combine NLO ME with PS are MC at NLO and POWHEG.

2.7 The Jet production

As seen before in QCD fundamental measurements are done with jets. In fact important ingredients to make correct predictions and to perform analysis in hadron colliders are:

- the definition of the jet and its corrections
- the parton distribution functions (PDFs)
- the event generator tuning

Having a correct jet definition and its corrections is fundamental because it permits to understand the relation between theory and experiment and between the long distance degrees of freedom observed in the detector onto the short distance colored partons. Any mismatch represents a limit of our measurement.

2.7.1 The Jet Definition

The idea of the jet [39] as a collimated spray of energetic hadrons is almost known to everybody, but it is worthwhile to set the rules to reconstruct it. It is important to define a jet well because a jet gives the closest idea of the original parton and has important consequences: for example using a jet definition sensitive to soft radiations can cause a misunderstanding of the original parton behavior and could end up in incoherent conclusions.

The jet definition is a set of rules for how to group particles into a jet and how to calculate the resulting quadrimomentum. There are two important concepts here: the jet algorithm, that are the rules to group the particles to have a stable jet, and the jet recombination that is how the jet quadrimomentum is evaluated combining the particles that are inside.

Along the different decades and through the different experiments a large combination between jet algorithm and jet recombination was used giving a large spectra of jet definitions. At the beginning of Run II TeVatron and before the LHC startup some groups working on QCD tried to reorganize and some fundamental criteria were established for a good jet definition.

The known Snowmass criteria [40] for an ideal jet definition are listed here. For these the jet definition should be:

- simple to implement in a different analysis;
- simple to implement in the theoretical calculation;
- defined at any order of perturbation theory;
- yields finite cross sections at any order of perturbative cross section;
- yields a cross section that is relatively insensitive to hadronization

In particular from a theoretical point of view it is fundamental that the ideal jet algorithm should be infrared safe, collinear safe and invariant under boost transformations.

The infrared and collinear safety are two important concepts which concern the singularity present in Feynman diagrams when a parton emits a soft gluon and an outgoing parton splits into two collinear partons. Being infrared safe (Figure 2.30) means that no infrared singularities appear in the perturbative calculations and that it is insensitive to the soft radiation in the event. This implies that an emission of a soft gluon does not change the number of jets reconstructed in the event.

On the other hand, to be collinear safe means that collinear singularities do not appear in the perturbative calculation and that jets are insensitive to collinear radiation in the events. This guarantees that the jets found in the event when splitting a particle with two collinear particles do not change (Figure 2.29).

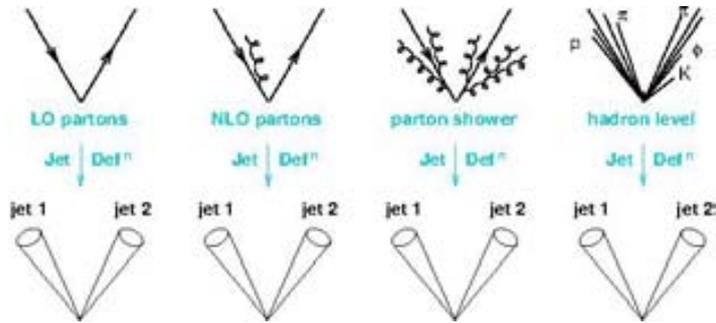


Figure 2.28: Jets at different orders of perturbation theory and at different points in the analysis.

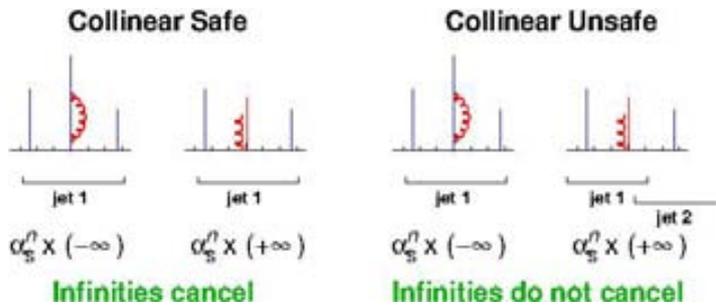


Figure 2.29: Illustration of collinear safety and collinear unsafety in an iterative cone algorithm, together with the implication for perturbative QCD calculation. Partons are drawn with vertical lines. Their height is proportional to their transverse energy, and the horizontal axis indicates the rapidity. Plot from [39]

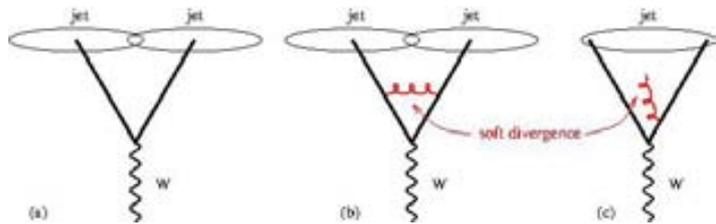


Figure 2.30: Illustration of Infrared unsafety for iterative cone algorithm in events with a W and two partons. The addition of a soft gluon converts the event from having two jets to just one jet. Scheme taken from [39]

2 QCD Theory

The collinear and infrared safety are really fundamental when we are going to compare with the theory in particular experimental measurement with leading order theory without PS. On the other hand, from the experimental point of view, it is crucial that jet algorithm should be detector independent, should not amplify the effects of resolution smearing and angles bias (minimization of resolution smearing) and should be implemented with a minimum of computer time.

There are two categories that group all the different kinds of jets: the cone algorithm that group the particles inside a stable cone and a clustering algorithm that works by grouping together nearby objects by pair-wise. In the following sections the advantages and the disadvantages of the different types are treated in detail.

2.7.2 Cone Algorithm

This algorithm takes inspiration from the first jet definition done by Sterman and Weinberg. Cone algorithms [41] form jets by associating together particles whose trajectories end up within a circle of specific radius R in $\eta \times \phi$ space. Starting with a trial geometric center for a cone in $\eta \times \phi$ space, the energy weighted centroid is calculated including contributions from all particles within the cone. This new point is used as the center for a new trial cone. As this calculation is iterated the cone center is followed until a stable solution is found, i.e. until the centroid of the energy depositions within the cone is aligned with the geometric axis of the cone.

At CDF a first jet cone algorithm (JetClu07) was developed. It used the centers of seed particle that passed a minimum energy cut as the starting points to look for the stable cone. In this way, there is no need to look everywhere and it is more efficient computationally, but the introduction of the seed has as a consequence: the jets are collinear and infrared unsafe, as it is illustrated in the example in Figure 2.30.

Since we have an infrared- and collinear-safe jet, the idea was to avoid the seeds that generate these problems and thus to move to the Seedless-Cone algorithm which is infrared and collinear safe. The only problem concerning this algorithm is its expensive computation even if now with some computational geometry techniques it is not so far from the most common one as explained in Figure 2.32.

A mid way idea was to implement Midpoint. This is a seed cone algorithm that considers also the midpoint between seeds as the starting points. This is computationally faster and it is infrared-safe at least of 3+1 order. An issue was that this algorithm leaves some unclustered energy. In order to avoid it, a search cone of about half a cone was inserted. This caused the jet to be completely infrared safe (Figure 2.31). For this reason in this thesis we are using the MidPoint without search cone

2.7.3 Clustering Algorithm

The cluster algorithm groups nearby objects pair-wise in relation of the generic distance between two object i, j . This is given by:

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \quad (2.17)$$

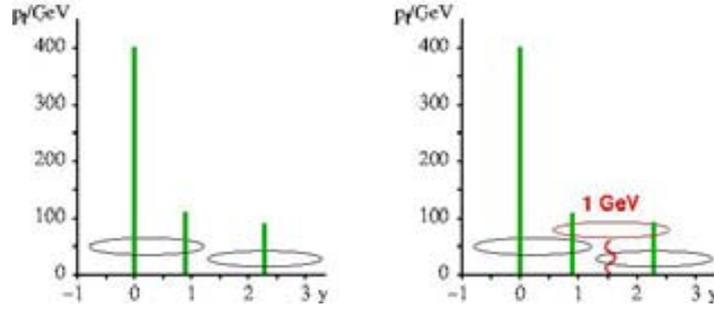


Figure 2.31: Configuration that is the source of Infrared unsafety in Midpoint algorithm with the diagram in the right showing the extra stable cone that can appear with the addition of a new soft seed. From [39]

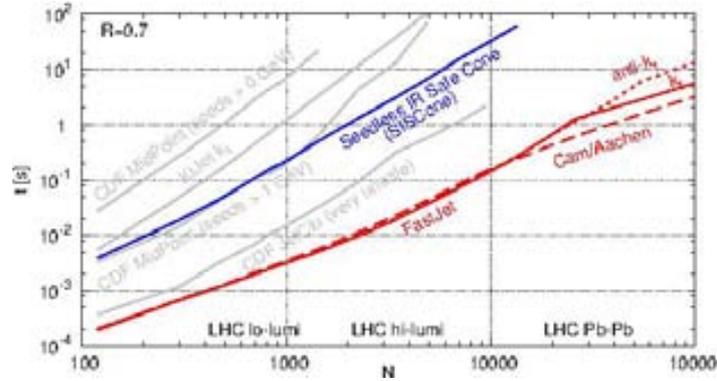


Figure 2.32: Timing for the different clustering of a simulated ~ 50 GeV dijet events. From [39]

where $\Delta R_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and $p_{t,i}$ is the transverse momentum of i object. There are three most important cluster algorithms that differ only based on the definition of the distance d reported before: for $p = -1$ anti kt algorithm, for $p = 0$ the Cambridge/Aachen algorithm and for $p = 1$ the kt one. All of these are Infrared and Collinear safe. As seen in Figure 2.33 the anti kt algorithm is more cone-like, it is infrared and collinear safe and it is not sensible to pile up. The computer time to build it is the lowest among the algorithms in use at hadron colliders. For this reason it was chosen by ATLAS and CMS collaborations to be the standard one.

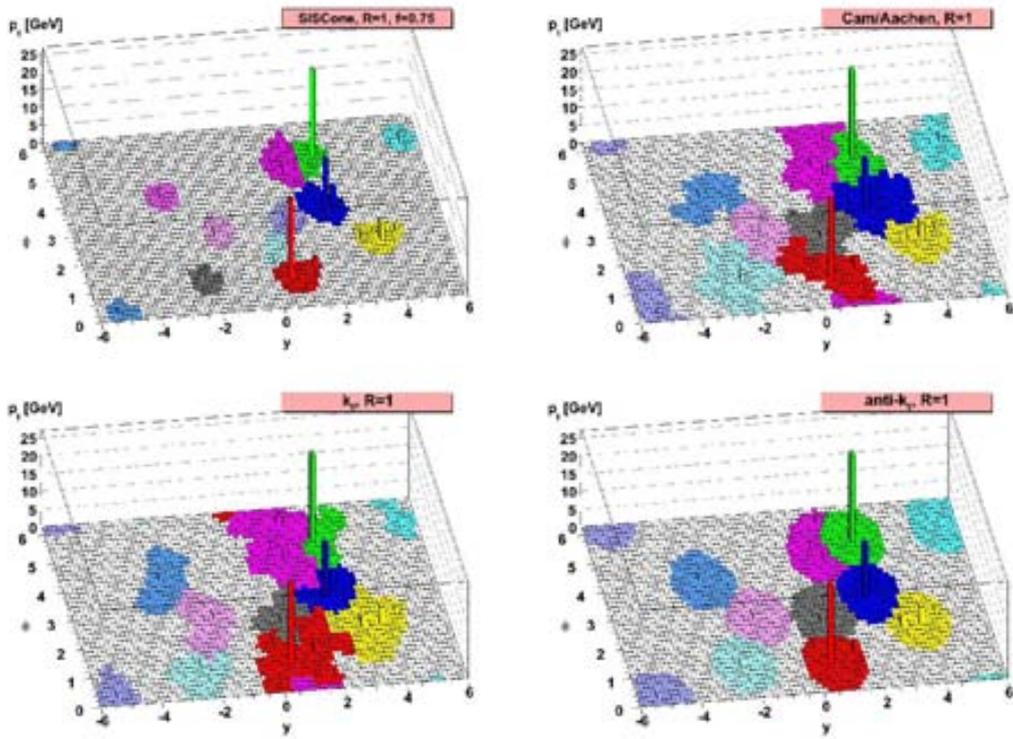


Figure 2.33: A sample parton-level event clustered with four different jet algorithm. Figure taken from [11]

3 $Z + b$ Theoretical predictions

The following section presents the challenges that calculation of $Z + b$ -jets final states have to be compared with and the tools developed for such purpose.

3.1 Introduction: the 4FNS and the 5FNS scheme

Processes involving b quarks are generally described in QCD by two theoretical schemes: the five flavor scheme (5FNS) or variable scheme (VFS) and the four flavor scheme or fixed scheme (4FNS or FFS).

In the FFS or 4FNS only 4 massless-quark densities are considered in the initial state and non-zero mass b quarks are arising in final states through a gluon splitting, while in a VFS or 5FNS¹ an initial state massless (ACOT approximation) b quark density is introduced. This b quark density is considered to be coming from a gluon splitting $g \rightarrow b\bar{b}$ where one heavy quark remains at low p_T and it is implicitly integrated out, while the other participates in the hard scattering and emerges at high- p_T . Thus its distribution function can be evaluated perturbatively with the DGLAP equation.

The two approaches are equal at all orders in perturbative theory but may give very different results at finite order and both schemes have been used to perform $Z + b$ predictions. Moreover one study is ongoing to merge the two configuration for the $Z + 1b$ jet calculation.

Below is a review of the predictions available for the different processes:

- $Z + 1$ jet+X: Z plus one single jet with one or more heavy flavor quarks. The complete details of the calculation can be found in [44]. It is performed with MCFM in the 5FNS scheme. This is perfectly suitable to describe $Z + 1$ jet events with 1 b -tag;
- $Z + 2$ jets+X: Z plus two jets with one or more heavy flavor quark, the calculation is done with MCFM and it is described in [48]. This is suitable for $Z + 2$ jets events with 1 b -tag.
- $Z + b\bar{b}$ NLO: Z in association with a bottom-antibottom pair. The calculation of the NLO radiative corrections is done in the massless hypothesis with MCFM and in 4FNS scheme [49]
- $Z + b\bar{b}$ NLO in 4FNS and in non-zero mass approach [50]
- $Z + b\bar{b}$ NLO in 4FNS with $m_b > 0$ plus parton shower, *aMCatNLO* [51].

¹Since the scale is much larger than the mass of the b quark the VFS and 5FNS coincide, for best explanation of VFS please refer to [46]. The b quark mass may be neglected for the b quark at initial state.

3 $Z + b$ Theoretical predictions

All these three processes are different since the Feynman diagrams that contribute in the calculation are also different. In the following sections we will analyze each process, showing the LO and NLO subprocesses.

3.2 NLO Associated production of a Z boson and a Single Heavy quark Jet

The main Leading-Order contribution for this process is $gb \rightarrow Zb$ and the tree Feynman diagrams are shown in Figure 3.1. They are of the order α_s in the 5FNS scheme since the b quark is present in the initial state.



LO diagrams for $gb \rightarrow Zb$. The massless b quark is in the initial state.

This process can be evaluated at NLO (α_s^2) taking into account 1 loop corrections and real corrections. The principal contributions are coming from the following subprocesses:

- $gb \rightarrow Zb$ 1 loop
- $qb \rightarrow Zbq$
- $gb \rightarrow Zbg$
- $gg \rightarrow Zbb$

Another process that contributes to $Z + b$ is $q\bar{q} \rightarrow Zb\bar{b}$ (Figure 3.2), where one or two b-jet can be detected. One b-jet detected can occur if the two quarks are collinear and they end up in the same jet or if one b jet falls outside the coverage of the detector. In this case it is necessary to introduce the mass of the quark ² to regulate the divergence arising from a gluon splitting.

It is worth nothing that the processes $q\bar{q} \rightarrow Zb\bar{b}$ correspond to two different Feynman diagrams: one when the Z is emitted from the initial state quarks and the other when Z is emitted from the final ones. In the former case it is more probable to have two high- p_T (tagged) jets, while in the latter is more probable that the two b quarks end up in the same jet.

²Due to its complexity in MCFM this process is available only at LO

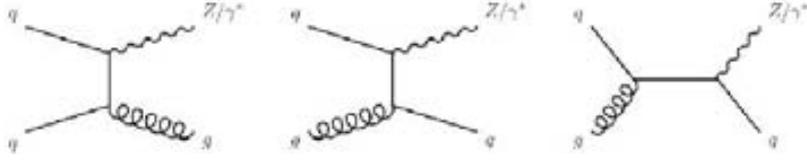


Figure 3.2: LO diagrams for $qq \rightarrow Zbb$

Therefore in MCFM the jet cross section estimation of a Z boson and at least one jet with one or more b quarks is given by:

$$\sigma_{Z+b\text{-jet}} = \sigma_{gb \rightarrow Zb} + \sigma_{q\bar{q} \rightarrow Zbb, 1tag} + C_{DC} + 2 \cdot \sigma_{q\bar{q} \rightarrow Zbb, 2tag} \quad (3.1)$$

where $\sigma_{gb \rightarrow Zb}$ is calculated at NLO, C_{DC} is a correction for double-counting and $\sigma_{q\bar{q} \rightarrow Zbb}$ are estimated at LO in b -mass hypothesis. In that way the LO component gives a prediction quite sensitive to the scale uncertainty.

In Alpgen this cross section is calculated considering only the tree level diagrams and in the b -quark mass hypothesis (4FNS or Massive ME). The complete sample is generated through the different multiplicity sub-samples: $Z + b\bar{b}$, $Z + b\bar{b} + 1jet$ up to almost three light jets in the final state.

3.3 $Z + 2$ jets with one b -quark jet

The $Z + 2$ jet with one or more b quarks is just an extension of the previous one³. Having one more parton in the final states, the processes that contribute to it are :

- $bg \rightarrow Zbg$ and $bg \rightarrow Zbq$
- $gg \rightarrow Zbb$ and $q\bar{q} \rightarrow Zbb$ (Figure 3.3)

The main contributions at NLO are coming from the processes listed above and from their real and virtual corrections:

- $q\bar{q} \rightarrow Zbb$ at tree level and one loop;
- $gg \rightarrow Zbb$ at tree level and one loop
- $bq \rightarrow Zbq$ at tree and one loop;
- $bg \rightarrow Zbg$ at tree level and 1 loop;

³The same scheme, 5FNS, is used.

3 $Z + b$ Theoretical predictions



Figure 3.3: Diagrams contributing to the associated production of a Z boson and two high- p_T jets, one of which contains a heavy quark. These are of the second order of α_s .

- $q\bar{q} \rightarrow Zb\bar{b}g$ at tree level and in non-zero mass hypothesis;
- $gg \rightarrow Zb\bar{b}g$ at tree level;
- $bg \rightarrow Zbgg$ at tree level;
- $bq \rightarrow Zbgg$ at tree level;
- $gq \rightarrow Zb\bar{b}q$ at tree level and in non-zero mass hypothesis;
- $bg \rightarrow Zb\bar{q}q$ at tree level;

All these processes are evaluated in the massless hypothesis except those where two b quarks end up in the same jet.

In particular, a few diagrams of real corrections are shown in Figure 3.4. One cannot uniquely identify them with any of the leading order processes, since they all contribute to the same amplitude and therefore they interfere with each other.

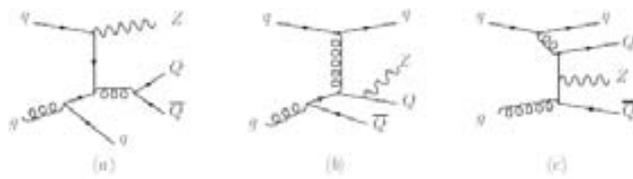


Figure 3.4: Real NLO corrections for $gb \rightarrow Zbg$.

3.4 $Z +$ two high p_T b-quark jet

The principal contribution of $Z +$ two b -jets final state is coming from $gg \rightarrow Zb\bar{b}$ and from $q\bar{q} \rightarrow Zb\bar{b}$. Some diagrams including real are found in Figure 3.5.

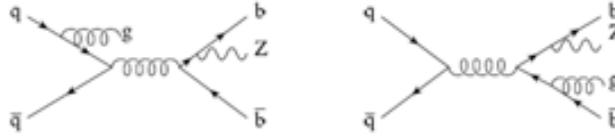


Figure 3.5: Diagrams of some real corrections for the two processes.

It is particularly challenging to predict the cross section for Z with two high p_T b -tag jets. One prediction can be done with MCFM in 4FNS massless approach, ignoring the low p_T phase space. Another is presented in Febres Cordero et al. [50] work, this prediction is based on the 4FNS scheme with non-zero mass hypothesis. Finally, it is interesting the approach of aMCatNLO that includes the NLO correction, bottom quark mass effects, spin correlation, showering and hadronization. Taking into account the b mass, it is possible to estimate the cases in which one of two b is not observed and can have small transverse momenta.

4 The Experimental Environment

The events analyzed in this thesis were produced as a result of proton - antiproton collisions at a center-of-mass energy of 1.96 TeV at CDF (Collider Detector at Fermilab), one of the two general purpose detectors within the Tevatron ring.

In this chapter, the CDF II detector and the Tevatron accelerator chain will be described in the detail.

4.1 The Fermilab Tevatron collider

The Tevatron [54] is an underground circular proton-synchrotron with 1 km of radius and it is the last stage of the accelerator system (Figure 4.1) located at Fermi National Accelerator laboratory (Fermilab) in Chicago (IL, USA). Before the LHC started, the Tevatron was the most powerful hadron collider in the world. While operating in collider mode bunches of protons, circulating clockwise and spaced by 396 ns, collide against a similar beam of antiprotons accelerated in the opposite direction, both at energies of 980 GeV.

The Tevatron performance, as a collider, is evaluated in terms of two parameters: the available center-of-mass energy \sqrt{s} , and the instantaneous luminosity, \mathcal{L} . The former defines the accessible phase-space for the production of particles in the final states, while the latter is

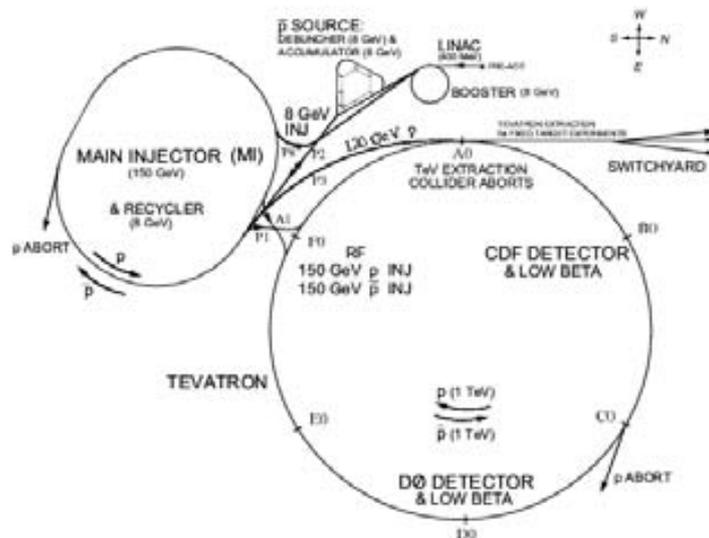


Figure 4.1: Illustration of the complete accelerator chain at Fermilab.

4 The Experimental Environment

the coefficient of proportionality between the rate of a given process $\frac{dN}{dt}$ and its cross section σ , as expressed by the following formula:

$$\frac{dN}{dt}[\text{events } s^{-1}] = \mathcal{L}[cm^{-2}s^{-1}] \times \sigma[cm^2]$$

The time-integral is therefore a measure of the expected number of events $N(T)$ produced in a finite time T :

$$N(T) = \int_0^T \mathcal{L} \sigma dt$$

Assuming an ideal head-on $p\bar{p}$ collision with no crossing angle between the beams, the instantaneous luminosity is defined as:

$$\mathcal{L} = 10^{-5} \frac{N_p N_{\bar{p}} B f \beta \gamma}{2\pi \beta^* \sqrt{(\epsilon_p + \epsilon_{\bar{p}})_x (\epsilon_p + \epsilon_{\bar{p}})_y}} F(\sigma_z / \beta^*)$$

where N_p ($N_{\bar{p}}$) is the average number of protons (antiprotons), B is the number of circulating bunches, f is the revolution frequency, $\beta\gamma$ is the Lorentz relativistic factor and $F(\sigma_z / \beta^*)$ is an empiric hourglass factor, which is a function of the ratio between the longitudinal r.m.s. width of the bunch (σ_z) and the beta function calculated at the interaction point (β^*), and the 95 % normalized emittance of the beams ($\epsilon_p \sim 18\pi$ mm mrad and $\epsilon_{\bar{p}} \sim 13\pi$ mm mrad after injection)¹.

The main parameters of Tevatron accelerator are summarized in the Table 4.1.

Parameter	value
energy of center-of-mass (\sqrt{s})	1.96 TeV
number of bunches (B)	36
space between bunches	396 ns
width of the bunch (σ_z)	60 cm
bunch average number of protons (N_p)	3×10^{11}
bunch average number of antiprotons ($N_{\bar{p}}$)	3×10^{10}
beta function (β^*)	31 cm
luminosity peak	$4.08 \times 10^{32} cm^{-2}s^{-1}$

Table 4.1: Summary of the main Tevatron characteristics.

The limiting factor for the luminosity is the capability to create a monochromatic beam of antiprotons that can be transmitted efficiently without dispersions into the entire accelerator

¹The hourglass factor is a parameterization of the longitudinal profile of the beams in the collision region, which assumes the shape of an horizontal hourglass centered in the interaction region. The beta function is a parameter convenient for solving the equation of motion of a particle through an arbitrary beam transport system. The emittance ϵ measures the phase-space occupied by the particles of the beam. Three independent two dimensional emittances are defined. The quantity $\sqrt{\beta\epsilon}$ is proportional to the r.m.s. width of the beam in the corresponding phase plane.

chain.

In the following paragraph the proton/antiproton production and the several acceleration steps to reach the energy of 980 GeV are explained in more detail.

4.1.1 Proton and Antiproton production

The proton production begins with hydrogen ionization: hot hydrogen gas is passed through a magnetron, which extracts a 50-55 mA current of 15-22 keV H^- ions, subsequently accelerated to 750 keV by a Cockcroft-Walton accelerator. The hydrogen ions beam, segmented into bunches, is then injected into a 150 m long Linac where hydrogen ions increase their energy up to 401.5 MeV before injection into Booster.

The Booster is an alternating gradient synchrotron with an orbit of 85.5 m that accelerates protons to 8 GeV in 33 ms. At injection, a thin carbon foil is used to strip the electrons from the H^- ions to obtain protons. Injecting H^- ions rather than protons into the Booster allows the injection to proceed over multiple revolutions of the beam around the Booster ring. If protons were used instead, the magnetic field used to inject new protons onto orbit in the Booster would also deflect the already revolving protons out of orbit.

Here proton and antiproton production processes become different and two basic modes are characteristic during the collider operations: antiproton accumulation and injection in the main ring.

In the antiproton production, one set of 84 proton bunches is extracted from the Booster at 8 GeV and injected into the Main Injector every 2.2 s. The Main Injector, a circular synchrotron, accelerates the protons up to 120 GeV. These are extracted and directed to impact against a rotating 7 cm thick target.

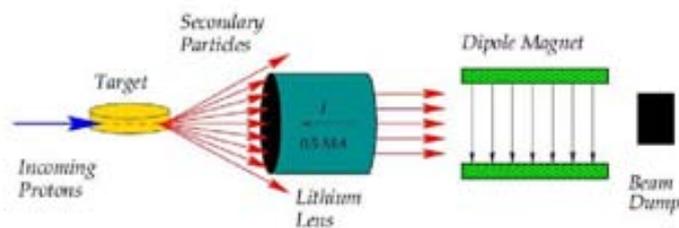


Figure 4.2: Antiproton production.

The particles produced in the interaction are spatially wide spread. They are collected and focused with a cylindrical lithium lens (Figure 4.2). 8 GeV/c negatively charged secondary particles are selected in momentum by a 1.5 T pulsed dipole magnet. The antiprotons created are delivered to the Debuncher storage ring, a triangular synchrotron that transforms the antiproton pulses in a continuous beam of monochromatic antiprotons. Stochastic cooling [55], electron cooling [56] and bunch rotation are applied during many cycles to collimate the beam. From the Debuncher antiprotons are transferred with 60 -70 % efficiency into the

4 The Experimental Environment

Accumulator where they are stacked and cooled with a variety of systems until the maximum antiproton intensity is reached. Then they are sent to Recycler that is stored in the same ring as the Main Injector and that is useful to maintain the antiproton momenta at 8 GeV, “stacking” the antiprotons which can then be injected into the Tevatron.

4.1.2 Injection and collisions

Every 10-20 h, antiproton accumulation is stopped in preparation for injection. A set of seven proton bunches is extracted from the Booster, injected into the Main Injector accelerated to 150 GeV, coalesced into a single bunch of 300×10^9 protons and then injected into the Tevatron. This process is repeated every 12.5 s, until 36 proton bunches, separated by 396 ns, are loaded into the Tevatron central orbit. Then four sets of 7-11 \bar{p} bunches are extracted from the Recycler to the Main Injector, accelerated to 150 GeV, coalesced into four $\sim 30 \times 10^9 \bar{p}$ bunches separated by 396 ns, and then injected into the Tevatron. Protons and antiprotons circulate in the same beam-pipe, sharing magnet and vacuum system. The injection process is repeated nine times until 36 antiproton bunches circulate in the Tevatron.

Sweeping the Tevatron RF by $\sim 1kHz$, the beam is then accelerated in about a minute from 150 to 980 GeV. Once the final energy is reached the two counter-rotating particle beams pass through each other colliding at the two instrumented interaction-points located along two straight sections of the Tevatron: D0 and B0, where the D0 and CDF II detectors respectively are situated. This stable situation of 980 GeV proton-antiproton collisions is called a store.

4.1.3 Tevatron performance

Since the beginning of Run II (2001) the Tevatron performance has been steadily increasing until its end of the activity (September 2011) when more than $10 fb^{-1}$ of data have been collected for each experiment. The plot in Figure 4.3 shows the integrated luminosity since the beginning of Run II.

4.2 The CDFII Detector

The CDFII detector [57] is a large multi-purpose solenoid magnetic spectrometer surrounded by 4π fast, projective calorimeters and fine-grained muon detectors. It is used to record the interactions resulting from the proton-antiproton collisions at a center-of-mass energy of 1.96 TeV. It is a detector designed to measure the energy, momentum and the identity of particles produced in Tevatron collisions combining all information coming from the different sub-detectors. A cross sectional view of half the detector is shown in Figure 4.4.

Particles produced in the collisions (Figure 4.5) first pass through the tracking detectors where the momentum of charged particles is measured from their curvature, after that they cross the calorimeters, where the energy of electrons and hadrons are deposited. At the end a few of these particles reach the external part where muon detectors detect the passage of any

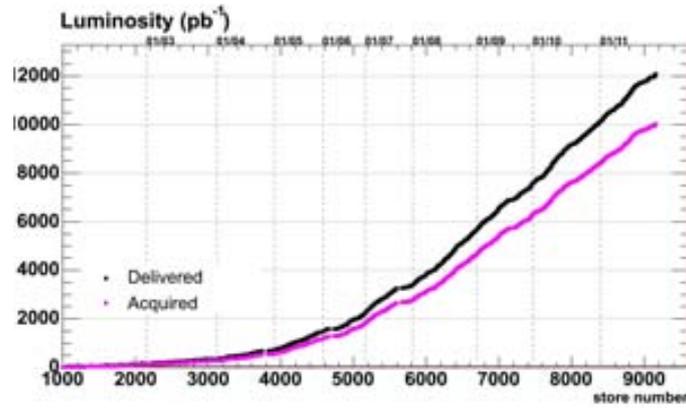


Figure 4.3: Integrated luminosity as a function of time for Run II data-taking. In black curve there is the delivered luminosity while the pink one represents the acquired luminosity, stored on tape.

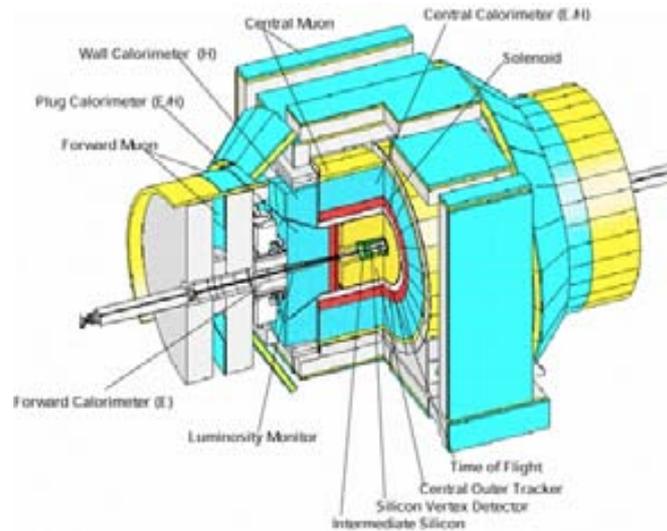


Figure 4.4: CDF II detector sketch in three dimensions.

4 The Experimental Environment

charged particles that escapes from the calorimeter. The combined responses of the various detectors permits to identify the different particles.

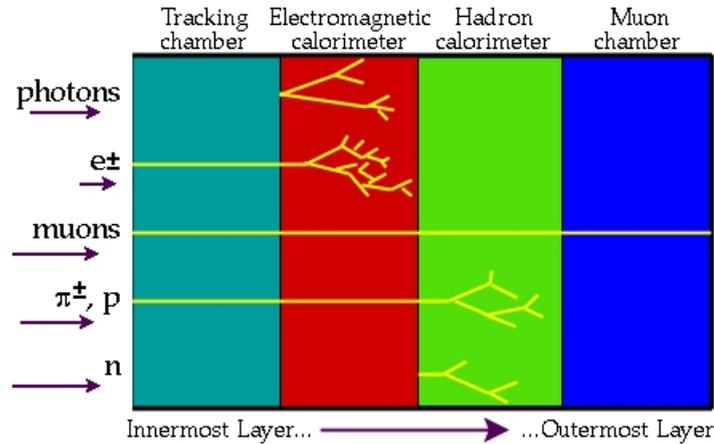


Figure 4.5: Particles identification. The passage of different kinds of particles through the CDFII sub-detectors. Combining all the information coming from the several sub-detectors we can identify whether the particle is a photon, an electron, a jet or a muon.

The tracking systems are contained in a super-conducting solenoid of 1.5 m in radius and 4.8 m in length that generates a 1.4 T magnetic field parallel to the beam axis. The calorimeter and the muon system are outside the solenoid.

The main detector characteristics are an excellent tracking performance, which provides high mass resolution and precisely reconstructed decay vertexes, good electron and muon identification capabilities combined with charged-hadron identification, and an advanced trigger system that fully exploits the high rate events.

Before explaining in detail the several sub-detectors it is worthwhile defining the coordinate system used at CDFII.

4.2.1 The CDFII Coordinate system

The CDFII detector uses a right-handed Cartesian coordinate system with the origin in the B0 interaction point and where the $+z$ -axis lies along the nominal beam-line pointing toward the proton direction (east). The (x, y) plane is therefore perpendicular to either beams, with positive y-axis pointing vertically upwards and positive x-axis in the horizontal plane of the Tevatron, pointing radially outward respect to the center of the ring.

Since the colliding beams of the Tevatron are unpolarized, the resulting physical observation is invariant under rotations around the beam line axis, for this reason a cylindrical coordinate system is frequently used to describe the detector geometry. Longitudinal and transverse means respectively parallel and perpendicular to the proton beam direction.

In hadron collision environments, it is common to use a variable invariant under z Lorentz

boosts as unit of relativistic phase-space, instead of polar angle θ . This variable, called rapidity Y , is defined as:

$$Y = \frac{1}{2} \log \left[\frac{E + p \cos \theta}{E - p \cos \theta} \right] \quad (4.1)$$

where (E, p) is the energy four-vector of the particle. However the problem with the rapidity is that its measurement still requires an accurate particle identification capabilities because of the mass term entering E . For practical reasons it is preferably to substitute Y with its approximate expression $\eta = -\log[\tan(\theta/2)]$, called pseudorapidity. They are equal in the ultra relativistic limit.

As the event by event longitudinal position of the actual interaction is distributed around the nominal interaction point with 30 cm r.m.s. width, it is useful to distinguish detector pseudo-rapidity, η_{det} , measured with respect to the (0,0,0) nominal interaction point, from particle pseudo-rapidity, η , which is measured with respect to the z_0 position of the real vertex where the particle originated.

4.2.2 The tracking system

The CDFII tracking system (Figure 4.6) is designed to reconstruct the three-dimensional charged particles tracks with high resolution and precision. The system consists in three silicon sub-detectors and in a drift chamber (COT). It is located inside a super-conducting solenoid which generates a 1.4 T magnetic field parallel to the beam axis.

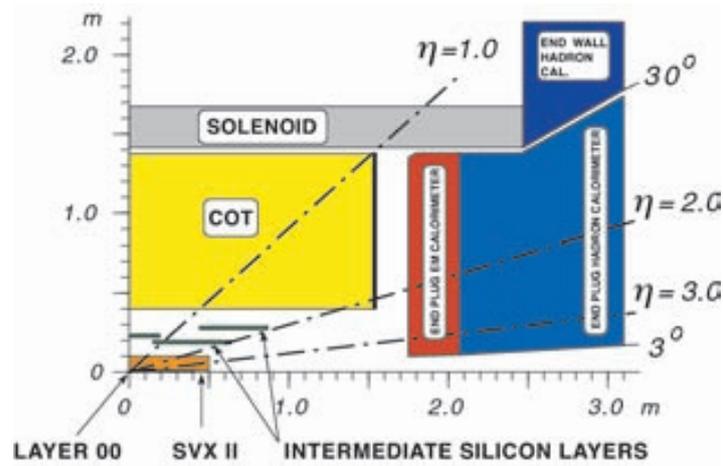


Figure 4.6: CDFII Tracking system. Longitudinal cross-sectional of the detector, showing the tracking system and the plug calorimeters.

Silicon detectors

The CDFII silicon detectors are designed to perform high precision tracking, which is very important for identifying long-lived particles, such as B hadrons. These B hadrons can travel several millimeters before decaying into several particles, and the precise reconstruction of the charged particles allows the extrapolation of their trajectories to find a common decay origin (secondary vertex) that is well displaced from the location of proton-antiproton collision (primary vertex).

The CDFII silicon detectors are composed of silicon micro-strip sensors that can be divided into three sub-detectors (Figure 4.7). The core is the Silicon Vertex (SVXII), then in the outer part there is the Intermediate Silicon Layers (ISL) while in the inner part there is the Layer 00. L00 is a light-weight silicon layer placed on the beam pipe. It recovers the degradation in resolution of the reconstructed vertex position due to the multiple scattering on the SVXII read-out electronics and cooling system, installed within the tracking volume. Micro-strip allows precise measurement and is based on the p-n junction that creates localized region where electric charges are formed by the passage of charged particles. The resolution is given from the distance d between the strip and it is around $d/12$.

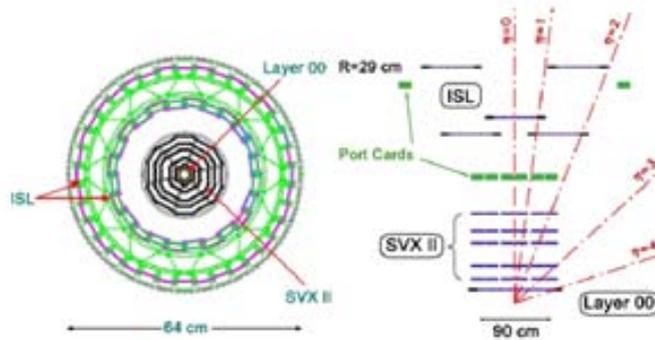


Figure 4.7: Silicon tracking sub-detectors projected in transversal and (r, z) plane.

L00

Starting from the center of the detector there is the L00 that consists of a single castellated layer of single-sided, AC-coupled silicon sensors mounted directly on the beam pipe at radii, alternating in ϕ , of 1.35 cm or 1.62 cm from the beam. It provides full azimuthal and $|z| \leq 47$ cm longitudinal coverage. The strips are parallel to the beam axis allowing sampling of tracks in the (r, ϕ) plane.

Silicon Vertex detector II

The SVXII (Figure 4.8) [58] is a fine resolution silicon micro-strip vertex detector which provides five three-dimensional sampling of tracks at 2.45, 4.1, 6.5, 8.2 and 10.1 cm of radial

distance from the beam with full pseudo-rapidity coverage in the $|\eta_{det}| \leq 2$ region. It has a cylindrical geometry coaxial with the beam, and its mechanical layout is segmented in three axial sections of 32 cm, called barrels. Moreover each radial layer is divided in twelve 30° parts, called wedges.

Sensors in a single layer are arranged into independent longitudinal read-out units, called ladders. Each ladder comprises two, double-sided sensors and a multi-layer electronic board, all attached on a carbon-fiber support.

The active surface consists of double-sided, AC-coupled silicon sensors with micro-strips implanted on a $300 \mu\text{m}$ thick, high resistivity bulk. On one side, all sensors have axial strips spaced approximately $60\text{--}65 \mu\text{m}$, for a precise reconstruction of θ coordinate. On the reverse side, the following combination of read-out pitch is used: $141 \mu\text{m}$ (90°), $125.5 \mu\text{m}$ (90°), $60 \mu\text{m}$ (1.2°), $141 \mu\text{m}$ (90°), $65 \mu\text{m}$ (-1.2°) from the innermost to the outermost layer for reconstructing the z-coordinate. The complete features of each layer are summarized in Table 4.2.

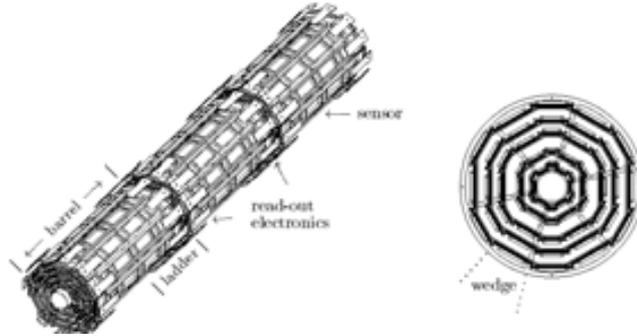


Figure 4.8: Schematic illustration of the three instrumented mechanical barrels of SVXII detector and, on the right, of the cross-section of a SVXII barrel in the (r, ϕ) plane.

Intermediate Silicon Layer

The ISL (Figure 4.9) [61] is a silicon tracker placed at intermediate radial distance between the SVXII and the drift chamber. At $|\eta_{det}| \leq 1$ a single layer of silicon sensors is mounted on a cylindrical barrel at radius of 22.6 cm. At $1 \leq |\eta| \leq 2.0$ two layers of silicon sensors are arranged into two pairs of concentric barrels (inner and outer). In the inner (outer) barrel, staggered ladders alternate at radii of 19.7 and 20.2 cm (28.6 and 29.0 cm). One pair of barrels is installed in the forward region, the other one is in the backward region. Each barrel is azimuthally divided into a 30° structure matching the SVXII segmentation. Each sensor has axial strip space by $112 \mu\text{m}$ on one side and 1.2° angled strip spaced $112\text{--}146 \mu\text{m}$ on the reverse.

4 The Experimental Environment

Propriety	Layer 0	Layer 1	Layer 2	Layer 3	Layer 4
number of strip ϕ	256	384	640	768	896
number of strip Z	256	576	640	512	896
number of ϕ chip	2	3	5	6	7
number of Z chip	2	3	5	4	7
stereo angle	90°	90°	1.2°	90°	-1.2°
pitch ϕ strip (μm)	60	62	60	60	65
z pitch strip (μm)	141	125.5	60	141	65
total arm length (mm)	17.140	25.594	40.300	47.860	60.170
total long extension (mm)	74.3	74.3	74.3	74.3	74.3
active length (mm)	15.300	23.746	38.340	46.020	58.175
active longitudinal extension (mm)	72.43	72.3	72.38	72.43	73.43
number of detectors	144	144	144	144	144

Table 4.2: Characteristics of SVX II layers

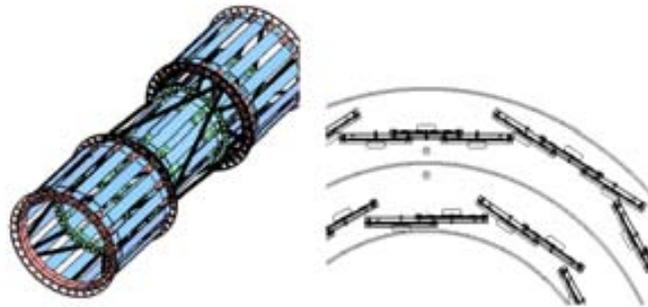


Figure 4.9: Intermediate Silicon Layer cartoon

Thanks to these three silicon sub-detector an excellent identification of secondary vertexes is possible with a $\sim 40 \mu m$ resolution on the impact parameter.

4.2.3 Central Outer tracker

The Central Outer Tracker (COT) [63], in Figure 4.10, is a cylindrical drift chamber located outside the silicon detectors at a radius from 40 cm to 137 cm and covers $|\eta_{det}| \leq 2$. It consists of 8 super-layers: 4 parallel to the beam-line (axial super-layers) and 4 with an angle $\pm 2^\circ$ with respect to the z axis (stereo). Each super-layer is made of varying number of cells (for instance super-layer 1 has 169 cells and super-layer 8 has 480 cells). Each cell consists of a field sheet and a wire plane with alternating sense wires and field wires. Each cell has a total of 12 sense wires. The chamber is filled with a 50:50 mixture of Argon and Ethan gas 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 of 35° . Super-cells are tilted by 35° with respect to the radial direction to compensate this effect.

Charged particle passing through the COT interacts and ionizes the gas mixture. Positively charged ions and free electrons are created. If an electric field is applied in the gas volume, electrons will drift toward the anode. In the high field region near the anode, the electron ionizes other atoms and produces an avalanche, which creates a large signal on the wire. Electrons, so created, are collected on the anode wire giving an indication of the passage of a particle near that volume. Electrons drift faster than ions due to their lower mass. The electron drift velocity depends on the electric field gradient and on the properties of the gas molecules, which for the COT is $\sim 50 \mu m/s$. Usually signals deposited by a particle are collected in less than 200 ns.

The COT is useful in measuring the momentum of the charged particles. Since the COT is placed in the 1.4 T magnetic field, charged particles travel in a helix with a radius $r = \frac{p_T}{|q|B}$ where p_T is the transverse momentum, q the particle charge and B the magnetic field. By reconstructing the track's curvature in the $r - \phi$ plane, p_T can be determined.

The technical properties of the tracker sub-detector are summarized in Table 4.3.

4.3 Track reconstruction

The arc of the helix, in Figure 4.11, described by a q charged particle in the magnetic volume of CDFII is parameterized using the following five variables, among which three are transverse and the other two are longitudinal:

C - signed helix half-curvature, defined as $C = \frac{q}{2R}$, where R is the radius of the helix. This is directly related to the transverse momentum: $p_T = \frac{cB}{2|C|}$;

φ_0 - ϕ direction of the particle at the point of closest approach to the z-axis;

4 The Experimental Environment

Layer 00	
r	from 1.35 to 1.65 cm
resolution	6 μm (axial)
number of channels	13824
SVX II	
r	from 2.4 a 10.7 cm
number of <i>layer</i>	5
read-out coordinates	r- ϕ one side for <i>layer</i>
other coordinates	r-z, r-z, r-uv, r-z, r-uv (uv=1.2°)
pitch resolution	60-65 μm r - ϕ , 60-150 μm stereo
resolution	12 μm (axial)
total length	96.0 cm
rapidity	$ \eta \leq 2.0$
number of channels	405 504
ISL	
r	from 20 to 28 cm
number of <i>layer</i>	one for $ \eta \leq 1$, two for $1 \leq \eta < 2$
read-out coordinates	r- ϕ and r - uv (1.2° stereo) for all <i>layer</i>
pitch resolution	10 μm (axial), 146 μm (stereo)
resolution	16 μm (axial)
total length	174 cm
rapidity	$ \eta \leq 1.9$
number of channels	268 800
COT	
r	from 44 to 132 cm
Number <i>super-layers</i>	8
Cells for <i>super-layer</i>	12
read-out Coordinates	+20°,-20°, +20°, +20°
drift distance	0.88 cm
resolution	180 μm
rapidity	$ \eta \leq 2.0$
number of channels	30340

Table 4.3: Summary of the technical details of the tracker sub-detectors.

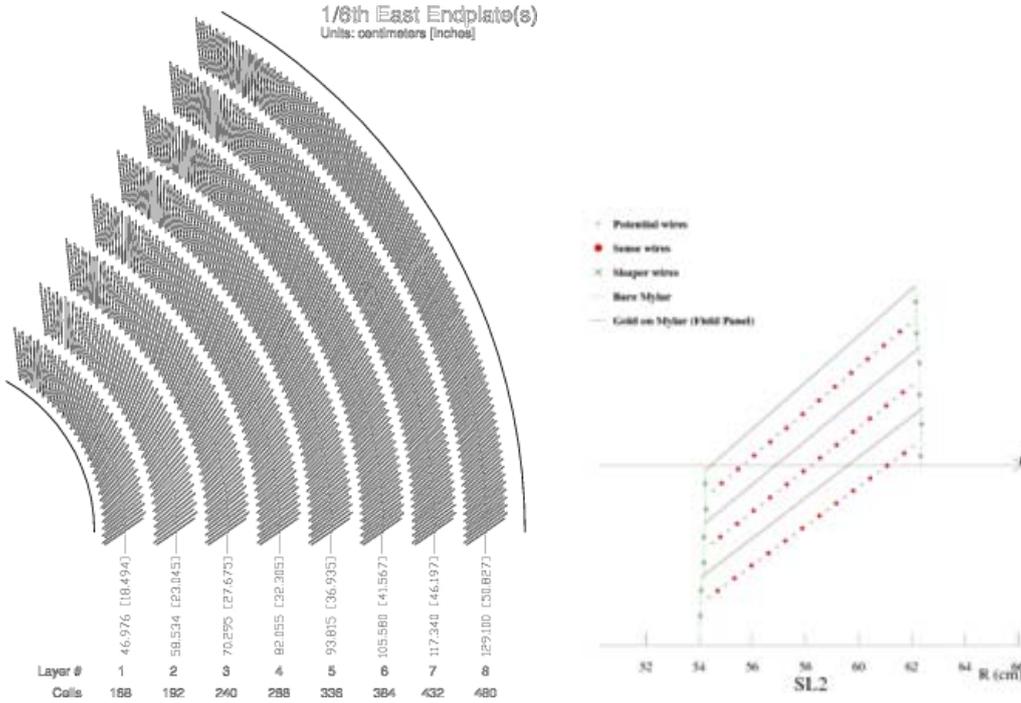


Figure 4.10: On the right: 1/6 of the COT end-plate. On the left: sketch of a axial cross-section of three cells in super-layer 3. The arrow shows the radial direction.

d_0 - signed impact parameter, i.e. the distance of the closest approach to the z-axis, defined as $d_0 = q(\sqrt{x_c^2 + y_c^2} - R)$, where (x_c, y_c) are the coordinates of the center-guide;

λ - the helix pitch, i.e. $\cot(\theta)$, where θ is the polar direction of the particle at the point of its closest approach to the z-axis. This is directly related to the longitudinal component of the momentum: $p_z = p_T \cot \theta$;

z_0 - the z coordinate of the point of closest approach to the z axis

The reconstruction of a charged particle trajectory consists in determining the above parameters through an helical fit of a set of spatial measurements (hits) reconstructed in the tracking detectors by clustering and pattern-recognition algorithms. The helical fit takes into account field non-uniformities and scattering in the detector materials.

The COT efficiency for tracks is typically 99 % and the single hit resolution is $140 \mu m$. The typical resolutions of track parameters are the following: $\sigma_{p_T}/p_T^2 \sim 0.0015 (GeV/c)^{-1}$, $\sigma_{\phi_0} \sim 0.035^\circ$, $\sigma_{d_0} \sim 250 \mu m$, $\sigma_{z_0} \sim 0.3 cm$. [67]. Including the silicon information improves the impact parameter resolution of tracks which, depending on the number of the silicon hits, may reach $20 \mu m$. This value combined with the $\sigma_T \sim 30 \mu m$ transverse beam size is sufficiently small with respect to the typical transverse decay length of heavy flavors to allow the separation of their decay vertexes from production vertexes. The silicon tracker

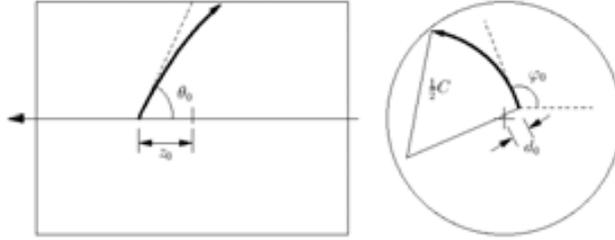


Figure 4.11: Track Reconstruction coordinates.

improves also the stereo resolutions up to $\sigma_{z_0} \sim 70 \mu m$, while the transverse momentum and azimuthal resolutions remain approximately the same as COT only tracks. [68]

The comparison between resolutions of tracks reconstructed with only COT informations and with silicon+COT is shown in Table 4.4.

Parameter	COT	COT+SVX II+ISL
$\sigma_{p_T/p_T^2} [(GeV/c)^{-1}]$	0.0015	10^{-3}
$\sigma_d [\mu m]$	250	20
$\sigma_{z_0} [\mu m]$	300	70
$\sigma_{cot\theta}$	0.17°	0.06°

Table 4.4: Track resolutions using COT only or Silicon information plus COT.

4.4 Time of Flight detector

Between the COT and the super-conducting solenoid there is a time-of-flight detector (TOF) [69], located at $r \sim 140$ cm from the beamline. It is a cylindrical array made of 216 scintillating bars of almost 3 m of longitude and located at $r \sim 140$ cm. Both longitudinal sides of the bars collect the light pulse into a photo-multiplier and measure accurately the timing of the two pulses. The time between the bunch crossing and the scintillation signal in these bars defines the β of the charged particle while the momentum is provided by the tracking system. Particle identification (PID) information is available through the combination of TOF information and tracking measurements. The measured mean time resolution is 110 ps. This guarantees a separation between charged pions and kaons with $p_T \leq 1.6$ GeV/c equivalent to 2σ , assuming Gaussian distributions. Unfortunately, in high luminosity conditions ($\mathcal{L} \geq 5 \times 10^{31} \text{ cm}^{-2}\text{s}^{-1}$) the occupancy of the single bars determines a degradation in efficiency, which is about 60 % per track.



Figure 4.12: On the right view of the Time of Flight. On the left particular of the link between scintillator and photo-multiplier.

4.5 Calorimeters

Outside the solenoid, scintillator-based calorimetry covers the region $|\eta_{det}| \leq 3.6$, and is devoted to the measurement of the energy deposition of electrons, photons and hadrons using the shower sampling technique.

The basic structure consists in alternating layers of passive absorbers and a plastic scintillator. Neutral particles and charged particles with a $p_T \geq 350 \text{ MeV}/c$ are likely to escape the solenoid's magnetic field and penetrate into the CDFII calorimeters. These are finely segmented in solid angle around the nominal collision point, and coarsely segmented radially outward from the collision point (in-depth segmentation). Angular segmentation is organized in projective towers. Each tower has a truncated-pyramidal architecture having the imaginary vertex pointing to the nominal interaction point. The base is a rectangular cell in the (η_{det}, ϕ) space. Radial segmentation of each tower instead consists of two compartments, the inner (closer to the beam) devoted to the measurement of the electromagnetic component of the shower, and the outer devoted to the measurement of the hadronic fraction of the energy. These two components are read independently through separated electronics channels.

A different fraction of energy released in the two compartments distinguished photons and electrons from hadronic particles. In total CDFII calorimetry consists of 1536 calorimeter towers. The light produced by the particles of the shower that cross the scintillating plate is collected by wavelength shifting (WLS) fibers that transport it to photo-multiplier tubes (PMT) located in the outermost part of the calorimeters. Every projective tower is read by one or two PMTs.

The sub-detectors that constitute the calorimeter of CDFII, are separated by the position with respect to the interaction point in two main groups: the central calorimeters, that approximately cover the region $|\eta_{det}| \leq 1.1$, and the plug calorimeters, that cover $1.1 \leq |\eta_{det}| \leq 3.6$. The central calorimeters consist of two separated halves that meet at $\eta_{det} = 0$. Due to this peculiar configuration, two gaps region exist around $\eta = 0$ and $\eta = 1.1$. Figure 4.13 shows the spacial disposition of the calorimeter and Table 4.5 lists the main characteristics of each calorimeter sub-detector.

4 The Experimental Environment

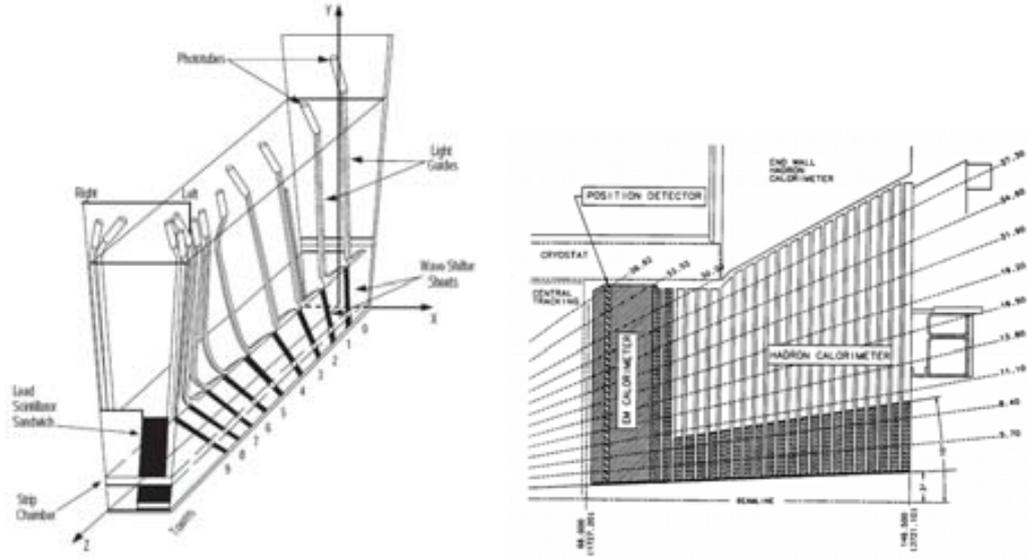


Figure 4.13: Schematic illustration of an azimuthal sector of the central electromagnetic calorimeter. On the left elevation view of one quarter of the plug calorimeter.

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

Table 4.5: The CDF II calorimeters with their acronym, η region, segmentation, thickness and energy resolution. X_0 represents the shower length and λ_I is the pion nuclear absorption length in $g \text{ cm}^{-2}$

4.5.1 Central region: CEM, CHA, WHA

The radial extension of the calorimeters in the central region is $1.73 \text{ m} \leq r \leq 3.5 \text{ m}$. The Central ElectroMagnetic Calorimeter (CEM) [70] [71] is constructed as four azimuthal arches (NE, NW, SE, SW), each of which subtends 180° and is divided into twelve 15° wedges. A wedge consists of 31 layers of 5 mm thick polystyrene scintillators interlayered with 30 aluminum-clad lead 3.2 mm thick sheets. It is divided along η_{det} into 10 towers. To maintain a constant thickness in X_0 (radiation length) and compensating the $\sin(\theta)$ variation between towers, the same lead layers are replaced with increasing amounts of acrylic as a function of η_{det} . The spacial resolution of the CEM is about 2 mm. The outer two towers in one wedge (known as chimney tower) are missing to allow solenoid access, for a resulting total number of 478 instrumented towers. At a radial depth of $5.9 X_0$, which is approximately the depth corresponding to the peak of shower development, the Central Strip multi-wire proportional chamber (CES) measures the transverse shower shape with 1.5 cm segmentation. A further set of multi-wire proportional chambers, the Central Pre-Radiator (CPR) [72] is located in the gap between the outer surface of the solenoid and the CEM. It monitors photon conversions started before the first CEM layer. Phototube gains are calibrated once per store using an automated system of Xenon or LED light flashers.

The hadronic compartment is the combination of two sub-systems: the Central HAdronic (CHA) and the Wall HAdronic (WHA) [73] calorimeters. Each CHA wedge is segmented into 9 η_{det} towers matching in size and position the CEM towers. The WHA wedge instead consists of 6 towers of which three are matching CHA towers. Radially a CHA tower is constructed of 32 layers of 2.5 cm thick steel absorber alternating with 1.0 cm thick acrylic scintillator. WHA tower structure is similar but there are only 15 layers of 5.1 cm thick absorber.

The total thickness of the electromagnetic section corresponds to approximately $18 X_0$ ($1 \lambda_I$, where λ_I is the pion nuclear absorption length in units of g cm^{-2}), for a relative energy resolution $\sigma_E/E = 14\%/\sqrt{E_T} \oplus 2\%^2$

4.5.2 Forward region: PEM and PHA

The coverage of the $1.1 \leq |\eta| \leq 3.6$ region relies on the scintillating tile Plug calorimeter [74] [75] [76] which is composed of two identical devices, one installed in $\eta_{det} \geq 0$ and the other in the $\eta_{det} \leq 0$. Each of these two halves has electromagnetic and hadronic compartments. In each half, the absorber of the Plug ElectroMagnetic calorimeter (PEM) consists of 23 doughnuts - shaped lead plates, 2.77 m in outer diameter, which have a central hole where the beam pipe is located. Each plate is made out of 4.5 mm thick calcium-tin-lead sandwiched between two 0.5 mm thick stainless-steel sheets. Between the absorber plates are inserted the 4 mm thick scintillator tiles organized azimuthally in 15° triangularly-shaped wedges. A Pre-shower detector consists of a thicker (10 mm) amount of scintillator installed in the first layer of PEM, while shower maximum sampling is performed at radial depth of $\approx 6X_0$ by two tilted layers of scintillator strips (pitch 5 mm).

²The first term is called the stochastic term and derives from the intrinsic fluctuations of the shower sampling process and of the PMT photo-multiplier yield. The second term, added in quadrature, depends on the calorimeter non-uniformities and in the uncertainty on the calibration. All energies are in GeV .

4 The Experimental Environment

Each half of the hadronic compartment, Plug HAdronic calorimeter (PHA), is azimuthally divided in 12 wedge-shaped modules each subtending 30° . In depth each module consists of 23 layers of 5 cm thick iron absorber alternated with 6 mm scintillator layers. Within each sampling layer the scintillator is arranged in tiles similar to those used in the PEM.

The total thickness of the electromagnetic section corresponds to approximately $23 X_0$ ($4.7 \lambda_I$), for an energy resolution of $\sigma_E/E = 16\%/\sqrt{E_T} \oplus 1\%$. The total thickness of the hadronic section corresponds to approximately $8 \lambda_I$, for an energy resolution of $\sigma_E/E = 75\%/\sqrt{E_T} \oplus 4\%$.

4.6 Muons systems

Muon system in Figure 4.14 is placed in the most outer part of the detector. This consists of scintillating counters and drift tubes installed at various radial distances from the beam to detect muons [77] [78]. Scintillators serve as a trigger and vetoes while the drift chambers measure the ϕ coordinate using the absolute difference of the drift electrons arrival time between 2 cells and the z coordinate by charge division.

These sub-detectors cover the whole range of pseudo-rapidity $|\eta_{det}| \leq 2$ and are used only to identify the penetrating muon reconstructing a small segment of their path (stub) sampled by the chambers. The moment measurement is performed by pointing back the stub to the corresponding track in COT.

Different muon sub-systems cover different geometrical regions. In the $|\eta_{det}| \leq 0.6$ region moving outward from the beam we encounter the inner Central MUon detector (CMU) chambers at radial distance of 3.5 m. Approximately $5.4 \lambda_I$ of material separate the luminous region from CMU resulting in about $1/220$ high energy hadrons passing through the calorimeter and reaching the muon detector. This defines also a p_T threshold for muons reaching the CMU which is approximately $1.4 GeV/c$. In order to recognize and discard them, the Central Muon uPgrade (CMP) chambers lie in the same η_{det} region separated radially from the CMU by a 60 cm thick wall of steel achieving a rejection of 95 % of the fake muons.

The muon coverage in the $0.6 \leq |\eta_{det}| \leq 1.0$ volume is ensured by the Central Muon eX-tension (CMX) chambers, embedded in scintillator counters and placed at radius of 3.5 m. The Intermediate MUon detectors (IMU) are instead drift tubes covering the pseudo-rapidity range $1.0 \leq |\eta_{det}| \leq 2.0$. CDFII triggers on muons only emerging at $|\eta_{det}| \leq 1.5$ where the muon coverage is segmented with sufficient granularity to survive high occupancies. The granularity of muon devices in the forward regions is less fine and not adequate for triggering, but sufficient for off-line muon assignment to high p_T tracks going through that region.

4.7 The Cerenkov Luminosity counters

The luminosity (\mathcal{L}) is inferred from the average number of inelastic interactions per bunch crossing (\bar{N}) according to :

$$\bar{N} \times f_{b.c.} = \sigma_{p\bar{p}-in} \times \epsilon \times \mathcal{L}$$

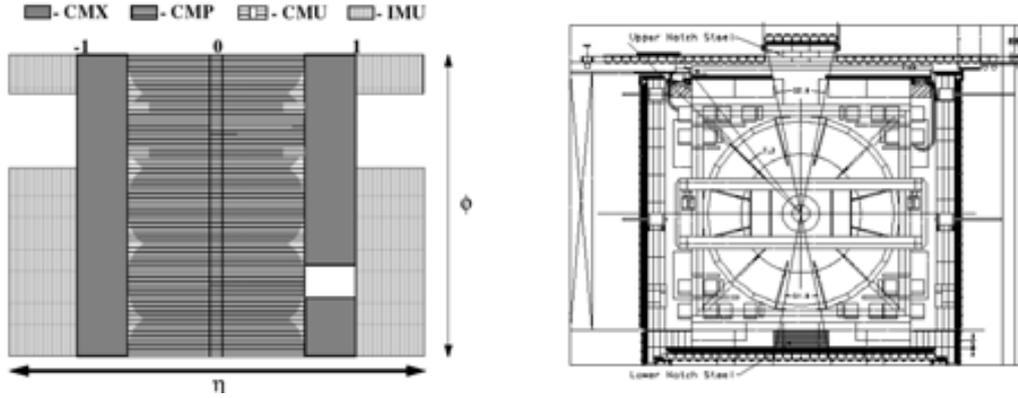


Figure 4.14: Muon sub-detectors

where the bunch-crossing frequency ($f_{b.c.}$) is precisely known from the Tevatron RF, $\sigma_{p\bar{p}-in} = 60.7 \pm 2.4 mb$ is the $p\bar{p}$ cross-section resulting from the average CDFII and E811 luminosity independent measurement at $\sqrt{s} = 1.8 TeV$, and extrapolated to $\sqrt{s} = 1.96 TeV$. ϵ is the efficiency for detecting an inelastic scattering [79] [80].

The Cerenkov luminosity Counters (CLC) are two separate modules, covering the $3.7 \leq |\eta_{det}| \leq 4.7$ range symmetrically in the forward and backward regions. Each module consists of 48 thin, 110-180 cm long, conical, ISO-butene-filled Cerenkov counters. They are arranged around the beam-pipe in three concentric layers and pointed to the nominal interaction region. The base of each cone, 6-8 cm in diameter and located at the furthest extremity from the interaction region, contains a canonical mirror that collects the light into a PMT, partially shielded from the solenoidal magnetic field. ISO-butane guarantees high refraction index and good transparency for ultraviolet photons. With a Cerenkov angle $\theta_C = 3.4^\circ$ the momentum thresholds for light emission are $9.3 MeV/c$ for electrons and $2.6 MeV/c$ for charged pions. Prompt charged particles from the $p\bar{p}$ interactions are likely to transverse the full counter length, thus generating large signals and allowing discrimination from the smaller signals of angled particles due to the beam halo or to secondary interactions. In addition, the signal amplitude distribution shows distinct peaks for different particles multiplicities entering the counters. This allows a measurement of \bar{N} with 4.2 % relative uncertainty in the luminosity range $10^{31} \leq \mathcal{L} \leq 10^{32} cm^{-2} s^{-1}$. This accuracy, combined with the 4 % relative uncertainty on the inelastic $p\bar{p}$ cross section, results in an instantaneous luminosity measurement with 5.8 % relative uncertainty.

4.8 Trigger and data acquisition system

Tevatron produced almost 1.7 millions collisions between bunches of proton and antiproton every second. Each collision recorded by CDFII is called event and dead-time is the percentage of events which are rejected because the trigger is busy during the acquisition of an event (due to the read-out of the entire detector that takes approximately 2 ms).

The most interesting processes constitute only a minimal fraction of the total events. For this

4 The Experimental Environment

reason a system is necessary to discriminate the events during their acquisition. This is the task of the trigger system, which evaluates the partial information provided by the detector and discards the uninteresting events on-line.

The CDFII trigger is a three-level system that selectively reduces the acquisition rate with virtually no dead-time, i.e. keeping each event in the trigger memory a time sufficient to allow for a trigger decision without inhibiting acquisition of the following events. Each level receives the accepted event from the previous one and, provided with detector information of increasing complexity and with more time for proceeding, applies a logical OR of several sets of programmable selection criteria to make its decision.

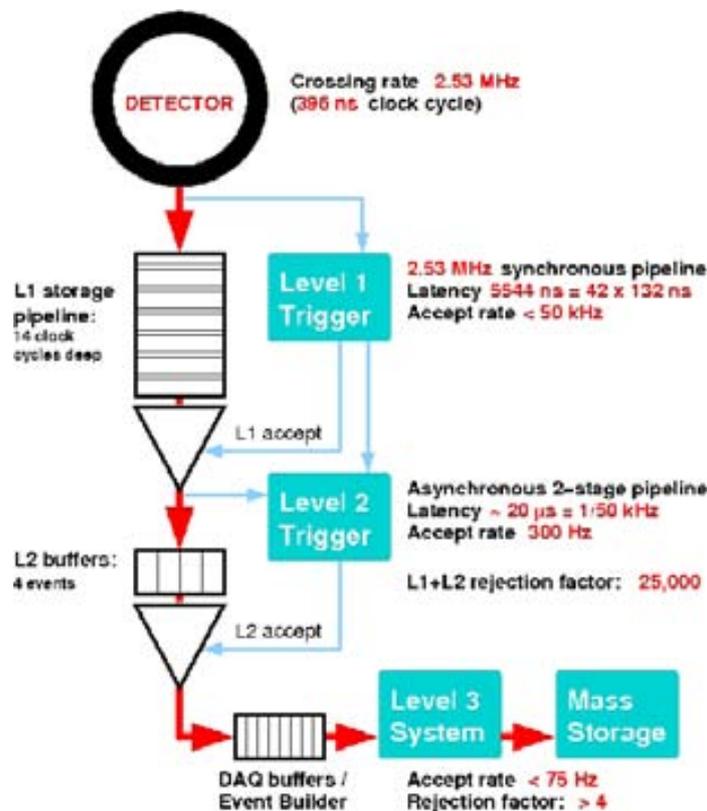


Figure 4.15: CDFII trigger system

4.8.1 Level 1

At Level-1 (L1), a buffered synchronous system of custom-designed hardware processes a simplified subset of data in three parallel streams to reconstruct information from the calorimeters (total energy and presence of single towers over threshold), the COT (two dimensional tracks in transverse plane) and the muon system (muon stubs in the CMX, CMU and CMP chambers). A decision stage combines the information from these low resolution physics objects, called primitives, into more sophisticated objects. For instance, track primitives are

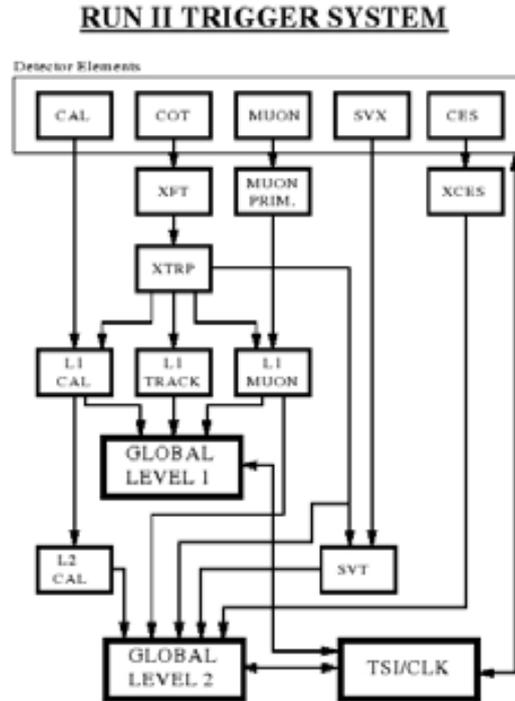


Figure 4.16: CDFII trigger block diagram.

matched with muon stubs or tower primitives to form muon, electron or jet objects, which then undergo some basic selections. This trigger can decide whatever to record the event in $5.5 \mu s$. In such a way it is able to reduce the rate of potentially interesting events to 30 kHz. The fundamental processor in this passage is the eXtremely Fast Tracker (XFT) [81] that identifies two dimensional tracks in the (r, ϕ) view of the COT (transverse plan) in time with L-1 decision.

4.8.2 Level-2

Level-2 is an asynchronous system of custom designed hardware that processes the time ordered events accepted by the Level-1. Additional information from the shower-maximum strip chambers in the central calorimeter and the axial hits in the SVXII is combined with Level-1 primitives to produce Level-2 primitives. A energy-clustering is done in the calorimeter by merging the energies in adjacent towers to the energy of a seed tower above threshold. Level-1 track primitives, matched with consistent shower maximum clusters, provide refined electron candidates whose azimuthal position is known with 2° accuracy. Information from the (r, ϕ) sides of the SVXII is combined with Level-1 tracks primitives to form two dimensional tracks with resolution similar to the offline one. Finally, an array of programmable processors makes the trigger decision, while the Level-2 objects relative to the following event accepted at Level 1 are already being reconstructed.

One important task at L-2 is performed by the Silicon Vetex Trigger.

Silicon Vertex Trigger

Reconstructing decay vertexes on-line is technically challenging and requires constrained geometrical fitting of high-resolution tracks at high events rates. The Silicon Vertex Trigger (SVT) [82] [83] computes instead the impact parameters of the charged particles, which is faster than fully reconstructing their decay vertexes, but still provides information on the lifetimes of the decaying particle. The full spatial resolution of silicon detectors is needed to discriminate $O(100 \mu m)$ impact parameters from the $O(10 \mu m)$ beam spot. Thus the SVT requires the coincidence of hits in four axial SVXII layers with a XFT track. Since the silicon signals are digitalized only after the Level 1 accept decision, the SVT is used at Level-2, whose average latency is around $20 \mu s$. Within this time, the SVT reconstructs two dimensional tracks in the bending plane of the spectrometer with off-line resolution. SVT speed is largely due to a high-parallelized structure and to the implementation of novel techniques both in pattern recognition and in track fitting.

The output of the SVT are the reconstructed parameters of the two dimensional track in the transverse plane: ϕ_0 , p_T and the impact parameter d_0 . The list of parameters for all found tracks is sent to Level-2 for trigger decision.

The SVT measures the impact parameter with a standard deviation of $\approx 35 \mu m$ and a average latency of $24 \mu s$, 9 of which spent waiting for the start of the read-out of silicon data.

4.8.3 Level 3

The digitalized output relative to the Level-2 accepted event reaches Level-3 via optical fibers and it is fragmented in all sub-detectors. It is collected by a custom hardware switch that arranges it in the proper order and transfers it to commercial computers. The ordered fragments are assembled in the event records, a block of data that univocally corresponds to a bunch crossing and it is ready for the analysis of the Level 3 software [86]. The event reconstruction benefits from full detector information and improved resolution with respect to the proceeding trigger levels, including three-dimensional track reconstruction, tight matching between tracks and calorimeter or muon information. If an event satisfies the Level-3 requirements, the corresponding event record is transferred to mass storage at a maximum rate of 250 MB per second. A fraction of the output is monitored in real time to search for detector malfunctions, to derive calibrations constants and to graphically display events. The Level-3 decision is made after the full reconstruction of an event is completed and the integrity of its data is checked.

4.8.4 Operation and data quality

During the data taking there are several procedures to check the data quality and the complete operativeness of the all sub-detectors.

Each time that at least one of the trigger path fires, an event is labeled with a progressive number. Events are grouped into runs, i.e. periods of continuous data taking in constant configurations of trigger table, set of active sub-detectors and so on. Several parameters of

the operation are stored in the database on a run-averaged format.

All data manipulation occurring some time after the data are written to permanent memories are referred to as off-line processes, as opposed to the online operations that take place in real time, during the data-taking. The most important offline operation is the processing with a centralized production analysis that generated collections of high-level physics objects suitable for analysis, such as tracks, vertexes, muons, electrons, jets etc ... from low level information such as hits in the tracking sub-detectors, muon stubs, fired calorimeter tower etc ... During the production, more precise informations about the detector conditions and more sophisticated algorithms are used than those ones available at the Level-3 of the trigger.

To ensure homogeneous data-taking conditions, each run undergoes a quality inspection. One line shift operators, offline production operators, and sub-detector experts certify in what fraction of data the running conditions for all relevant sub-detectors are compliant to physics quality standards.

5 Physics Objects Reconstruction

Events passing trigger requirements are reconstructed offline in a process during which collections of high-level physics objects are generated. This chapter presents the standard CDF algorithms to reconstruct objects, such as electrons, muons and jets. Particular attention is given to the secondary vertex finder algorithm utilized to identify jets originated from b -quarks.

5.1 Primary Vertex

5.1.1 z primary vertex

The z -coordinate of the primary vertex is reconstructed using an algorithm called *ZVertexFinder* [?]. This combines all the track information of an event and it is based on an iterative procedure. The reconstruction starts determining the zero approximation vertex coordinate z_V from the median of the z_{0i} track coordinates of a given track collection. After that a χ^2 is calculated according to the formula:

$$\chi^2 = \sum_{i=1}^{N_T} \frac{(z_V - z_{0i})^2}{\sigma_i^2} \quad (5.1)$$

where N_T is the number of tracks and σ_i is the uncertainty of z_{0i} . Tracks whose contributions of χ^2 is greater than 3 are excluded and the vertex coordinate is recalculated according to the following formula:

$$z_V = \frac{\sum_{i=1}^{N_T} \frac{z_{0i}}{\sigma_i^2}}{\sum_{i=1}^{N_T} \frac{1}{\sigma_i^2}} \quad (5.2)$$

This procedure is repeated until no track is excluded anymore. The algorithm may return more than one vertex. In that case the one with the highest $\sum p_T$ is selected, where the sum runs over the assigned tracks.

5.1.2 3-D Primary Vertex

A precise three dimensional vertex is fundamental to look for displaced secondary vertexes, as it will be discussed later.

The 3-D vertex reconstruction starts with an input seed defined by the x and the y positions

of the run average beam-line and from the z_V found with ZVertexFinder algorithm. The *PrimeVtx* selects good quality tracks that pass cuts in Table 5.1 and are compatible with the seed vertex. Using this tool a fit is performed to find a new primary vertex. At this point the algorithm starts removing tracks that contributing with the worst χ^2 (≥ 10) relative to the fit. The iterative process stops when there are not anymore tracks with $\chi^2 \geq 10$ in the vertex. The resolution of the primary vertex coordinate obtained depends on the number of the tracks, but it is of the order of $\sim 50\mu m$.

$ z_0 - z_{seed} $	≤ 1 cm
p_T	≥ 0.5 GeV/c
$ d_0 /\sigma_{d_0}$	≤ 3.0
Axial COT SL with at least 6 hits	≥ 2
Stereo COT SL with at least 6 hits	≥ 2
Silicon hits	≥ 3

Table 5.1: Track quality cuts for the primary vertex

5.2 Electrons

Electron candidates are reconstructed [?] from the electromagnetic information in the calorimeters. Towers in an event with a minimum transverse electromagnetic energy E_T^{EM} of 100 MeV are sorted in E_T and are stored in a clusterable tower bank. Then they are grouped starting from a seed among the clusterable towers with transverse energy $E_T \geq 10$ GeV, adding only the closest ones adjacent in η_{det} within the same wedge in ϕ for the central calorimeter. In the plug, the clustering can extend to a 2×2 tower configuration in $\eta - \phi$. Towers can only be part of one cluster. After this first step, electron candidates are built by matching shower-max and pre-showering information to the cluster. When possible, also, the maximum p_T track, among all the tracks pointing to the shower-max cluster, is associated.

Tracks associated to electron candidates are beam constrained using COT beam position and their p_T is corrected for a ϕ dependent corrections to take care of misalignments of the COT tracking system.

The total energy of an electron is given by the sum of the hadronic and electromagnetic energy of all the towers in the cluster. The centroid and the respective E_T , η , ϕ are evaluated according the Snowmass principles:

$$\eta = \frac{E_{EM} \times \eta_{EM} + E_{Had} \times \eta_{Had}}{E}$$

$$\phi = \frac{E_{EM} \times \phi_{EM} + E_{Had} \times \phi_{Had}}{E}$$

where the angular variables are:

$$\eta_{EM} = \frac{\sum_i E_{EM}^i \times \eta^i}{\sum_i E_{EM}^i}$$

$$\eta_{had} = \frac{\sum_i E_{had}^i \times \eta^i}{\sum_i E_{Had}^i}$$

$$\phi_{EM} = \frac{\sum_i E_{EM}^i \times \phi^i}{\sum_i E_{EM}^i}$$

$$\phi_{had} = \frac{\sum_i E_{had}^i \times \phi^i}{\sum_i E_{Had}^i}$$

In order to identify electrons among the selected candidates, characteristic variables able to discriminate electrons from fake signals will be used later in the analysis. These are:

$E_T = E \sin \theta$: The transverse electromagnetic energy, E_T , is obtained from the energy of the EM cluster and the polar angle θ of the associated track.

$p_T = P \sin \theta$: The transverse momentum of the track associated with the electron. P is the momentum of the track.

E_{Had}/E_{Em} : The ratio between the hadronic and electromagnetic energy deposition in the calorimeter. This is particularly important to discriminate electrons from jets.

E/P : The ratio between the transverse energy of the EM cluster and the transverse momentum of the track. This ratio could be large if electrons radiate photons. For high energy electrons the value of the fraction is close to 1.

$\Delta x_{CES}Q, \Delta z_{CES}$: Distances between the extrapolated track and the best matching CES cluster respectively in the plane (r, ϕ) and (r, z) . Δx_{CES} is usually multiplied for the track charge Q .

χ_{CES}^2 : The χ^2 comparison of the CES shower profile with the shower profile obtained from the test beam measurement.

L_{shr} : This is a variable useful to discriminate electrons and photons from hadronic showers faking these particles in the central electromagnetic calorimeter. It compares the distribution of adjacent CEM tower energies in the cluster to shapes derived from test-beam, as function of the seed energy.

$Isol04$: The isolation is defined as:

$$Isol04 = \frac{E_T^{cone} - E_T^{electron}}{E_T^{electron}}$$

where $E_T^{electron}$ is the energy of the electron cluster and E_T^{cone} is the transverse energy in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$ around the electron cluster.

$\chi_{3 \times 3}^2$: The χ^2 comparison of the PEM shower profile with the shower profile obtained from the test beam measurement.

$PES U$ energy: The energy for the U strip of the best matched cluster in PES

$PES V$ energy: The energy for the V strip of the best matched cluster in PES

5 Physics Objects Reconstruction

$\Delta R(PEM, PES)$: The difference between the PEM and PES coordinates of the electron cluster

N_{hits}^{Si} : The number of the silicon detector hits

N_{trk} : Number of tracks associated to the cluster

$E_{ne_{3 \times 3}}$: Energy of 3×3 PEM cluster

5.3 Muons

Muons, as minimum ionizing particles, leave only residual energy in the calorimeters. They are identified by signals left in the tracking system and, when fiducial in the muon chambers, also by the stubs present in those detectors. Depending on the muon chamber where the stub is found, muons are classified as CMU, CMP, CMX and BMU muons.

The muon reconstruction [?] begins from the information collected in drift chambers and scintillators of the muon sub-detectors, looking for a stub that links a few hits. This is obtained with an iterative fit of the hits and the criteria used depend on the particular muon detector. Tracks are then associated to the stubs. Only one track should be assigned for each stub. When more than one track is matched, tracks are sorted based on the quality of the fit extrapolation and the best track is assigned. For the candidates that passed this step, calorimeter information is added by extrapolating the track trajectory into the calorimeter. Tracks that remain without a stub are integrated with calorimeter information and are stored as a stub-less muon candidates.

The main variables used in the analysis to identify muons are:

p_T, E_{Em}, E_{Had} : The transverse momentum p_T of the best matched track associated to the muon, the EM (HAD) calorimeter energy corresponding to the muon

N_{SL} : Number of COT super-layers passed through by the track associated to the muon

z_0 : The z-coordinate of the muon associated track at the distance of the closest approach to the beam-line

d_0 : The impact parameter, that is the distance of the muon associated track to the primary vertex in the (r, ϕ) plane.

$Isol_{04}$: The ratio between the energy deposited in the calorimeter towers within a cone ($\Delta R = 0.4$) and the muon track p_T

$\frac{\chi^2}{n.d.f}$: The reduced χ^2 for the stub-track association

N_{hits}^{Si} : The number of the silicon detector hits

N_{hits}^{COT} : The number of the COT detector hits

N_{SL}^{AX} : The number of axial COT SL with at least 5 hits

N_{SL}^{ST} : The number of stereo COT SL with at least 5 hits

5.4 Jets

As defined in Chapter 2, jets [89] consist of collimated spray of high-energy hadrons. The signature of jets are deposits of energy in electromagnetic and hadronic calorimeters. Their reconstruction starts from the physical towers in the calorimeter and depends on the jet algorithm used.

A physical tower is created from the detector tower using the following formulas:

$$\begin{aligned} p_x &= E_{EM} \sin \theta_{EM} \cos \phi_{EM} + E_{HAD} \sin \theta_{HAD} \cos \phi_{HAD} \\ p_y &= E_{EM} \sin \theta_{EM} \sin \phi_{EM} + E_{HAD} \sin \theta_{HAD} \sin \phi_{HAD} \\ p_z &= E_{EM} \cos \theta_{EM} + E_{HAD} \cos \theta_{HAD} \\ E &= E_{EM} + E_{HAD} \end{aligned}$$

where E_{EM} (E_{HAD}) is the energy deposited in the electromagnetic (hadronic) compartment of the detector tower. θ_{EM} and ϕ_{EM} (θ_{HAD} , ϕ_{HAD}) represent the direction from the interaction point to the shower maximum position of an electromagnetic (hadronic) shower (Figure 5.1).

In this analysis we are using jets reconstructed with the Midpoint algorithm in a cone of radius $R = 0.7$.

5.4.1 The CDF MidPoint Jet clustering Algorithm

As explained in Chapter 2, the Midpoint is a seed based algorithm. The towers with an energy greater than 100 MeV are selected and sorted in p_T . Towers belonging to identified electrons or muons are excluded from the towers available to reconstruct the jet. Starting from the highest p_T seed towers (with a momentum greater than 1 GeV/c), a protojet is created by adding adjacent towers within a cone of radius 0.7 in $\eta - \phi$ space. This is an iterative process, every time that a tower is added a cluster centroid is recalculated, a new cone is drawn and new towers are added, until the protojet becomes stable¹. The determination of the centroid is done in E-massive scheme that consists in adding the four-momenta. This makes the jet massive, contrary to the Snowmass scheme where the jets are considered massless.

After that, a list of midpoints is generated between protojets separated by less than twice the cone radius and new stable protojets are found around midpoints. Adding midpoint seeds between all stable cones reduces the sensitivity of the algorithm to soft radiation [41].

The protojets are then passed to splitting/merging steps where the overlapping protojets are separated based on the percentage of p_T shared by the lower p_T protojet. Protojets that are sharing a fraction greater than 75 % will be merged, otherwise they will split and towers assigned to the closer cone.

What we have described is the detector level jet reconstruction. In this analysis we also use jets defined at hadron level in MC where four-vectors of the stable particles are the basic elements to be clustered. The reconstruction procedure follows the steps listed above.

¹This means that the process is iterated until the cone axis and the centroid coincide, indicating that the configuration results stable.

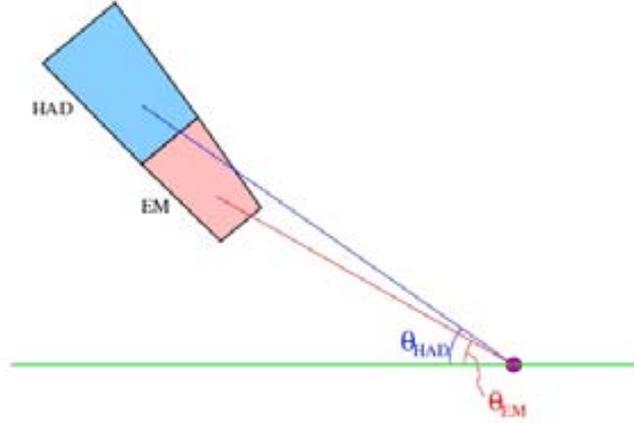


Figure 5.1: Scheme of a single calorimeter tower with electromagnetic (pink) and hadronic compartments. θ_{EM} and θ_{HAD} are the respective directions with respect the interaction point.

5.4.2 Jet Corrections

Jets reconstructed at calorimeter level are affected by losses in the gaps, multiple interactions, instrumental effects and detector non linearity in addition to the extra energy added by multiple interactions. In order to match the corresponding particle energy, jets are corrected[90] in the following way:

$$p_T^{particle} = (p_T^{jet} \times C_\eta - C_{MPI}) \times C_{abs}$$

where p_T^{jet} is the transverse momentum of the jet and C_η , C_{MPI} , C_{ABS} are the corrections due to the rapidity dependence of the calorimeter response, the multiple interactions and the absolute energy, which are described below.

Eta corrections

C_η , called relative corrections or L1, are performed to flatten the η dependence of the calorimeter response. These corrections are determined using the dijet balancing method which assumes the two jets to be balanced in p_T in absence of hard QCD radiations. A “trigger jet” is required to be in the central calorimeter and the other jet, called “probe jet”, could be anywhere in η . The correction factors, defined as $\beta = \frac{p_T^{probe}}{p_T^{trigger}}$, are determined separately for data and MC and for different p_T regions.

The systematic uncertainty is obtained by varying the event selection requirements and the fitting procedure. Plots of Figure 5.2 (5.3) show β as a function of rapidity and for different p_T regions before (after) the L1-corrections.

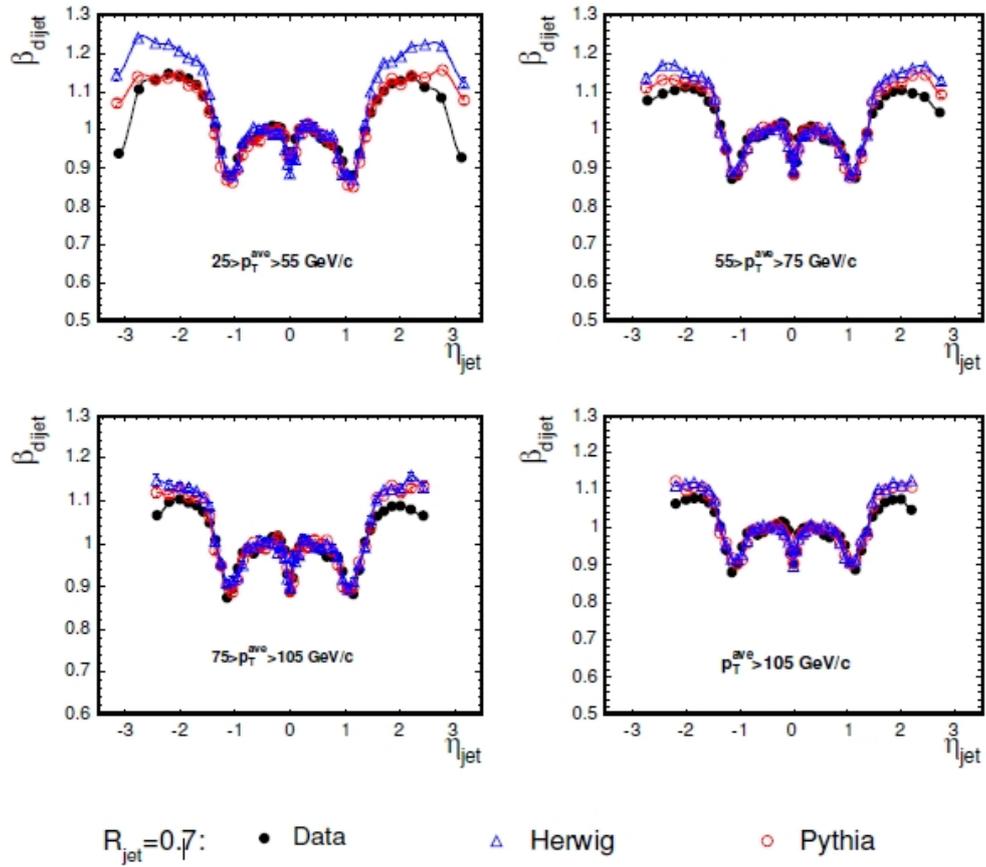


Figure 5.2: Dijet p_T balance as a function of η_{det} in data, HERWIG and PYTHIA MC samples for jets of cone size $R = 0.7$. Shown are the correction factor for several p_T regions. Figure taken from [90].

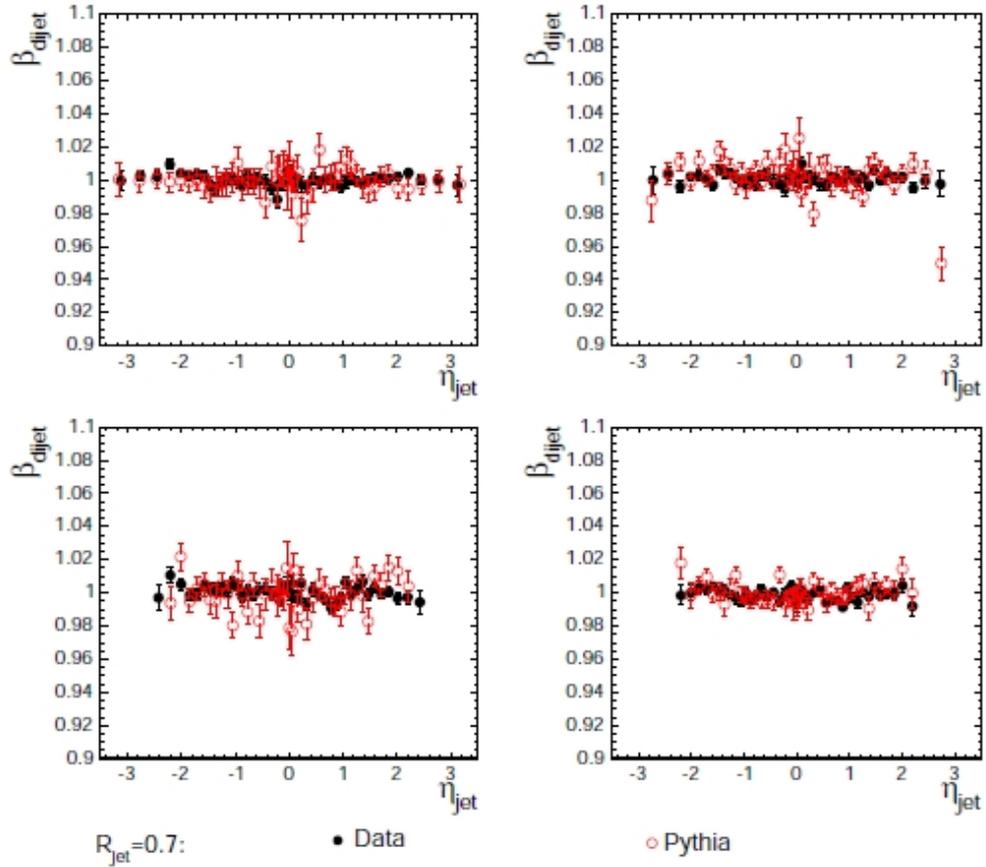


Figure 5.3: Dijet p_T balance as a function of η_{det} in data and PYTHIA MC samples for jets of cone size $R = 0.7$. Shown are the correction factor for several p_T regions after applying the L1-corrections. Figure taken from [90].

Multiple Interaction Corrections (C_{MPI})

The multiple interaction corrections, or L4, subtract from the jet energy the amount of energy coming from extra $p\bar{p}$ interactions taking place in the same bunch crossing and estimated by the number of reconstructed z -vertexes (N_{vtx}). These corrections are determined using a minimum bias data sample (collected requiring at least one $p\bar{p}$ interaction). The average transverse energy in a cone is measured in the central calorimeter as a function of the number of vertices and for three cone sizes. Data are parameterized using a fitted straight line (Figure 5.4) where the slope parameters give the extra transverse energy per interaction as a function of N_{vtx} . The p_T is corrected according to this formula:

$$p'_T = p_T - \epsilon(N_{vtx} - 1)$$

The systematic uncertainty from this correction is approximately 15 % and it is due to dependence on instantaneous luminosity and event topologies.

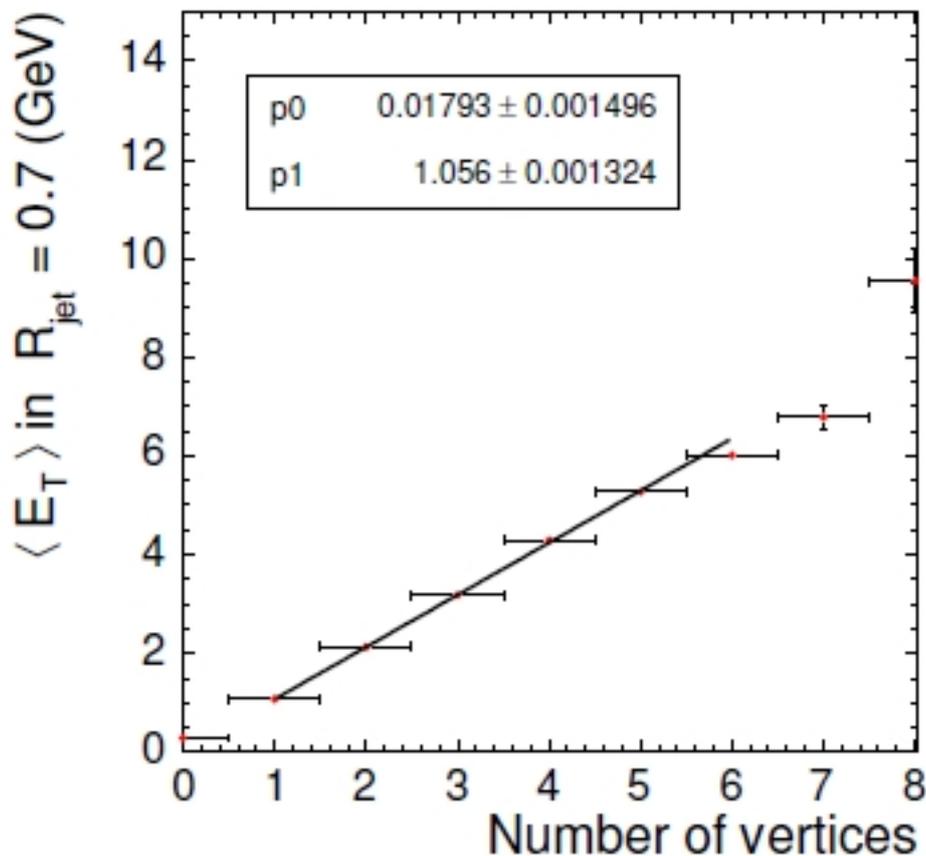


Figure 5.4: $\langle E_T \rangle$ versus the number of vertexes for jets of cone size $R = 0.7$. A linear fit is also shown. Figure taken from [90]

Absolute Correction (C_{abs})

This last correction that is considered is the absolute jet energy scale correction or L5. It transforms the jet energy measured in the calorimeter into the corresponding to the underlying particle jet so that the jet energy scale can be independent from the detector. This is done to compensate for non linearity and energy loss in the un-instrumental regions of the calorimeter. The correction is obtained in MC by mapping the total p_T of the particle level jet to the p_T of the calorimeter level jet. Jets reconstructed at calorimeter and particle level use CDF standard clustering, and they are required to be in a central region and to be one of the two leading jets. Particle jets are matched to calorimeter jets within $\Delta R \leq 0.1$.

The difference between the particle and calorimeter jet p_T is shown in Figure 5.5 for four p_T ranges. In Figure 5.6 the absolute corrections are shown for different cone size jets.

The main source of systematic error is uncertainty on the simulation of the calorimeter response to charge hadrons. The overall uncertainty is approximately of 2 % for low p_T^{jet} and rises to 3 % at high p_T^{jet} .

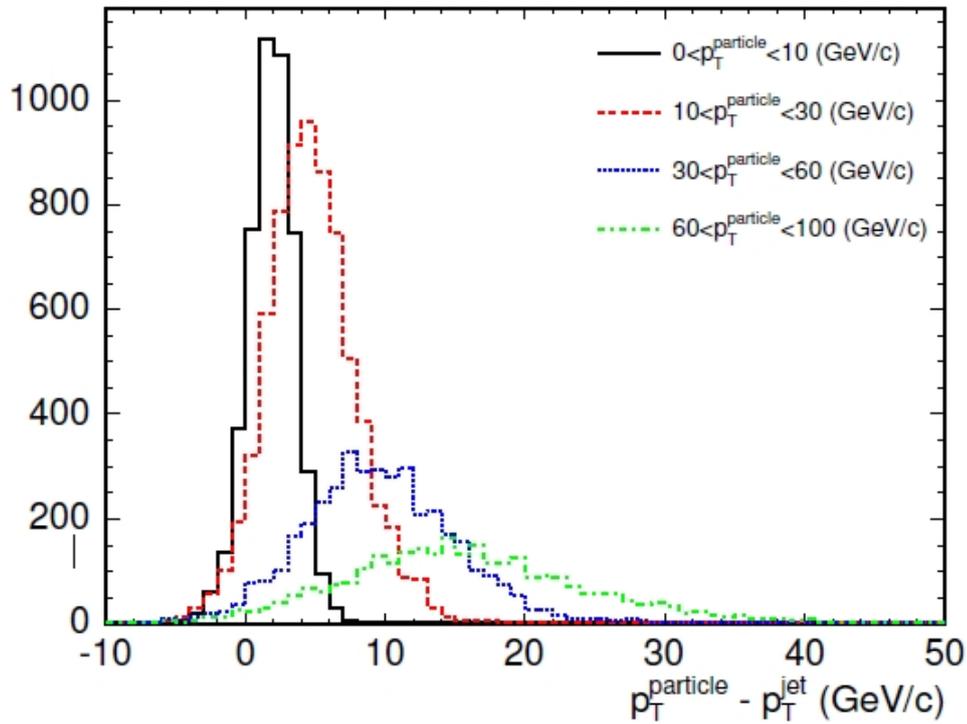


Figure 5.5: p_T difference between particle and calorimeter jet for different p_T range. Image taken from [90].

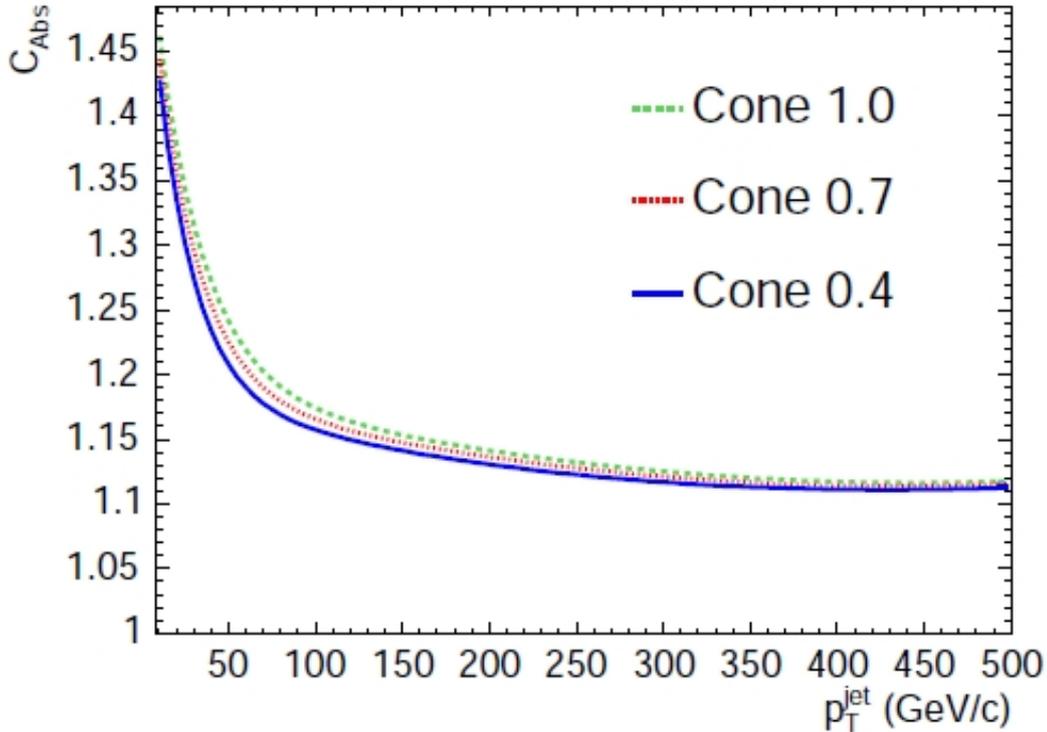


Figure 5.6: Absolute corrections for different cone sizes as a function of calorimeter jet p_T . Taken from [90].

5.4.3 b-tagging : bottom jet identification

In this analysis, it is key to identify jets coming from b -quarks. At CDF several algorithms that exploits different qualities of B hadrons have been developed. We use SecVtx algorithm which takes advantage from the relatively large B lifetime of ≈ 1.5 ps and its large mass ≈ 5 GeV/ c^2 . B -hadrons fly on average 0.5 mm before decaying, which is a larger distance than the intrinsic beam size. Thus their decay products are characterized by a non-zero impact parameter. The algorithm [?] looks for a reconstructed secondary vertex displaced from the primary vertex inside a jet.

Secondary vertex tagging operates on a per-jet basis where only tracks within the jet cone ($|\Delta R| \leq 0.4$) are considered for each jet in the event. A set of cuts involving the transverse momentum, the number of silicon hits attached to the tracks, the quality of those hits, and the χ^2/ndf of the track fit are applied to reject poorly reconstructed tracks. Only jets with at least two of these good tracks can produce a displaced vertex and are defined as “taggable”. Displaced tracks in the jet are selected based on the significance of their impact parameter with respect to the primary vertex and they are used as input to the SecVtx algorithm that used a two-pass approach to find secondary vertices. In the first pass, using tracks with $p_T \geq 0.5$ GeV/ c and the significance of d_0 (S_{d_0}) greater than 2.0, it attempts to reconstruct a secondary vertex which includes at least three tracks (at least one of the tracks must have p_T greater than 1 GeV/ c). If the first pass is unsuccessfully, it performs a second pass which

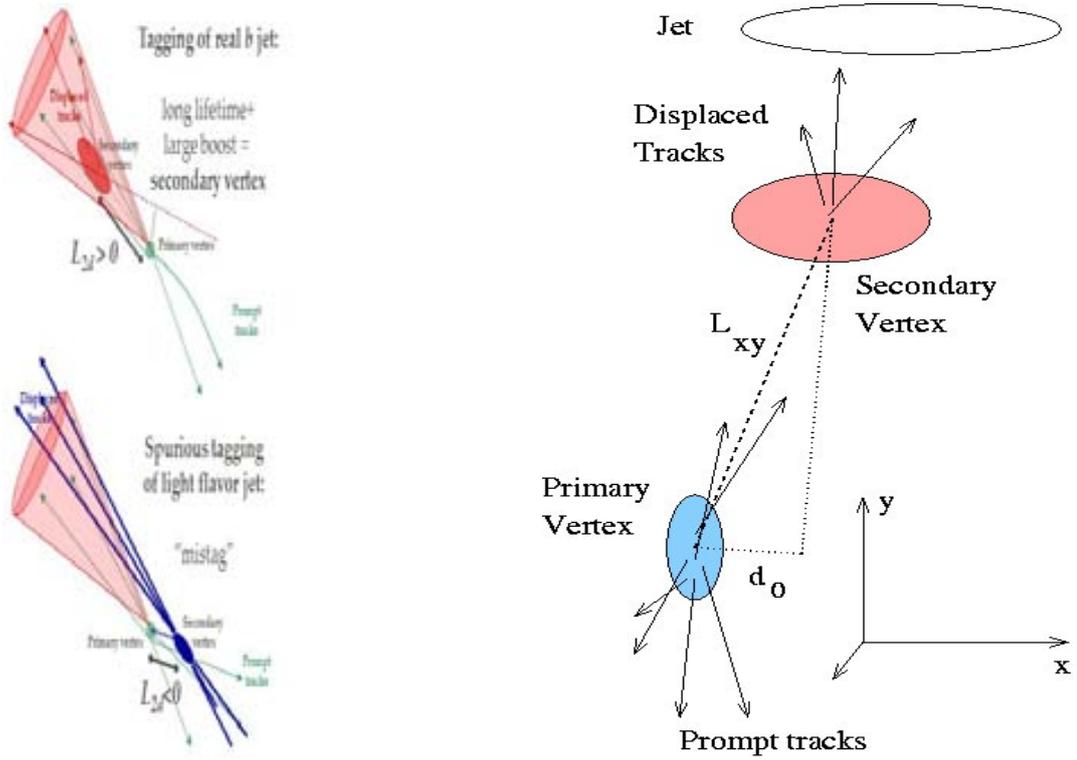


Figure 5.7: Scheme of a tagging variable, in particular the geometrical interpretation of L_{xy} and d_0 is shown.

makes tighter track requirements ($p_T \geq 1 \text{ GeV}/c$ and $S_{d_0} \geq 3.5$) and attempts to reconstruct a two-track vertex (one track with p_T greater $1.5 \text{ GeV}/c$). Tracks consistent with K_0 or Λ are not considered by the algorithm. In Table 5.2 are summarized the most relevant parameters. Once a secondary vertex is found inside a jet, two dimensional decay length of the secondary vertex L_{xy} (Figure 5.7) is calculated as the projection onto the jet axis, in r - ϕ view only, of the vector pointing from the primary vertex to a secondary vertex. The sign of L_{xy} is defined relative to the jet direction. To reduce backgrounds, jets not associated to a b -quark, a cut on L_{xy} significance is required and according to the value of this cut a loose and a tight Tagger is defined.

Tagging efficiency

The b -tag algorithm is calibrated in data. However, since MC simulations are widely used the b -tag efficiency is also evaluated in a controlled MC sample and then a SF is calculated to take into account for possible discrepancies between data and MC. Efficiency results as a function of jet p_T and jet rapidity are shown in Figure 5.8. The average scale factor for the Tight SecVtx Tagger is $SF_{tag} = 0.96 \pm 0.05$.

Variable	SecVtx Tight	
	Pass 1	Pass 2
Tracks criteria		
p_T (GeV/c)	> 0.5	1.0
SVXII layer with hits	> 3	
$\frac{d_0}{\sigma_{d_0}}$	> 2.0	3.5
d_0 (cm)	< 0.15	
$\frac{\chi^2}{n.d.f.}$	< 8.0	
χ^2	< 45	30
$\delta(z_0)$ (cm)	< 2.0	
Vertex criteria		
$\frac{d_0}{\sigma_{d_0}}$ of third track	> 4.0	
at least 1 track with p_T (GeV/c)	> 1.0	1.5
χ^2 of primary vertex	< 50	
χ^2 of the fit vertex	< 50	
L_{xy} (cm)	< 2.5	1.2
$\frac{L_{xy}}{\sigma_{L_{xy}}}$	> 7.5	

Table 5.2: SecVtx Tight parameters for Pass 1 and for Pass 2.

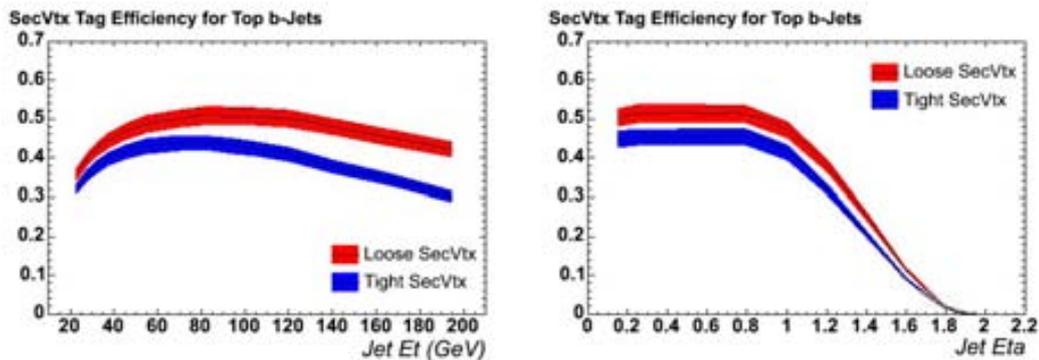


Figure 5.8: Efficiency for SecVtx tagger algorithm as a function of jet p_T and jet rapidity. The measurements were done using $t\bar{t}$ samples.

5 Physics Objects Reconstruction

The algorithm tagged also jets not coming from heavy flavor quarks. These jets that fake the SecVtx, are called mistag and their rate is estimated using a parameterization done in a dijet data sample. The results as a function of p_T^{jet} and of Y^{jet} are shown in Figure 5.9.

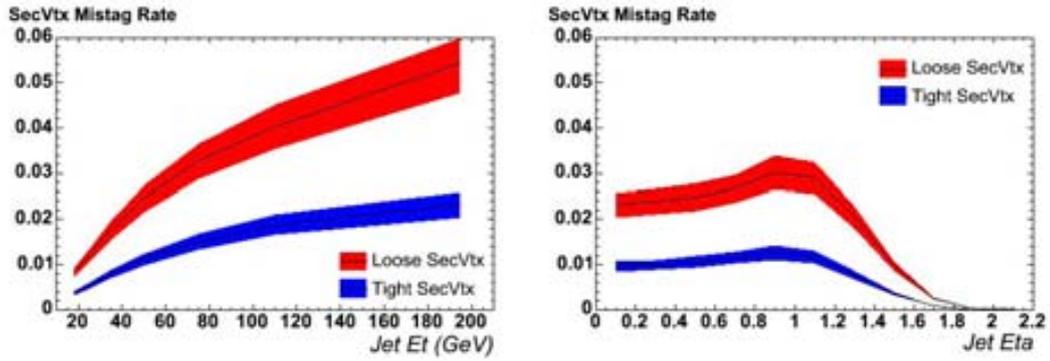


Figure 5.9: SecVtx tagger mistag as a function of jet p_T and jet rapidity. The measurement were perform in data dijet sample.

6 The $Z/\gamma^* + b$ -jet cross section measurement

This chapter describes the b -jet cross section measurement produced in association with a Z/γ^* boson. The description includes the methodology, the identification of the physics objects utilized in the analysis and the full treatment of the systematics uncertainties. Particular attention is given to the method used to estimate the b -jet fraction on data and its uncertainties.

6.1 Measurement definition

In general, a production cross section σ is defined as

$$\sigma = \frac{N_{data} - N_{bkg}}{A \cdot \mathcal{L}} \quad (6.1)$$

where N_{data} and N_{bkg} are respectively the number of events observed in the data and of expected background. A is the acceptance times the event selection efficiency for the process being observed and \mathcal{L} is the total integrated luminosity of the data.

Using this formula, a per jet cross section can be defined replacing the number of events with the number of jets, in this case the number of b -jets. By defining a per jet cross section in a well defined phase space the measurement of $Z/\gamma^* + b$ -jet cross section would not depend on possible flaws of the $Z/\gamma^* + b$ modeling, such as the number of jets per event or the extrapolation outside the detector acceptance. Moreover by performing a cross section ratio measurement with respect to Z/γ^* inclusive and $Z/\gamma^* + jets$ cross sections some systematics (such as luminosity, lepton ID efficiencies) largely cancel in the ratio. The per jet cross section ratio with respect Z/γ^* inclusive cross section is written as:

$$R = \frac{\sigma(Z/\gamma^*_{(\rightarrow l+l^-)} + b \text{ jet})}{\sigma(Z/\gamma^*_{\rightarrow l-l^+})} = \frac{A_{Z/\gamma^*}^{MC}}{A_{Z/\gamma^*+b-jet}^{MC}} \cdot \frac{N_{Z/\gamma^*+b-jet}^{data} - N_{bkg}}{N_{Z/\gamma^*}^{data} - N_{bkg}} \quad (6.2)$$

where l denotes electron or muon, A_{Z/γ^*}^{MC} and $A_{Z/\gamma^*+b-jet}^{MC}$ are respectively the Z/γ^* and $Z/\gamma^* + b$ -jet acceptance times the efficiency which are evaluated from MC.

The definition with respect $Z/\gamma^* + jets$ is similar and is obtained replacing A_{Z/γ^*} and N_{Z/γ^*}^{data} with the corresponding for $Z/\gamma^* + jet$.

The b jets in data are identified using the Tight SecVtx Tagger. Since the tagged jet sample is not only composed by b -jets, $N_{Z/\gamma^*+b-jet}^{data}$ is determined from a fit to the secondary vertex mass distribution M_{SecVtx} :

6 The $Z/\gamma^* + b$ -jet cross section measurement

$$N_{Z/\gamma^*+b\text{-jet}}^{data} = \frac{n_{data}^{tagged} \cdot f_b}{\epsilon_{tag}^b} \quad (6.3)$$

where n_{data}^{tagged} is the number of tagged jets, f_b the fraction of b -jets in the tagged sample obtained from the fit and ϵ_{tag}^b the b -tagged efficiency.

The measurement is defined for events with a Z/γ^* boson in the mass range $66 \leq M_{Z/\gamma^*} \leq 116$ GeV/ c^2 and a b hadron level jet with a $p_T \geq 20$ GeV/ c and a $|Y^{jet}| \leq 1.5$.

Event selection, secondary vertex mass fit, background subtraction and acceptance calculation are discussed in the following sections.

6.2 Definition of the dataset

We have analyzed the full dataset collected at CDF during Run II (March 2002-September 2011), that consists of $\sim 10 \text{ fb}^{-1}$ of integrated luminosity. Events are required to fire one of the following trigger paths for high- p_T central leptons. Each path consists of different criteria for each trigger level.

- ELECTRON_CENTRAL_18;
 - L1 - A central electron cluster with $E_T > 8$ GeV, $E_{HAD}/E_{EM} < 0.125$, and an associated $p_T > 8$ GeV XFT track.
 - L2 - A central electron cluster with $E_T > 16$ GeV, $E_{HAD}/E_{EM} < 0.125$, and an associated $p_T > 8$ GeV XFT track
 - L3 - A central electron cluster with $E_T > 18$ GeV, $E_{HAD}/E_{EM} < 0.125$, $L_{SHR} < 0.4$ and an associated $p_T > 9$ GeV L3 track that extrapolates to the CES within 8 cm in z cluster position.
- MUON_CMUP18
 - L1 - An XFT track with $p_T > 4$ GeV associated with both a CMU and a CMP stub
 - L2 - An XFT track with $p_T > 15$ GeV associated with both a CMU and a CMP stub
 - L3 - A minimum ionizing track with $p_T > 18$ GeV associated CMU and CMP stubs, $|\Delta X_{CMU}| < 20$ cm, and $|\Delta X_{CMP}| < 10$ cm.
- MUON_CMX18
 - L1 - An XFT track with $p_T > 8$ GeV associated with a CMX stub and CSX scintillator information
 - L2 - An XFT track with $p_T > 15$ GeV associated with a CMX stub
 - L3 - A minimum ionizing track with $p_T > 18$ GeV associated with a CMX stub with $|\Delta X_{CMX}| < 10$ cm

From all the data collected during Run II only those events acquired with a functional silicon tracker, electron systems and central muon systems (CMU, CMP and CMX) are considered. The total integrated luminosity per channel is $\sim 9.1 \text{ fb}^{-1}$.

6.3 Monte Carlo Samples

Monte Carlo simulated samples are used to model $Z/\gamma^* + \text{jets}$ events, to estimate background contributions, to evaluate the acceptance and to build the secondary vertex mass templates. These templates are used to discriminate the different jet flavors in the tagged sample.

MC events are produced according to its respective generator algorithm and then passed through the GEANT 3 based [94] CDF detector simulator. Several checks are performed to guarantee the agreement between data and MC distributions (section 6.6).

6.3.1 Alpgen+Pythia $Z + \text{jets}$ MC

ALPGEN v2.10' interfaced to PYTHIA v6.325 tune BW with CTEQ5L PDFs [95] is used to simulate $Z/\gamma^* + \text{jets}$ events. Samples were generated for $Z/\gamma^* + 0, 1, 2, 3, 4$ partons with the built-in mechanism, called "MLM matching" (see section 2.6.5), to remove overlap between jets from parton showers (PS) and from matrix element (ME) at the generator level. Since this procedure is not applied to heavy flavor (HF) quarks, when combining the samples, we remove events in each sample in order to avoid double-counting.

Alpgen generates ME heavy flavor separately ($Z/\gamma^* + Np$, $Z/\gamma^* + bb + Np$, $Z/\gamma^* + cc + Np$). In the ME light flavor sample ($Z/\gamma^* + Np$) Alpgen generates only up, down, strange and massless charm quarks, while the parton shower, done by Pythia, can generate all five flavors with mass. A heavy flavor double-counting can occur: for example $Z/\gamma^* + b\bar{b} + 1p$ and $Z/\gamma^* + 1p + (bb)^{PS}$ can occupy the same phase space. The procedure applied to avoid this overlap is a jet-based removal that allows HF from ME only if they are in a different reconstructed jet and from PS only if they end up in the same jet.

Since the MC samples are generated using a luminosity profile representative of only a fraction of the data sample, a reweight procedure was applied to match that of the full dataset. This was done converting the luminosity profile in a multiple interactions profile and using the latter to reweight the MC.

6.3.2 Pythia MC

Pythia Tune A MC samples are used to simulate inclusive $Z/\gamma^* \rightarrow l^+l^-$ production, top pair production and dibosons (WW , WZ , ZZ). Λ_{QCD} is set to 0.146 and the PDFs used are CTEQ5L [95].

6.4 Event Selection

The event selection starts at online level when the trigger requirements described in Section 6.2 are applied. Data is then analyzed offline to reconstruct the physics objects of interest. In this case the event signature contains two high- p_T leptons (electrons or muons) and at least a b -jet.

The full event selection is explained in this section. The same selection is applied on data and on simulated events from Monte Carlo programs.

Events passing trigger criteria are required to have a high-quality z -vertex with certain qualities and with a z position within 60 cm from the center of the detector. In the same way all the MC events are filtered requiring the generated primary vertex to be within 60 cm. This suppress non-collision backgrounds and removes events with abnormal calorimeter topologies which might violate assumptions implicit in the standard jet energy corrections [90].

Z/γ^* boson candidates are reconstructed via the identification of the high- p_T leptons. In the case of the decay into a pair of electrons, those are required to have $E_T \geq 20$ GeV, being central ($|Y| \leq 1.0$) or one of them central and the other in the forward region ($1.2 \leq |Y| \leq 2.8$). For muons, they should have $p_T \geq 20$ GeV/c, opposite charges and invariant mass $66 \leq M_{\mu\mu} \leq 116$ GeV/c² (the same for M_{ee}).

The procedures explained in Chapter 5 are utilized to reconstruct leptons and jets. Further cuts are made and hereby described to identify these objects reducing contributions from false signals and background sources.

The standard way to identify leptons is to apply a rectangular cuts on a set of variables suitable to differentiate fake leptons from real ones while in this analysis they are identified by means of an Artificial Neural Network (ANN). This allows to use the full distribution of variables, increasing the lepton identification efficiency while keeping a similar fake rejection rate.

6.4.1 Muon identification and $Z/\gamma^* \rightarrow \mu^+\mu^-$ reconstruction

Two ANNs are trained to discriminate between high- p_T muons coming from a $Z/\gamma^* \rightarrow \mu^+\mu^-$ decay and two different sources of fake (misidentified) muons. The first category of fakes comes from tracks of charged particles originated within jet fragmentation. This kind of fakes can be distinguished from real muons because they have similar probability to have same charge or opposite charge with respect to a muon identified in the same event, and for this reason this category is defined as *Same Charge* (SC) fakes. The other category of fake muons comes from low p_T charged particles which undergo a decay in flight, and because of a kink in their trajectory are incorrectly reconstructed as high p_T tracks. This category is defined as *Decay in Flight* fakes (DIF) and can be distinguished from real muons for their high impact parameter and poor quality of the tracking fit.

Three different samples of muons corresponding to real muons, SC fakes and DIF fakes are selected from data in the high- p_T sample using the following criteria.

- Real muons from $Z/\gamma^* \rightarrow \mu^+\mu^-$ decay
 - One tight CMUP or CMX
 - * At least 2 COT SL
 - * $|d_0| \leq 0.2$ cm if no silicon track or $|d_0| \leq 0.02$ cm if silicon track

- * $Isol04 \leq 0.1$
- * $E_{HAD} \geq 6 \text{ GeV}$
- * $E_{Em} \geq 2 \text{ GeV}$
- W + jet veto: discard events with $\cancel{E}_T > 20 \text{ GeV}$ ¹ and $p_T^{jet} > 15 \text{ GeV}/c$
- Select probe muons passing kinematic requirements associated to an identified tag muon so that the two muons have
 - * Opposite charge
 - * $\Delta z_0 < 4 \text{ cm}$
 - * Invariant mass within $86 - 96 \text{ GeV}/c^2$
- Fake SC muons
 - One tight CMUP or CMX muon
 - Select probe muons passing kinematic requirements and $\frac{\chi^2}{n.d.f.} \leq 4$ associated to an identified tag muon so that the two muons have
 - * Same charge
 - * $\Delta z_0 < 4 \text{ cm}$
- Fake DIF muons
 - One tight CMUP or CMX muon
 - W + jet events: $\cancel{E}_T > 20 \text{ GeV}$ and $p_T^{jet} > 15 \text{ GeV}/c$
 - Select muons passing kinematic requirements and $|d_0| > 0.2 \text{ cm}$

Figure 6.1 shows some relevant distributions for the three data driven muon and fakes samples.

The ANN used to differentiate SC fakes from signal muons employs the following variables:

- Energy in the EM calorimeter
- Energy in the HAD calorimeter
- Number of Axial COT SL with at least 5 hits
- Number of Stereo COT SL with at least 5 hits
- Number of COT hits
- Number of Silicon hits
- Distance from z -vertex

¹Missing E_T : is the magnitude of the missing transverse momentum. This is a signature of neutrinos and is evaluated from the imbalance in the measurements of the energy and momenta of all the other particles in the transverse plane. As a consequence of the conservation of the moment and of the fact that incoming proton and antiproton have negligible transverse momenta, also the final transverse momenta should be zero.

6 The $Z/\gamma^* + b$ -jet cross section measurement

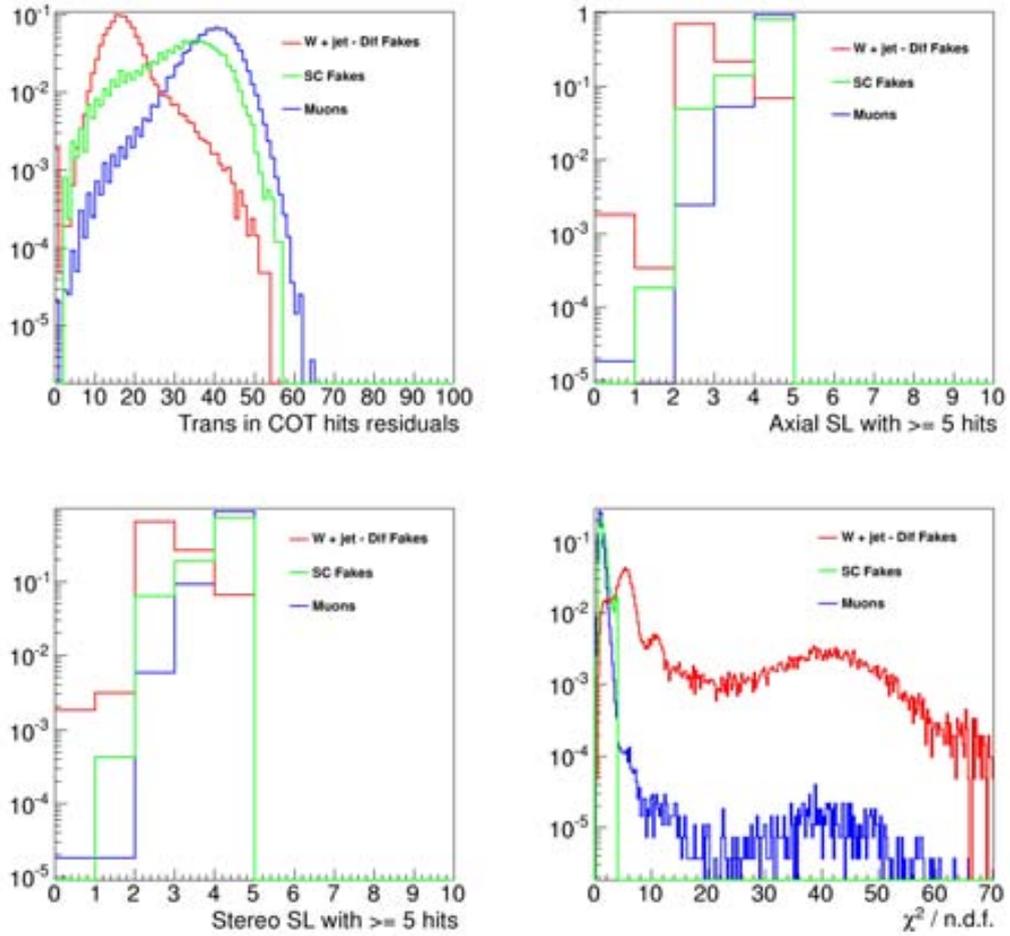


Figure 6.1: Comparison between probe and fake muons for some kinematical distributions used by muon ANNs to discriminate between real muons and fakes. The distributions are normalized to 1.

- Impact parameter d_0
- Isolation

The DIF fake ANN makes use of:

- Energy in the EM calorimeter
- Energy in the HAD calorimeter
- Number of Axial COT SL with at least 5 hits
- Number of Stereo COT SL with at least 5 hits
- Number of COT hits
- Number of Silicon hits
- Number of transitions in the residuals of the COT fit
- $\frac{\chi^2}{n.d.f}$
- Maximum number of consecutive residuals on the same side of the track in the COT fit
- Distance from z -vertex

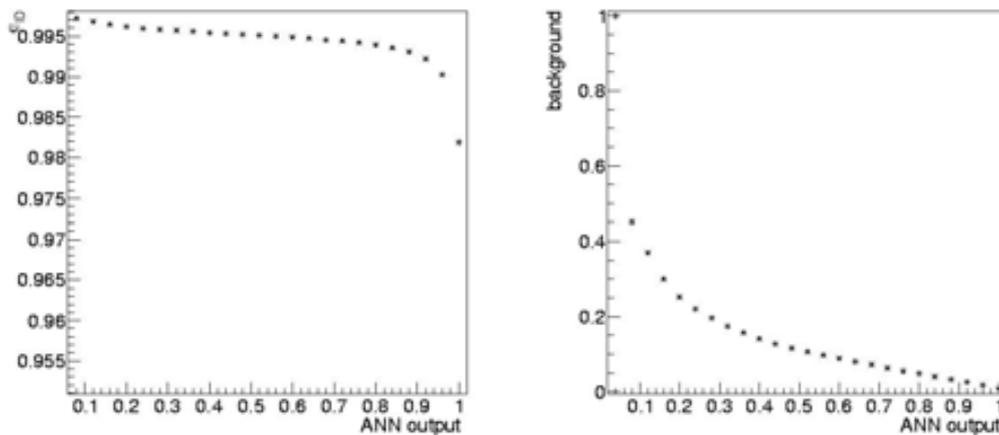


Figure 6.2: Muon ϵ_{ID} and background reduction rate as function of Same Charge ANN output. The cut on this ANN output in the analysis is 0.875.

Figure 6.2 (6.3) shows the muon efficiency, ϵ_{ID} , and fake muon survival rate as function of the output of the trained SC (DIF) ANN and Figure 6.4 presents the output of the trained ANNs. The ϵ_{ID} are high, close to 99 % over almost all the ANN output range, allowing to obtain high efficiency with very low background. Cuts for each ANN were found optimizing the significance $\sigma = \frac{S}{\sqrt{S+B}}$ with S the number of reconstructed $Z/\gamma^* \rightarrow \mu^+\mu^-$ and B the

6 The $Z/\gamma^* + b$ -jet cross section measurement

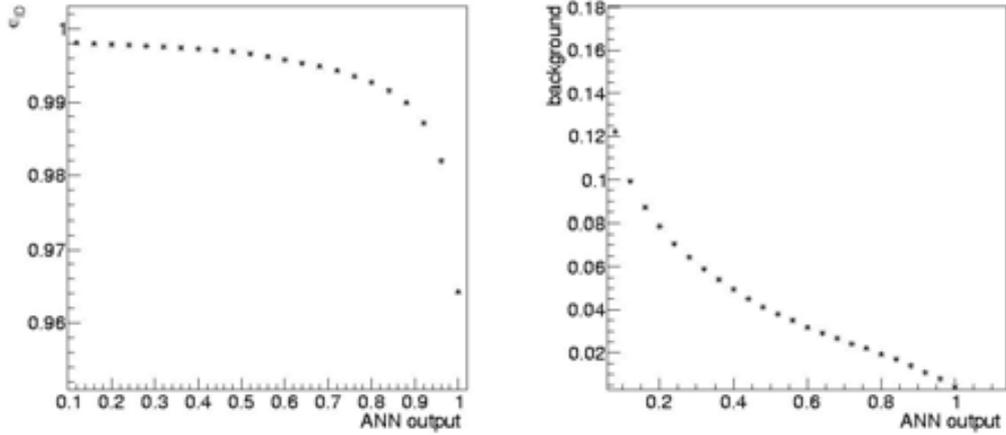


Figure 6.3: Muon ϵ_{ID} and background reduction rate as function of Decay In Flight ANN output. The cut applied in the analysis on this ANN output is 0.725.

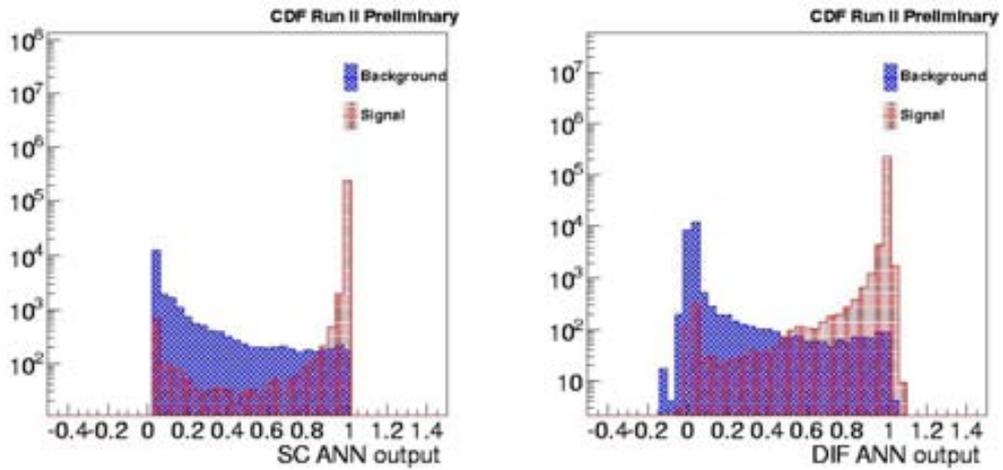


Figure 6.4: SC and DIF ANNs output.

background (Figure 6.5). Muons are selected if the SC ANN output is higher than 0.875 and the DIF ANN output is higher than 0.725. With this selection the Z/γ^* reconstruction efficiency is increased by $\sim 34\%$.

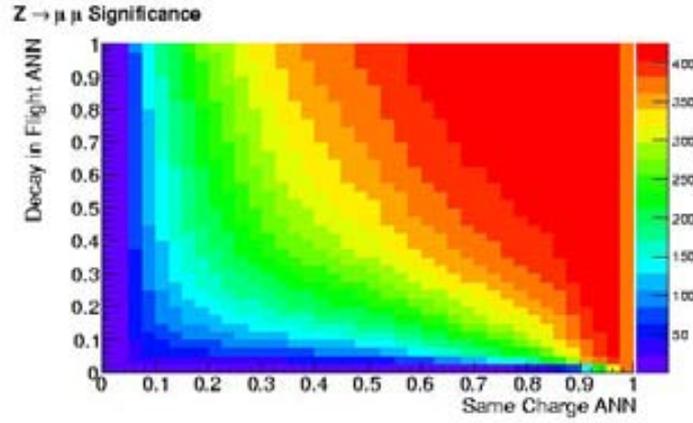


Figure 6.5: $Z/\gamma^* \rightarrow \mu^+\mu^-$ inclusive significance as a function of the output of the SC and DIF ANNs.

6.4.2 Electron identification and $Z/\gamma^* \rightarrow e^+e^-$ reconstruction

The electron identification is done similarly to the muon case using the help of two artificial neural networks. It is based on the discrimination between real high- p_T electrons coming from $Z/\gamma^* \rightarrow e^+e^-$ decay (probes) and jets that could fake the electron signal. Both samples are selected from data, probe electrons from high- p_T dataset and fakes from a dijet dataset according to the following criteria:

Probes electrons from $Z/\gamma^* \rightarrow e^+e^-$

- One very tight electron (tag electron)
 - $E_T \geq 20$ GeV
 - $p_T \geq 10$ GeV/c
 - $L_{shr} \leq 0.1$
 - $E/P \leq 1.2$
 - $E_T(\Delta R \leq 0.4)/E_T(e) \leq 0.05$
 - firing *ELECTRON_18* trigger

- Discard events with $\cancel{E}_T \geq 20$ GeV, to reduce W +jets contribution

6 The $Z/\gamma^* + b$ -jet cross section measurement

- Select probe electron passing basic kinematic cuts:
 - $E_T \geq 20$ GeV for central or plug electron
 - $p_T \geq 10$ GeV/c only for central electrons
 - associated to a tag electron so that the two electrons have invariant mass in [86–96] GeV/c² range and opposite charge if both are central

Fake electrons:

- candidates from dijet dataset, where there is at least one jet with $p_T \geq 20$ GeV/c and one fakeable object (electron matched with a jet)
- discard events with $\cancel{E}_T \geq 20$ GeV
- the fakeable object should pass some basic kinematic cuts:
 - $E_T \geq 20$ GeV for central and plug
 - $p_T \geq 10$ GeV/c and $|z_0| \leq 60$ cm only if it is central
- should exist a jet with $p_T \geq 20$ GeV/c that is matched with the fakeable object. This jet should not be the leading jet to avoid trigger bias.
- combination of the leading jet and fakeable object should be outside the Z mass peak

Two different ANNs are trained to differentiate real electrons from fakes: one for central and the other for plug electrons.

The central ANN employs the following variables:

- CES χ^2 ²
- E_{Had}/E_{Em}
- Isolation/ E_T
- Number of Si Hits
- Number of Axias COT SL with at least 5 Hits
- Number of Stereo COT SL with at least 5 Hits
- Number of tracks associated to the electron candidate
- E/P

while the plug ANN utilizes:

²Comparison of the CES shower profile on the $r - z$ plane to the expected profile as measured for electrons in the test beam

- Pem χ^2 ³
- Pes U energy/ Pes average E ⁴
- Pes V energy/ Pes average E
- ΔR Pem-Pes
- Pem cluster energy / E_T
- E_{Had}/E_{Em}
- Isolation/ E_T

Figures 6.6 and 6.7 show the distribution comparison between fakes and probes for the variables in the ANN, while the output of the two ANNs is presented in Figure 6.8. Central electrons are selected using a cut of 0.8 on the ANN output, while for the plug electrons a cut of 0.4 on the corresponding ANN is applied. These cuts are chosen based on the ϵ_{ID} and background rate shown in Figures 6.9 and Figure 6.10. With these values, the background rate is low and comparable to that obtained using rectangular cuts but the Z/γ^* reconstruction efficiency is larger by $\sim 42\%$.

6.4.3 Jet selection

Jets are identified using the MidPoint algorithm with a cone size of $R=0.7$ and a merging/splitting fraction set to 0.75. In data and MC, jets are clustered using calorimeter towers with transverse momentum above 0.1 GeV/c and seed towers of 1 GeV/c after excluding the towers associated to leptons from the Z/γ^* boson decay. In addition we also require a minimum distance between jet and leptons: $\Delta R_{lepton-jet} \geq 0.7$. After reconstruction the momentum of the jet is corrected following the prescription explained in Chapter 5. Jets are not corrected for underlying event or energy loss due to out-of-cone parton radiation. In fact, the measurements are performed at hadron level and these corrections (UE and hadronization) are applied to the theoretical predictions (see Chapter 8).

Jets are required to have $p_T \geq 20$ GeV/c and $|Y| \leq 1.5$. The rapidity cut is due to the reduction of b -tagging efficiency at larger rapidity values. It falls rapidly since a requirement of COT hits is made by the SecVtx algorithm to maintain a good track purity.

b tagged jets

Jets are further required to have a reconstructed secondary vertex within a cone of 0.4 around the jet axis. The secondary vertex is reconstructed using the TIGHT SECVTX package [98]. The efficiency of reconstructing a b -jet has been measured in data and Monte Carlo [?] and varies between 30 % and 40% for the jet p_T and rapidity range of interest. Following these studies the Monte Carlo efficiency is adjusted by a scale factor of 0.96 ± 0.05 .

³Value of the 3x3 PEM cluster energy distribution as compared to the hypothesis that EM object is an electron cluster

⁴ Pes average energy is calculated as : (Pes U energy +1.2 * PES V energy)/2.2

6 The $Z/\gamma^* + b$ -jet cross section measurement

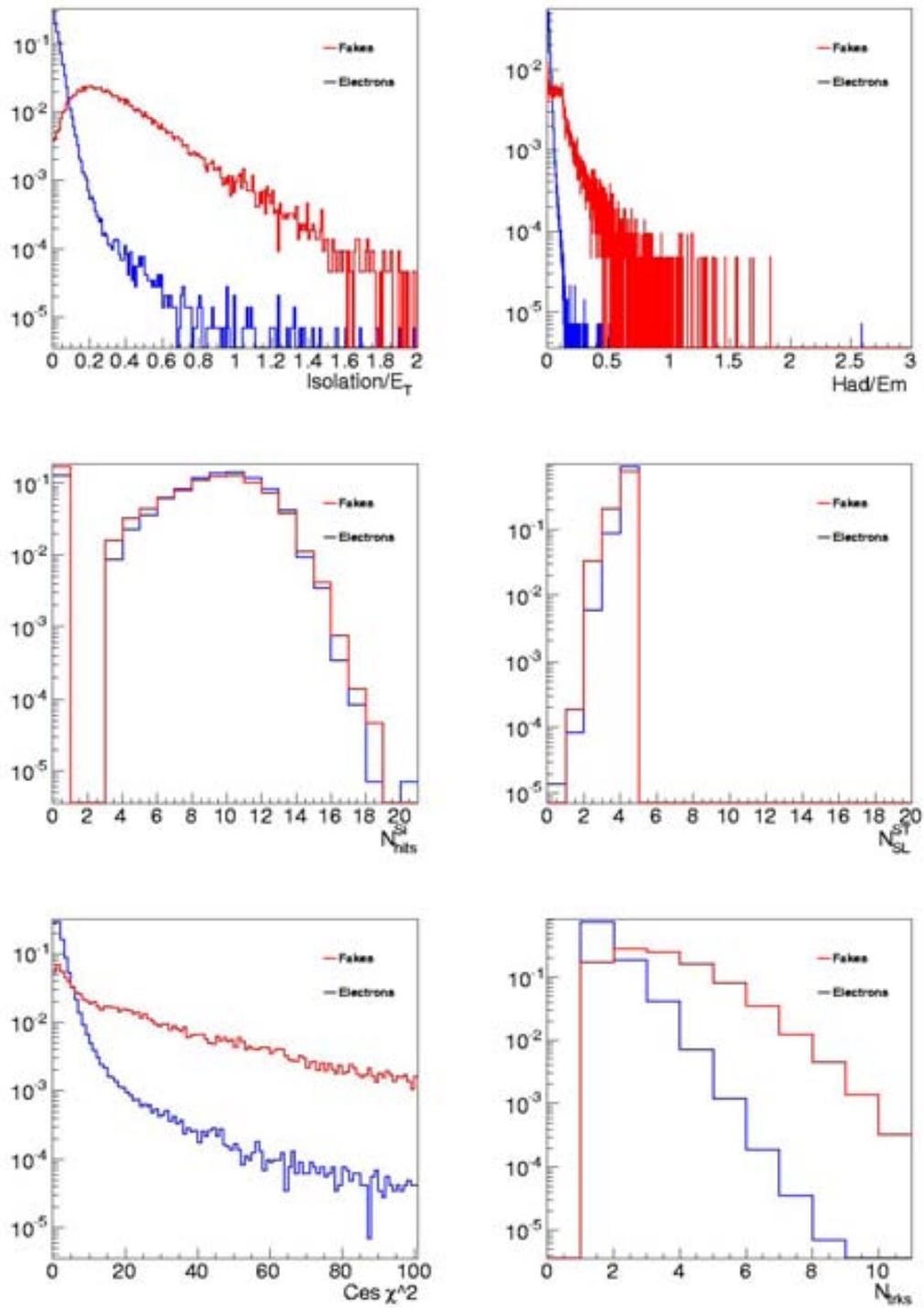


Figure 6.6: Comparison between fake and probe central electron distributions for the variables used in the central ANN.

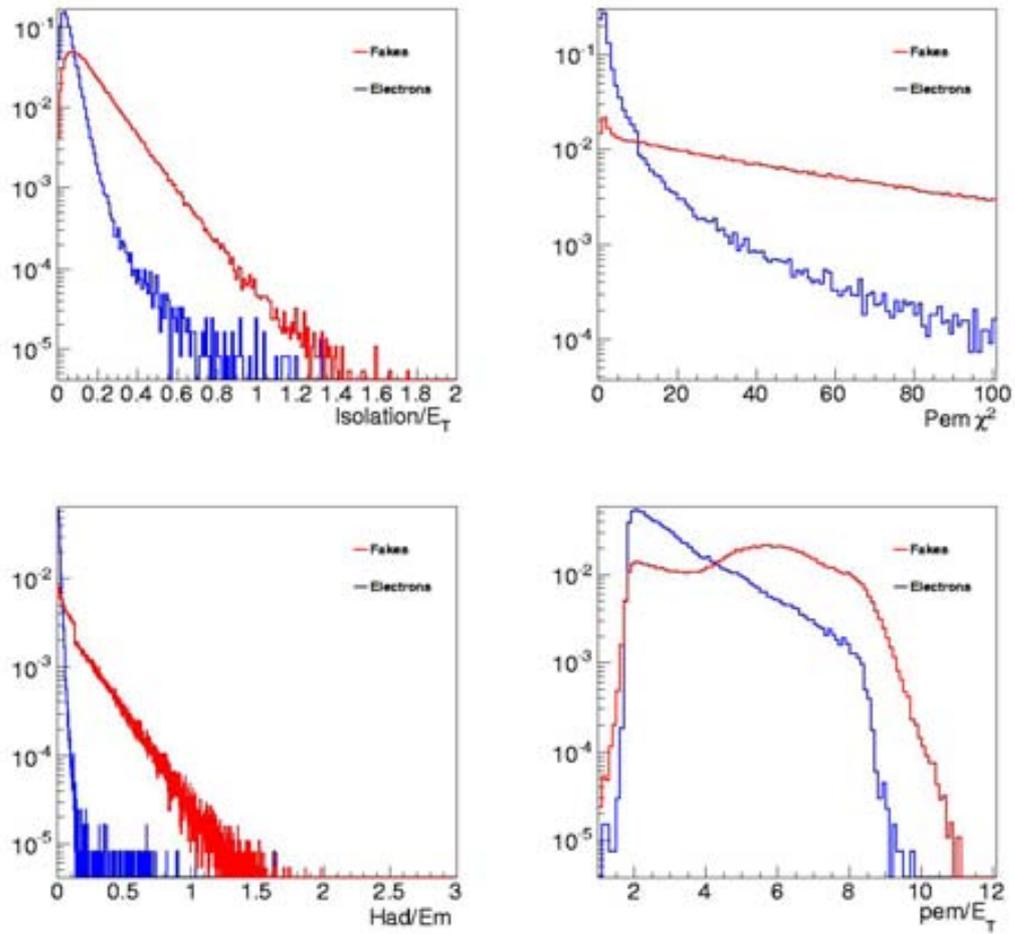


Figure 6.7: Comparison between fake and probe plug electron distributions for the variables used in the plug ANN.

6 The $Z/\gamma^* + b$ -jet cross section measurement

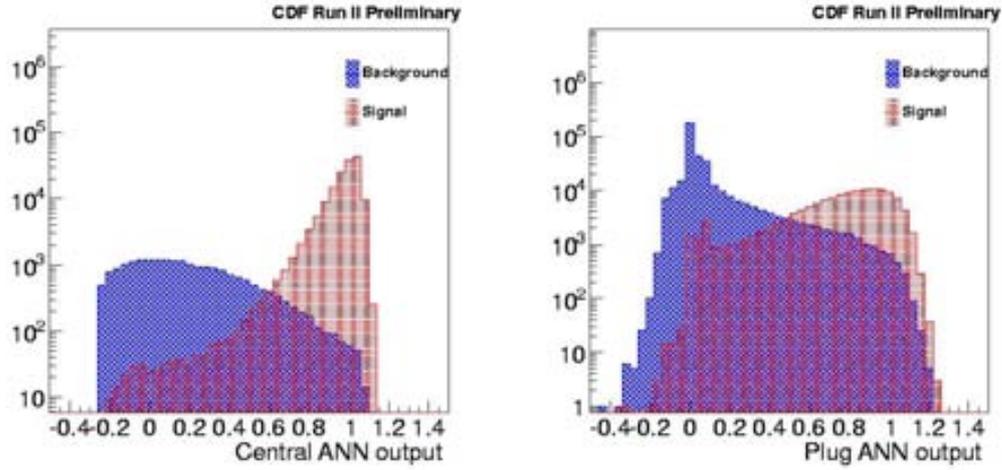


Figure 6.8: Output of the central and plug ANNs

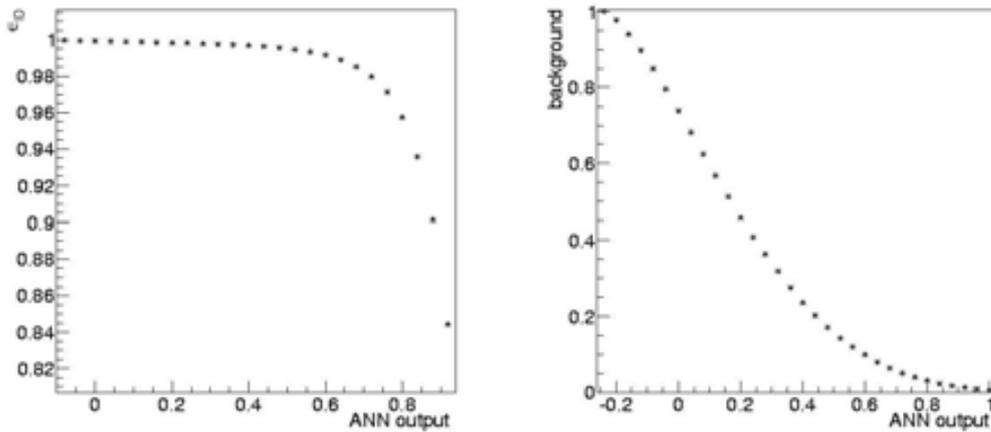


Figure 6.9: Electron ϵ_{ID} and background reduction rate as a function of central ANN

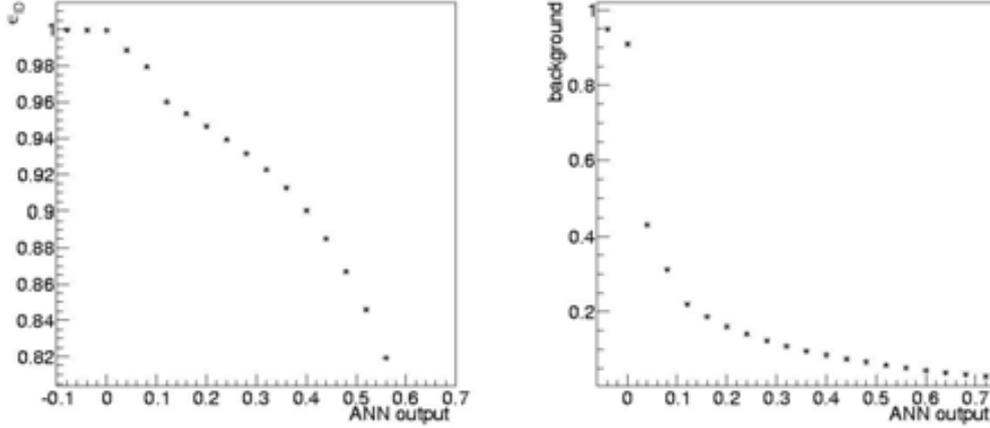


Figure 6.10: Electron ϵ_{ID} and background reduction rate as a function of plug ANN

6.4.4 Event selection Summary

From the full dataset collected during Run II, a sample of $\sim 9.1 fb^{-1}$ of integrated luminosity, passing high- p_T lepton triggers requirements is analyzed. Events are selected using the following criteria:

- at least one reconstructed primary vertex with z -position within 60 cm from the nominal interaction point
- a $Z/\gamma^* \rightarrow l^+l^-$ ($l = e, \mu$) boson reconstructed
 - in muon channel requiring two high p_T central muons with opposite charges
 - in electron channel asking for one central electron and a second electron that can be central or plug
 - with $66 \leq M_{ll} \leq 116 \text{ GeV}/c^2$
 - $p_T^l \geq 20 \text{ GeV}/c$
- at least one jet (MidPoint cone size $R=0.7$) Tight SecVtx tagged with $p_T \geq 20 \text{ GeV}/c$, $|Y| \leq 1.5$

Table 6.1 summarizes the observed candidate events for different stages of the event selection.

6.5 Background Modeling

Several physic processes have signals that mimic the $Z/\gamma^* + b$ -jet one. These processes include diboson (ZZ, ZW, WW) and top pair production and are modeled using MC samples. Another source of contamination, which may be important in the electron channel if jets are

6 The $Z/\gamma^* + b$ -jet cross section measurement

	Data	
	$Z/\gamma^* \rightarrow \mu^+\mu^-$	$Z/\gamma^* \rightarrow e^+e^-$
N_{Z/γ^*}	303 194	540 734
N_{Z/γ^*-jet}	53 941	84 519
$N_{Z/\gamma^*-b-tag-jet}$	856	1085

Table 6.1: Observed events for the electron and muon channels at different stages of the event selection.

misidentified as electrons, is originated from multi-jet production and W +jets events (QCD) and is estimated using data driven technique.

6.5.1 Diboson and $t\bar{t}$

The diboson and top pair production background are obtained using Pythia MC samples scaled to the data integrated luminosity and normalized to their theoretical cross sections (Table 6.2).

Backgrounds	Monte Carlo	Integrated luminosity
ZZ	Pythia	21518 fb^{-1}
WZ	Pythia	28677 fb^{-1}
WW	Pythia	5967 fb^{-1}
$t\bar{t}$	Pythia	860 fb^{-1}

Table 6.2: MC samples used to estimated the expected background in the analysis.

6.5.2 QCD and W + jets backgrounds

In the muon channel QCD, W +jets and DIF backgrounds are evaluated using data-driven techniques. Events are selecting using the same muon criteria as described in section 6.4.1 but instead of requiring two muons tracks to be opposite in charge they are required to have the same charge⁵. This is done because fakes and real muons have similar probability to have same charge or opposite charge with respect to another muon identified in the same event.

For the electron case the evaluation is more involved. QCD multi jet events are important backgrounds, since two hadronic jet can be misidentified as electrons and fake the Z/γ^* signal. W +jets also contributes significantly as a background, in particular when an hadronic jet is

⁵In this case we are requiring only same charge muons without the additional cuts in $\frac{\chi^2}{n.d.f.}$ and in Δz_0 and so also DIF fakes are included

misreconstructed as electron and forms a fake Z/γ^* with the real electron from W decay. Also here a data-driven method is used to evaluate these contributions. This consist in measuring the fake rate, that is the probability of a jet to be identified as an electron in data. The expected background is obtained applying the fake rate to jets in the high- p_T lepton data sample.

Fake rate.

The fake rate is defined as the probability that a jet passes the electron selection and is identified as an electron candidate. This probability is evaluated in a dijet sample where events with more than one electron and missing $\cancel{E}_T \geq 15$ GeV are excluded to reject electrons coming from the decay of the Z/γ^* boson. Therefore the fake rate is defined as:

$$fake_rate = \frac{N_{jets}(that\ pass\ electron\ cuts)}{N_{fakeable}(that\ are\ suitable\ to\ pass\ electron\ cut)} \quad (6.4)$$

where a fakeable is a jet that satisfies the kinematic requirements of the electron selection and the numerator is the number of such jets passing the electron identification cuts. A different fake rate is calculated for each electron category (central or plug). The energy of the jet is smeared to better model that of the electron. Due to the differences in the reconstruction algorithms of the jets and electrons, a jet of E_T^{jet} would fake an electron of E_T^{ele} with $E_T^{ele} \leq E_T^{jet}$. The energy scale factors E_T^{ele}/E_T^{jet} are shown in Figure 6.11.

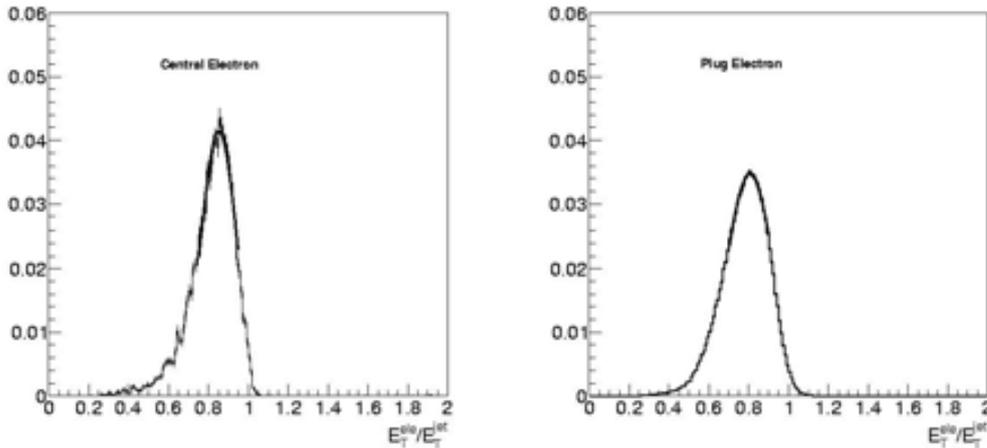


Figure 6.11: Scale factors for central and plug electron fakes. The distributions are fitted with a Gaussian distribution.

Once the fake rates (in Figure 6.12) and scale factors are obtained, the high p_T electron sample is used to estimate the background from QCD and W +jets. Real Z/γ^* events are rejected selecting events with one and only one electron. Every electron-jet pair is considered

6 The $Z/\gamma^* + b$ -jet cross section measurement

with a weight equal to the fake rate of the jet, adjusted by the probability for the others to not fake an electron.

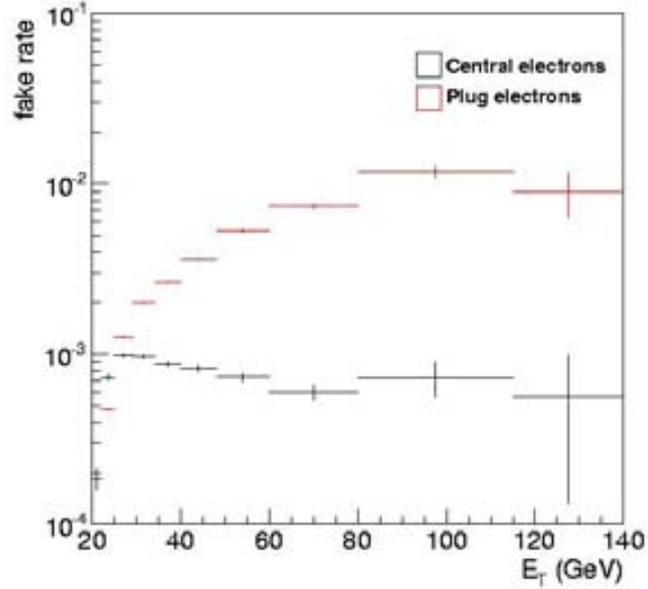


Figure 6.12: Fake rate for central and plug electrons.

6.6 Pretag Sample: Data - Monte Carlo Comparison

The pretag sample was used to compare Data and Monte Carlo in order to validate the modeling. Distributions of variables, such as p_T and absolute rapidity of the jets, as well as the invariant mass of the two leptons are presented in Figure 6.13. A good agreement is observed in all cases. For this comparison Alpgen+Pythia MC prediction has been normalized to data.

6.7 Composition of b -tagged sample

The tagged jet sample is contaminated by diboson, $t\bar{t}$ and QCD events, however the largest contributions are due to $Z/\gamma^* + \text{jets}$ events where charm and light/gluon jets (LF) are misidentified as b -jets by the SecVtx tagger. In order to estimate the fraction of b -jets we perform a fit to the invariant mass of the tracks that define the secondary vertex (M_{SecVtx}).

6.7 Composition of b -tagged sample

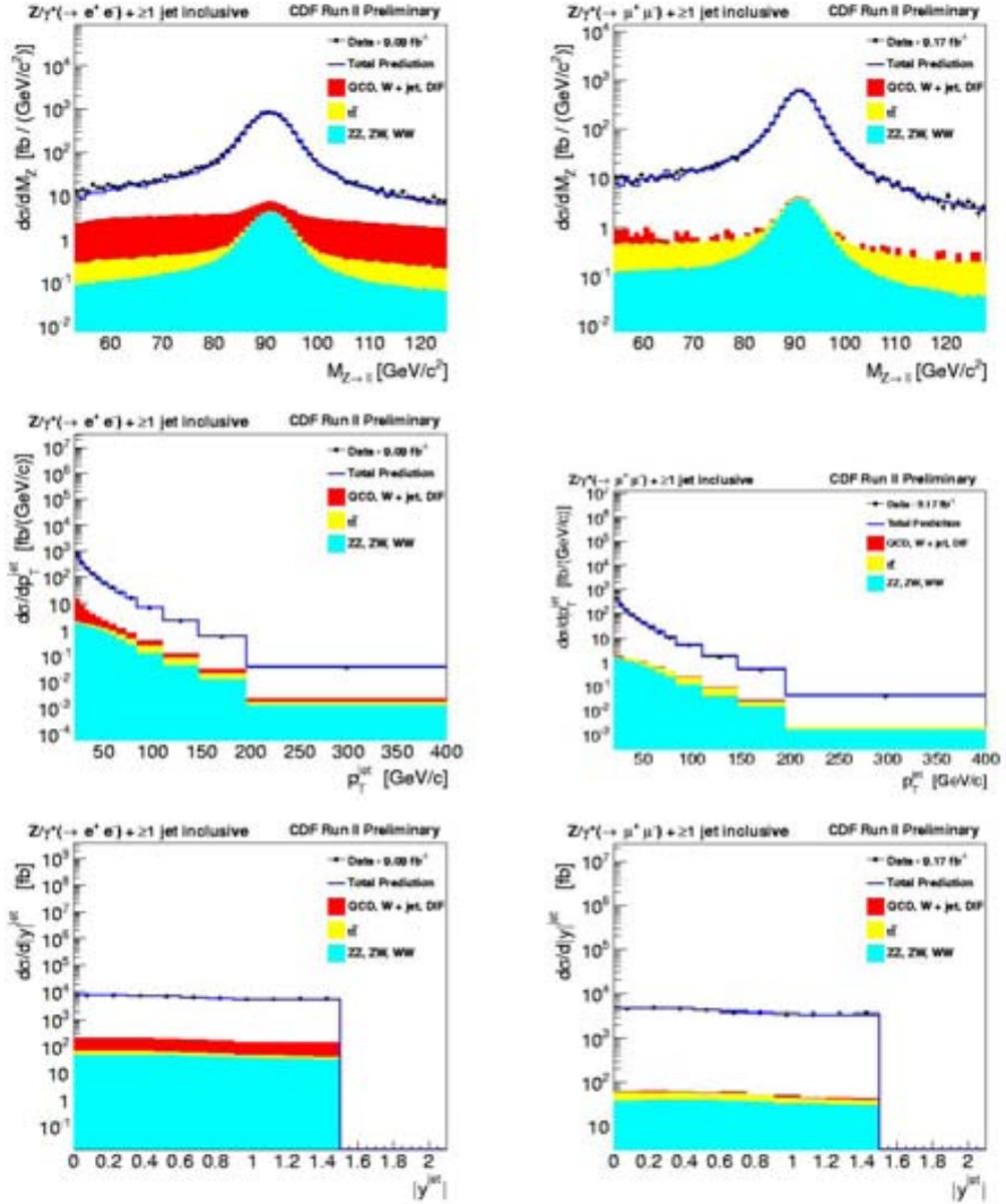


Figure 6.13: Data-MC comparison for M_{Z/γ^*} , p_T and rapidity of jets for electron and muon channel in the pretag sample.

6.7.1 Fitting procedure

As can be seen in Figure 6.14 the mass of the tracks constituting the secondary vertex is related to the mass of the particles decaying at that point. Templates of tagged b -jet, c -jet, LF -jet are built using Alpgen+Pythia $Z/\gamma^* + Np$ sample.

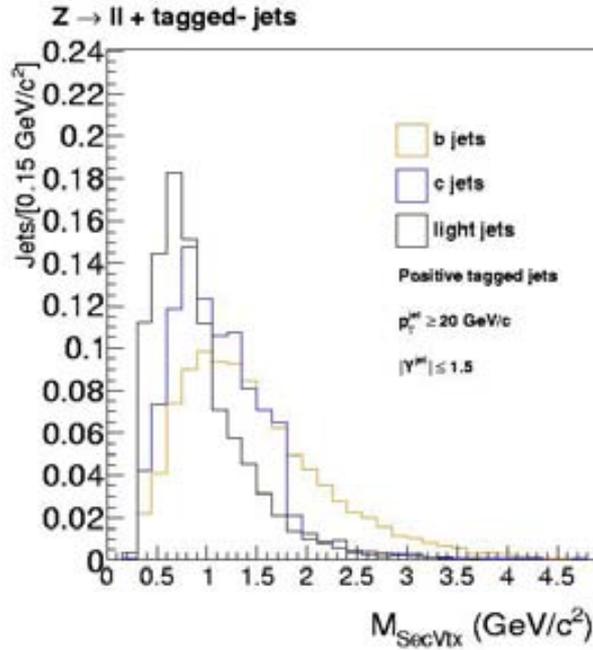


Figure 6.14: Secondary vertex mass templates for TIGHT SECVTX tagged b , c , LF jets.

Tagged jets have been matched within $\Delta R \leq 0.4$ to either bottom and charm hadrons. A tagged jet that has not been matched is assigned as a LF jet. To avoid contaminations in the LF sample from HF , samples that have HF in the matrix element are vetoed. The main contribution of tags LF jets comes from the $Z/\gamma^* + 1p$ sample.

The feature in the charm shape near $M_{SecVtx} = 1.8 \text{ GeV}/c$ is attributable to D (mass = $1.865 \text{ GeV}/c$) and D^\pm (mass = $1.869 \text{ GeV}/c^2$) vertices for which the invariant mass very nearly reproduces that of the parent. The feature can be seen also in the b shape. These cases correspond to $B \rightarrow D^\pm/D^0 X$ decays, whose secondary vertex contains tracks from the tertiary charm state decay.

When building the templates, the contribution from each Alpgen+Pythia sample is weighted according to its prediction. The weighted templates are shown in Figure 6.15 and the normalized ones are in Figure 6.16. Normalized templates show that M_{SecVtx} shapes vary only slightly with the process of origin.

Having found no significant differences between electron and muon templates (Figure 6.17), the final templates have been obtained by merging both samples. In this way we benefit from a larger statistics which in turn reduce the uncertainties on the b -fraction.

6.7 Composition of b-tagged sample

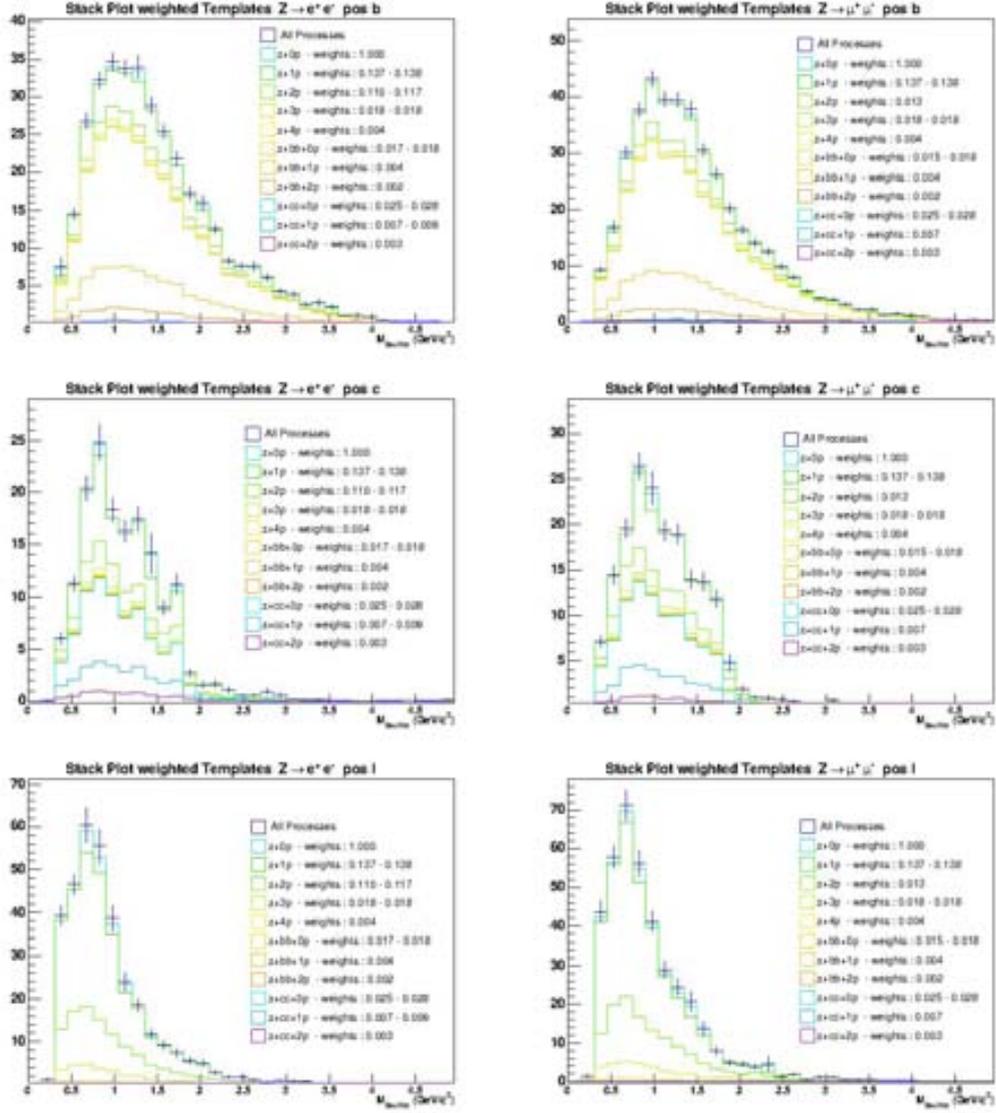


Figure 6.15: Secondary vertex mass templates for Tight SECVTX tagged b, c, LF jets. Plots show the contribution from each process and its corresponding weight.

6 The $Z/\gamma^* + b$ -jet cross section measurement

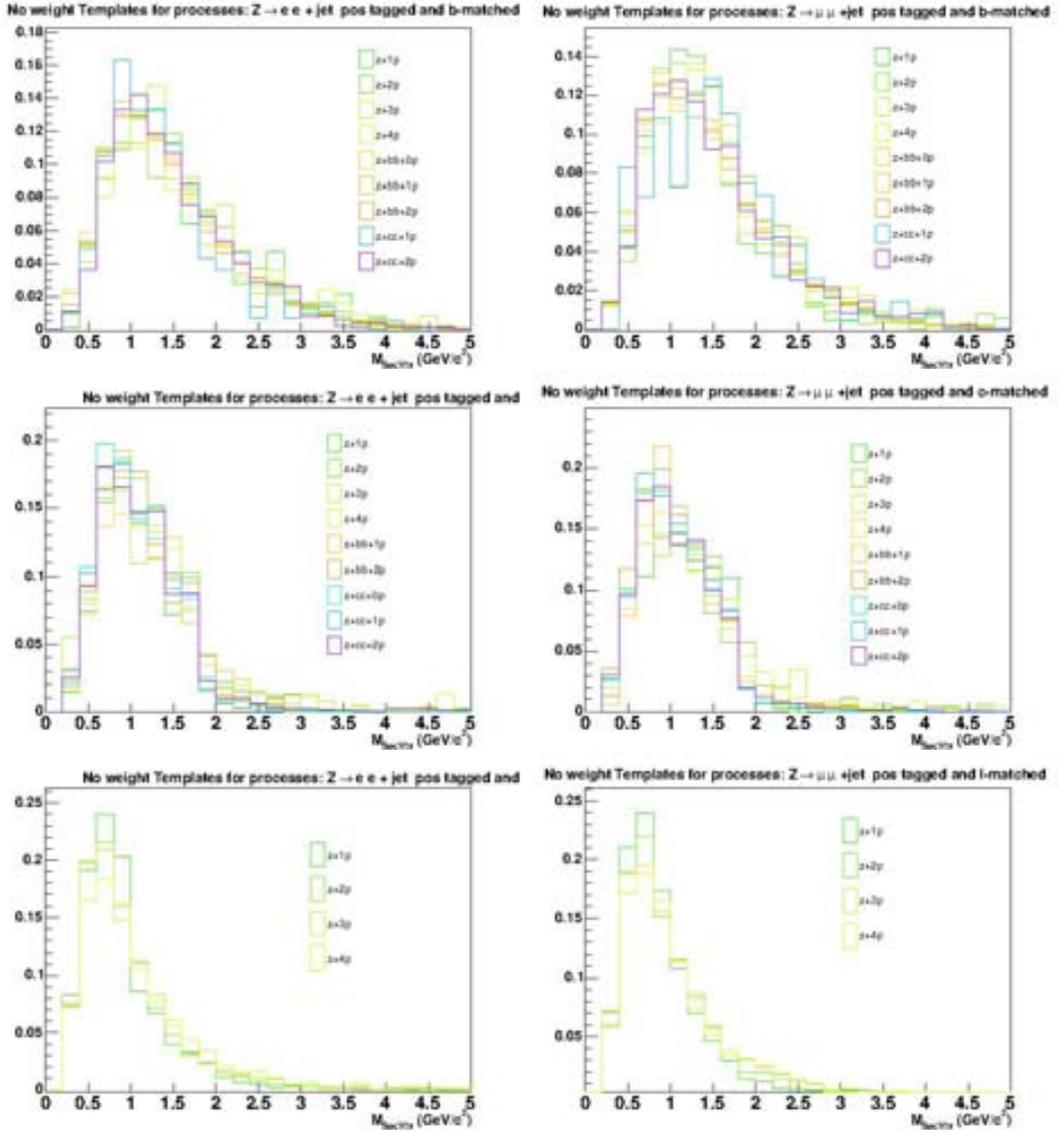


Figure 6.16: Secondary vertex mass templates for Tight SECVTX tagged b, c, LF jets for different processes, normalized to 1.

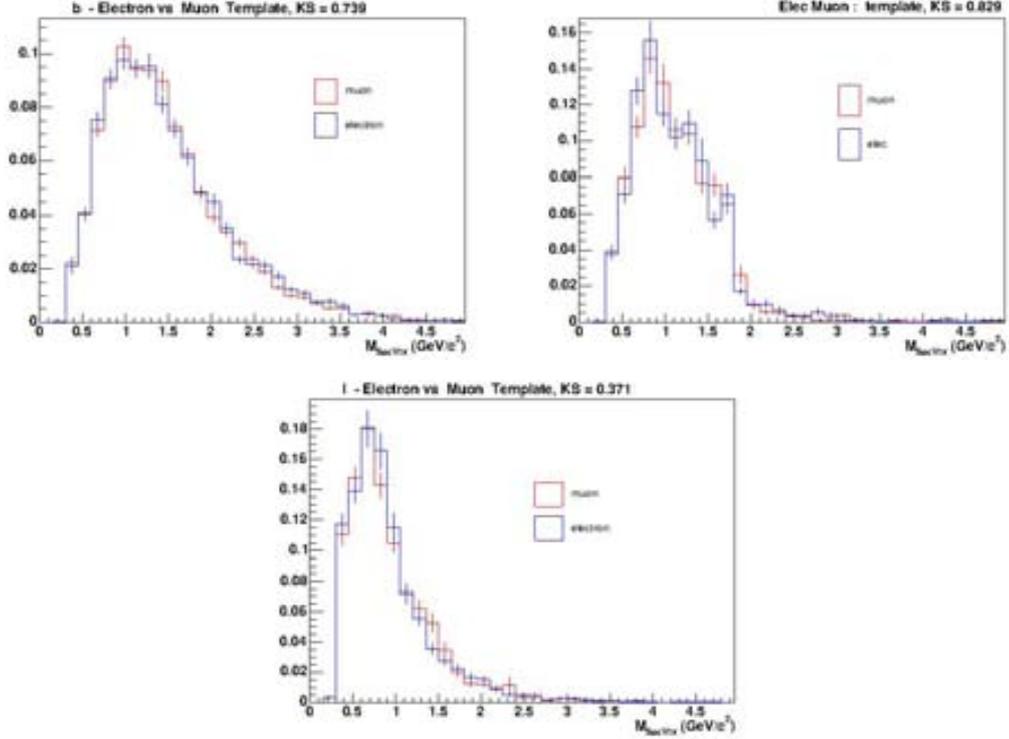


Figure 6.17: Comparison between electron and muon channel SecVtx templates for the b , c and LF jet flavors. The Kolmogorov statistical test (KS) on the top of the plots shows that there is a good agreement.

Some differences were found on the b template between contributions from Monte Carlo signal and $t\bar{t}$ background, as it is shown in Figure 6.18. For this reason the $t\bar{t}$ expected contribution is incorporated to the b -template.

Once the templates are built, the components from b , c and LF jets are obtained via a binned maximum likelihood fit. The expected number of total tagged jets in bin i can be expressed as:

$$\mu_i = N_{Tot} \cdot [f_b^{fit} \cdot N_b^i + f_c^{fit} \cdot N_c^i + f_{LF}^{fit} \cdot N_{LF}^i] \quad (6.5)$$

where N_j^i is the normalized contribution of flavor j ($j = b, c, LF$) given by the M_{SecVtx} template, f_j is the fit fraction for species j and N_{Tot} is the number of tagged jets in data sample. The sum of the fit fractions is constrained by:

$$f_b^{fit} + f_c^{fit} + f_{LF}^{fit} = 1 \quad (6.6)$$

The Poisson probability $P(n_i|\mu_i)$ of observing n_i tagged jets in bin i of a secondary vertex distribution given μ_i , is:

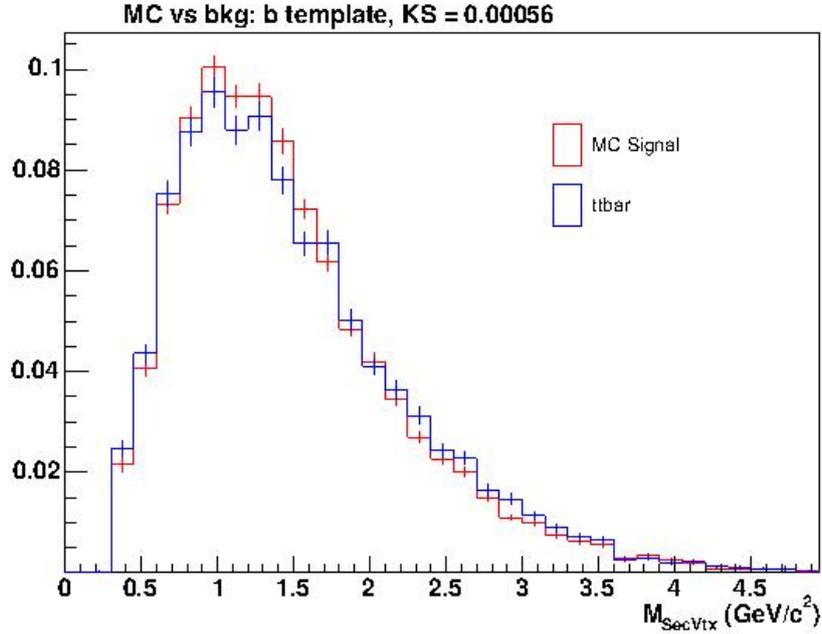


Figure 6.18: b SecVtx Mass template comparison between MC signal and background. The Kolmogorov statistical test (KS) shows that there is not a good agreement.

$$P(n_i|\mu_i) = \frac{e^{-\mu_i} \cdot \mu_i^{n_i}}{n_i!} \quad (6.7)$$

Therefore the Likelihood function \mathcal{L} can be constructed as:

$$\mathcal{L} = \prod_{i=1}^{N_{bins}} P(n_i|\mu_i) = \prod_{i=1}^{N_{bins}} \frac{e^{-\mu_i} \cdot \mu_i^{n_i}}{n_i!} \quad (6.8)$$

for the N_{bins} of the M_{SecVtx} templates.

The fit fractions are obtained by maximizing the likelihood:

$$\ln(\mathcal{L}) = \ln \left(\prod_{i=1}^{N_{bins}} \frac{e^{-\mu_i} \cdot \mu_i^{n_i}}{n_i!} \right) = \sum_{i=1}^{N_{bins}} [-\mu_i + n_i \ln \mu_i + const] \quad (6.9)$$

The MINUIT package [?] implemented in ROOT is used for this purpose.

The result of the fit is shown in Figure 6.19 where the fraction of b -jets obtained is 0.47 ± 0.04 , thus the number of b -tagged jets in the data are 908 ± 80 .

6.7.2 Bias Checks

The fit procedure was validated using pseudo-experiments (PE). Pseudodata are drawn from the various species of M_{SecVtx} templates with the same statistics as data for a particular

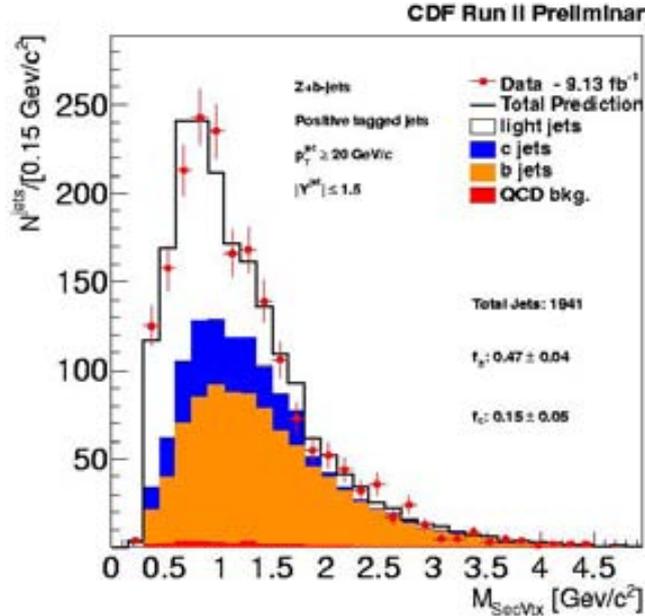


Figure 6.19: SecVtx Mass distribution for $Z/\gamma^*x \rightarrow l^+l^-$. In colors is shown the contributions from each flavor (as result of the fit)

choice of $b/c/LF$ fractions. The new M_{SecVtx} distribution, built from the pseudodata, is fitted with the same procedure described above. For each PE, f_b is obtained and histograms of the fitted f_b , its uncertainty and its pull $((f_b^{fit} - f_b^{input})/\sigma_{f_b}^{fit})$ are filled. The procedure is repeated for 5000 pseudo-experiments.

The pull distribution for the scenario $b/c/LF = 47/15/38$ is shown in Figure 6.20. The width of the pull, $RMS \sim 1.0$, as expected for properly defined uncertainties. No bias is observed in the mean of the distribution. The dependence on the input fraction was studied: a linear dependence, which shows no bias, is observed in Figure 6.21 where the value of the measured fractions is drawn as a function of the input.

6.7.3 Background Expectation

As explained previously, background contributions are coming from QCD (inclusive jets, $W + jets$ and decay in flight) and from other processes such as diboson (WW , ZZ , ZW) and $t\bar{t}$ jets. For the case of Z/γ^* and Z/γ^*+jets cross sections, the expected backgrounds are subtracted from the number of observed data events passing the corresponding selection requirements. In the case of the b -tagged jet sample, the predicted number of b -jet from tagged diboson and top pair production (Table 6.3) are subtracted from the fitted number of b -jets. The background originated from QCD (fakes) is included directly in the SecVtx Mass fit through a template obtained from data by weighting events by the fake rate (section 6.5.2).

6 The $Z/\gamma^* + b$ -jet cross section measurement

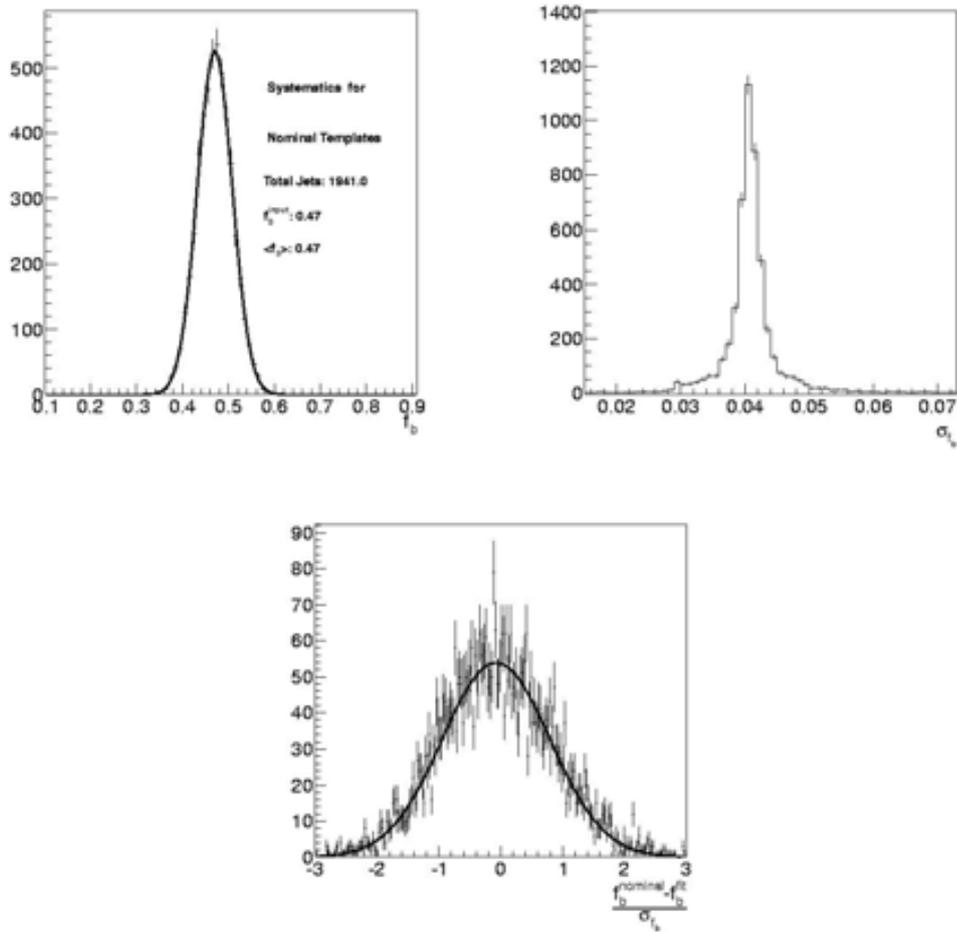


Figure 6.20: Fraction of b -jet, its uncertainty (σ_{f_b}) and pull obtained from pseudo-experiments for the scenario in the data, i.e. $b/c/LF=0.47/0.15/0.38$.

6.7 Composition of b -tagged sample

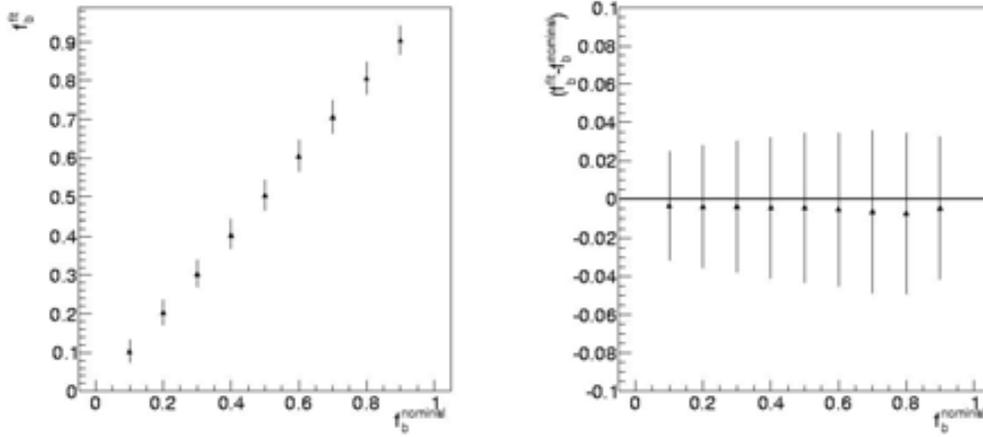


Figure 6.21: Results from pseudo-experiments using different input scenarios. Left: PE f_b mean as a function of input f_b . Right: residual means as a function of input f_b , the uncertainty is coming from the width of the Gaussian centered in f_b .

Background expectations are summarized in Table 6.4.

process	bkg $Z/\gamma^* \rightarrow \mu^+\mu^-$	process	bkg $Z/\gamma^* \rightarrow e^+e^-$
dibosons	14.9 ± 3.5	dibosons	15.9 ± 3.3
$t\bar{t}$	68.7 ± 6.9	$t\bar{t}$	65.0 ± 6.6
total	83.6 ± 10.4	total	80.9 ± 9.9

Table 6.3: The background contributions for $Z/\gamma^* + b$ tagged jets in muon and electron channel.

	Data		Fakes		MC Bkg.	
	muon	electron	muon	electron	muon	electron
N_{Z/γ^*}	303 194	540 740	115	4872	724 ± 75	798 ± 85
N_{Z/γ^*-jet}	53 941	84 520	43	1542	677 ± 70	745 ± 80
$N_{Z/\gamma^*-b-tag-jet}$	856	1085	1	27	83.6 ± 10.4	80.9 ± 9.9

Table 6.4: Events observed in data and expected backgrounds for all processes in the electron and muon channels

6.8 Acceptance

The acceptance for each process considered in this analysis, Z/γ^* inclusive cross section, $Z/\gamma^* + \text{jets}$ and $Z/\gamma^* + b$ -jets, was evaluated separately for the muon and electron channel and later combining weighting by its corresponding integrated luminosity. Details on the calculation are described below.

Acceptance for the Z/γ^* inclusive cross section

It is evaluated using the Pythia Z/γ^* inclusive sample. The numerator is defined as the number of the events that pass the Z/γ^* reconstruction cuts listed in previous section, while the denominator is represented by the number of generated events that have a Z/γ^* with a M_Z between 66 and 116 GeV/c^2

$$A_{Z/\gamma^*} = \frac{N \text{ } Z/\gamma^* \text{ reconstructed}}{N \text{ } Z/\gamma^* \text{ generated within mass range cut}}$$

Acceptance for $Z/\gamma^* + \text{jets}$

For the calculus of $Z/\gamma^* + \text{jets}$ acceptance, we considered the same selection as for the Z/γ^* inclusive cross section to which we have added the requirement of the presence of at least a jet that at both, calorimeter and hadron level, should pass the requirements $p_T \geq 20 \text{ GeV}/c$ and $|Y| \leq 1.5$.

$$A_{Z/\gamma^* - \text{jet}} = \frac{N \text{ } Z/\gamma^* \text{ reconstructed} + \text{calorimetric jets in phase space}}{N \text{ } Z/\gamma^* \text{ generated within mass range cut} + \text{hadron level jets in phase space}}$$

Acceptance for $Z/\gamma^* + b$ -jets

In this case the acceptance is obtained from the ratio between the number of b -tagged calorimeter jets found in events with a reconstructed Z/γ^* passing selection requirements and the number of b -matched hadron jets in events generated with a Z/γ^* passing selection requirements.

$$A_{Z/\gamma^* - b\text{tag-jet}} = \frac{N \text{ } Z/\gamma^* \text{ reconstructed} + \text{calorimetric jets in phase space tagged positively and } b - \text{matched}}{N \text{ } Z/\gamma^* \text{ generated within mass range cut} + \text{hadron level jets in phase space } b - \text{matched}}$$

6.9 Systematics

Sources of systematic uncertainties may affect the analysis by altering the shapes of M_{SecVtx} or the acceptances or background rates. Their contributions to the uncertainty on the cross section measurement are summarized in Table 6.7 and their estimation is discussed below.

Process	$Z/\gamma^* \rightarrow e^+e^-$	$Z/\gamma^* \rightarrow \mu^+\mu^-$
A_{Z/γ^*}	0.2754 ± 0.0002	0.1637 ± 0.0001
A_{Z/γ^*+jet}	0.3159 ± 0.0007	0.2224 ± 0.0005
$A_{Z/\gamma^*+b_tag_jet}$	0.0763 ± 0.0008	0.0666 ± 0.0007

Table 6.5: Acceptance table for electron and muon channel for different processes. The increment seen on the muon acceptance with respect to the electron one when adding the requirement of the presence of a jet in the event, is due to changes on the Z/γ^* rapidity distribution, as it becomes predominantly central.

6.9.1 Vertex Mass shape systematics

Since templates are built using MC, possible source of mismodeling have been analyzed. For each of them new templates are built to take into account the corresponding effect in M_{SecVtx} . Contributions are due to:

- **Track reconstruction inefficiency:** A track reconstruction inefficiency of 3 % was observed in MC [4]. New templates are built by randomly rejecting 3 % of tracks and recomputing the value of M_{SecVtx} . The differences between the templates are shown in Figure 6.22.
- **Single/Double B/C hadron in a jet** In Alpgen the fraction of $b\bar{b}$ to b -jets (and $c\bar{c}$ to c jets) is 0.23 (0.68). However previous studies[4] have presented differences of a factor of 3 (2) times. In order to take into account this effect we built the new b and c templates by weighting Alpgen events in such a way that the double fraction varies between 0 to 0.7 for b and between 0 to 1.36 for c . Figure 6.23 shows the modified templates.
- **Light flavor template systematic** The default light flavor template is made of positively tagged jets that have not been matched to any HF hadron in $Z/\gamma^* + Np$ Alpgen+Pythia MC samples. For mistag jets, the probability to be positively or negatively tagged is expected to be approximately the same. In order to evaluate a systematic uncertainty, a template from negatively tagged jets from data (mostly populated by mistags) is built. The data sample utilized is defined requiring a high- p_T lepton and high E_T (Figure 6.24).

The method to evaluate these systematics consists in performing pseudo-experiments (PE), building pseudo-data from the new templates and fitting them using the nominal ones. For example, let us consider the track reconstruction inefficiency systematic, the procedure is the following:

- build pseudo-data from the new M_{SecVtx} template, that includes the rejection of 3 % of the tracks

6 The $Z/\gamma^* + b$ -jet cross section measurement

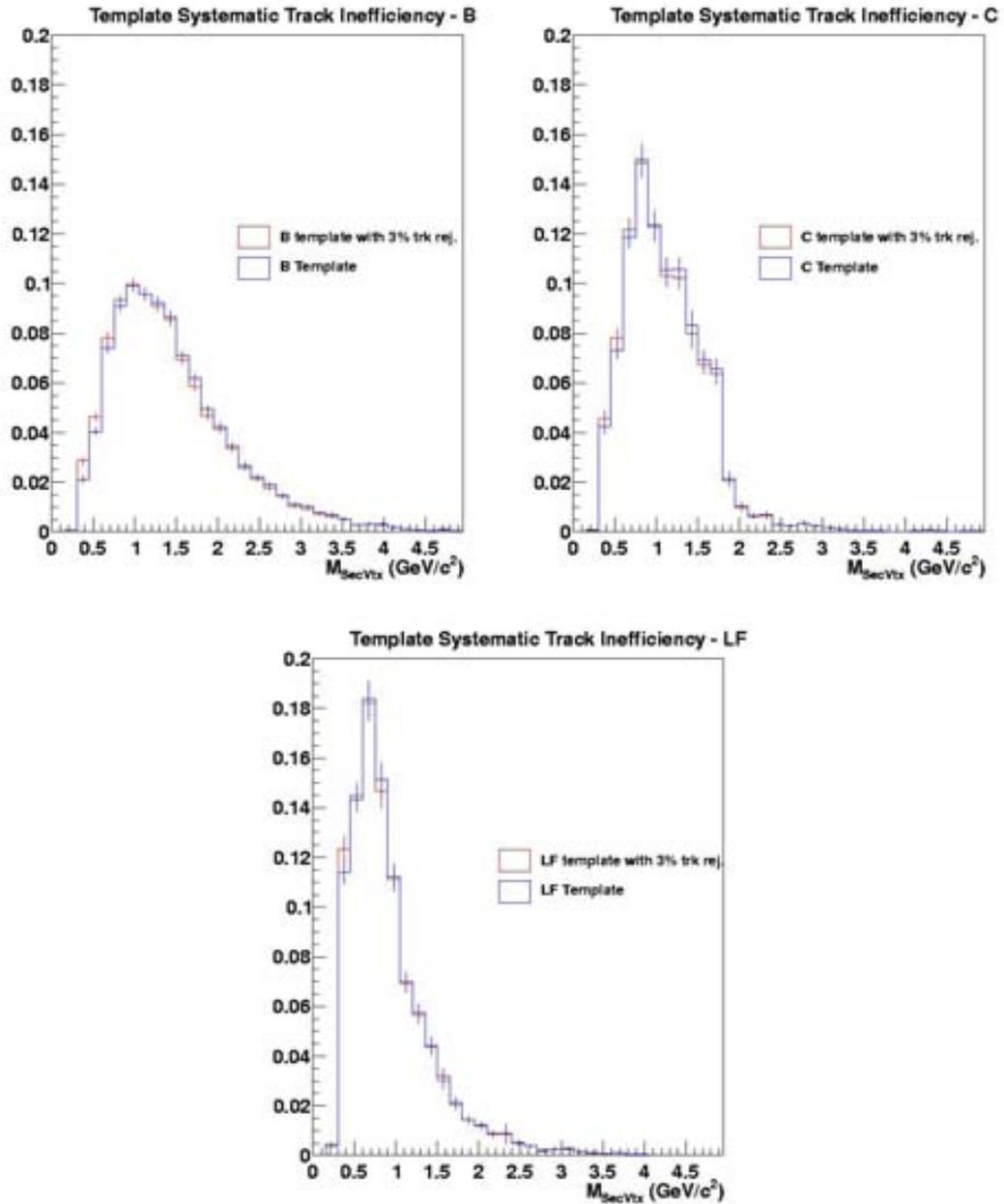


Figure 6.22: SecVtx Mass templates built with and without the 3% track rejection.

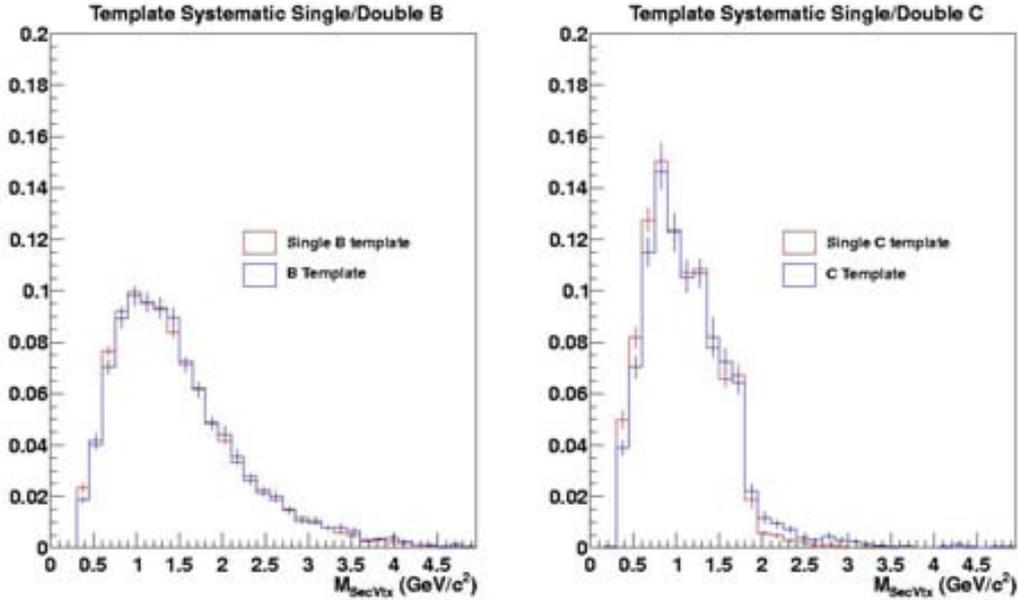


Figure 6.23: SecVtx Mass templates comparison for Single/Double b/c hadron in tagged jets

- fit pseudo-data using the original templates and find f_{b-syst}
- repeat the procedure 5000 times
- assign as systematic the shift between the mean value of the f_{b-syst} distribution and the nominal one obtained from PE drawn from nominal templates⁶.

6.9.2 Bootstrap

Uncertainties due to the limited MC statistics have been also estimated using the bootstrapping technique. This method consists in using the MC dataset to create multiple datasets by random selection from the same parent distribution. f_{b-syst} distributions are obtained using these new datasets. The RMS of the f_{b-syst} mean distribution is a measure of the uncertainty due to the limited MC statistics. We verified that these uncertainties are negligible. However it was not the case for the first attempt to build a negative tagged jets data template. The first sample used was the $Z/\gamma^* + \text{jets}$ sample but the statistics proved to be too low. A new template was obtained from a larger dataset as explained in the previous subsection.

⁶For the single/double b/c the half difference obtained by varying the $b\bar{b}/b(c\bar{c}/c)$ ratio from 0 to 3 (2) is used as systematic uncertainty.

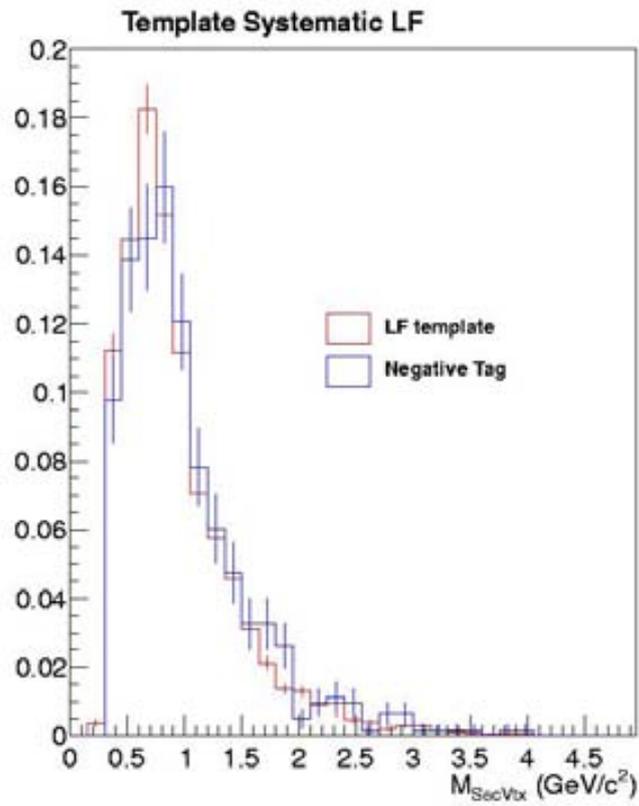


Figure 6.24: Comparison of SecVtx Mass templates constructed using LF MC jets and Negative tags from data

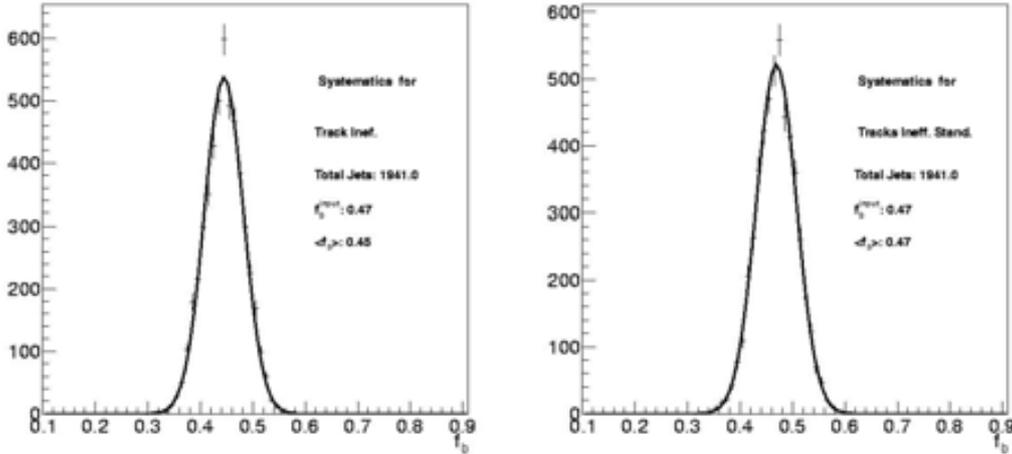


Figure 6.25: Example of the template systematic uncertainty evaluation where are presented f_b distributions for PE using nominal and track reconstruction inefficiency templates. The shift in $\langle f_b \rangle$ is used as systematic.

6.9.3 Acceptance systematics

The acceptance in the cross section measurement is affected by the JES and the b -tag SF uncertainties.

The uncertainty on the jet energy corrections is described in Chapter 5. These corrections are designed to properly scale back to particle level the energy of jets measured by the CDF II detector. To evaluate this uncertainty, we shift the jet energies in the signal MC by $\pm 1\sigma$ for each of the jet energy corrections individually. For each correction the uncertainty is assigned as half the difference between the two shifted new acceptance values. The total jet energy scale uncertainty is taken as the quadrature sum of the uncertainty from all 3 jet energy corrections. The relative effect on $A_{Z/\gamma^*+b-jet}$ is of 2% while in the case of A_{Z/γ^*+jet} is larger, $\sim 8\%$, due to differences on the transverse momentum spectra.

The second source of systematic error in the acceptance rises from the imprecise knowledge of the b -tag efficiency in data. This uncertainty is introduced in the measurement through the error of the b -tag SF (0.96 ± 0.05). Other systematics that affects the acceptances are the errors on the trigger, lepton ID and z vertex efficiency, but these are negligible and canceling in the ratio.

The final acceptances uncertainties are found in the table 6.6.

6.9.4 Background subtraction systematics

A systematic error was estimated due to the background normalization uncertainty. For background evaluated with MC the main part of the uncertainty is due to the cross section of the physics process ($\sim 20\%$ for di-boson, and $\sim 10\%$ for top) and the remaining from the luminosity uncertainty ($\sim 5.8\%$). The effect in the cross section is less than 0.1 %.

6 The $Z/\gamma^* + b$ -jet cross section measurement

Process	$Z/\gamma^* \rightarrow e^+e^-$		$Z/\gamma^* \rightarrow \mu^+\mu^-$	
	tag(%)	JES(%)	tag(%)	JES(%)
A_{Z/γ^*+jet}	/	9.3	/	7.9
$A_{Z/\gamma^*+b_tag_jet}$	5.2	2.2	5.2	1.5

Table 6.6: Systematic uncertainties for the acceptances due to jet energy scale (JES) and tag efficiency (tag)

On the other side, for data-driven backgrounds, in the muon case, a 100 % uncertainty is applied but its effect is negligible. For electrons, distributions in Figure 6.12 are fitted and error of 15 % is applied to cover possible deviations from the fit but also in this case the effect is negligible.

Systematics	$\frac{\sigma_{Z_bjet}}{\sigma_Z}$ (%)	$\frac{\sigma_{Z_bjet}}{\sigma_{Zjet}}$ (%)
Acceptance Systematics		
Jet Energy Scale abs	1.1	2.8
Jet Energy Scale mpi	1.4	5.9
Jet Energy Scale eta	0.6	2.0
b tag efficiency	5.2	
Templates Systematics		
Light Templates - data	8.7	
Double 1b/2b	1.5	
Double 1c/2c	2.7	
Tracks Inefficiency	5.1	
Others		
Background subtraction	0.01	0.01

Table 6.7: Summary of the systematics that affects the cross section measurement.

The total systematic uncertainty is obtained summing quadratically each contribution and is $\sim 11\%$ for the ratio with respect to the Z/γ^* inclusive cross section, and 13 % for the ratio with respect to Z/γ^*+jets . The systematics is comparable to the statistical uncertainties. The largest one is due to the tracks reconstruction inefficiency.

7 $Z/\gamma^* + b$ -jet Differential Cross Sections

The $Z/\gamma^* + b$ -jet cross section ratios is also measured differentially as a function of jet transverse momentum and jet rapidity. This chapter describes how the measurement was performed.

7.1 Measurement definition

The strategy utilized for these measurements is the same as for the integrated one (described in Chapter 6), but in this case data is divided in p_T^{jet} and $|Y^{jet}|$ bins and for each bin the complete analysis chain is repeated, including a reevaluation of the systematic uncertainties. Differential cross sections normalized to the inclusive Z/γ^* cross section are defined as:

$$\frac{d(\sigma_{Z/\gamma^*+b-jet}/\sigma_{Z/\gamma^*})}{d\alpha} = U_{had}^{cal}(Z/\gamma^*+b-jet)(\alpha) \times \frac{N_{Z/\gamma^*+b-jet}^{data} - N_{Z/\gamma^*+b-jet}^{bkg}}{\frac{N_{Z/\gamma^*}^{data} - N_{Z/\gamma^*}^{bkg}}{A_{Z/\gamma^*}}} \frac{1}{\Delta\alpha} \quad (7.1)$$

where:

- $\alpha = p_T^{jet}, Y^{jet}$
- $N_{Z/\gamma^*+b-jet}^{data(bkg)}$ and $N_{Z/\gamma^*}^{data(bkg)}$ are the number of $Z/\gamma^* + b$ -jet and Z/γ^* in data (expected background)
- A_{Z/γ^*} is the Z/γ^* acceptance times the event selection efficiency
- $U_{had}^{cal}(\alpha)$ the unfolding factors.

In order to remove effects from the experimental environment, the measurement is unfolded back to hadron level. The Unfolding factors, $U_{had}^{cal}(\alpha)$, are derived for each distribution using Monte Carlo and applied bin-by-bin.

7.2 Sample composition

Events are selected following the same prescription described in Chapter 6, but in this case divided in jet p_T and jet rapidity bins. Table 7.1 shows the number of tagged jets and expected background for each bin in electron and muon channels.

For each bin, the amount of b -tagged jets is obtained via a fit of the M_{SecVtx} distribution. Templates are created for each bin, combining electron and muon channel, and the fit is then performed via a maximum binned Likelihood. Results of the fit for different p_T jet bin

7 $Z/\gamma^* + b$ -jet Differential Cross Sections

p_T bins	$Z/\gamma^* \rightarrow \mu^+\mu^-$			$Z/\gamma^* \rightarrow e^+e^-$		
	data	fake	bkg	data	fake	bkg
20 - 27 GeV/c	173	1 ± 1	6.7 ± 1.1	223	5.4 ± 1.6	7.6 ± 1.2
27 - 35 GeV/c	155	0	9.0 ± 1.3	224	5.5 ± 1.6	9.6 ± 1.4
35 - 45 GeV/c	163	0	13.7 ± 1.9	194	5.0 ± 1.5	13.3 ± 3.6
45 - 60 GeV/c	117	0	19.4 ± 2.4	196	4.7 ± 1.4	18.7 ± 2.3
60 - 100 GeV/c	167	0	31.9 ± 3.4	175	3.0 ± 0.9	30.3 ± 3.4
$ Y^{jet} $ bins						
0.0 - 0.3	233	1 ± 1	25.8 ± 3.1	279	7.2 ± 2.2	25.6 ± 3.1
0.3 - 0.6	214	0	25.5 ± 3.2	265	7.6 ± 2.3	25.2 ± 3.1
0.6 - 1.0	247	0	25.8 ± 3.3	334	8.1 ± 2.4	25.9 ± 12.5
1.0 - 1.5	152	0	14.5 ± 1.9	207	4.5 ± 1.4	13.6 ± 1.8

Table 7.1: Number of observed tagged jets and expected background contributions. The muon fake error since this background is negligible is quoted to be 100 %. For the electron fake the uncertainty is coming from the fit of fake rate.

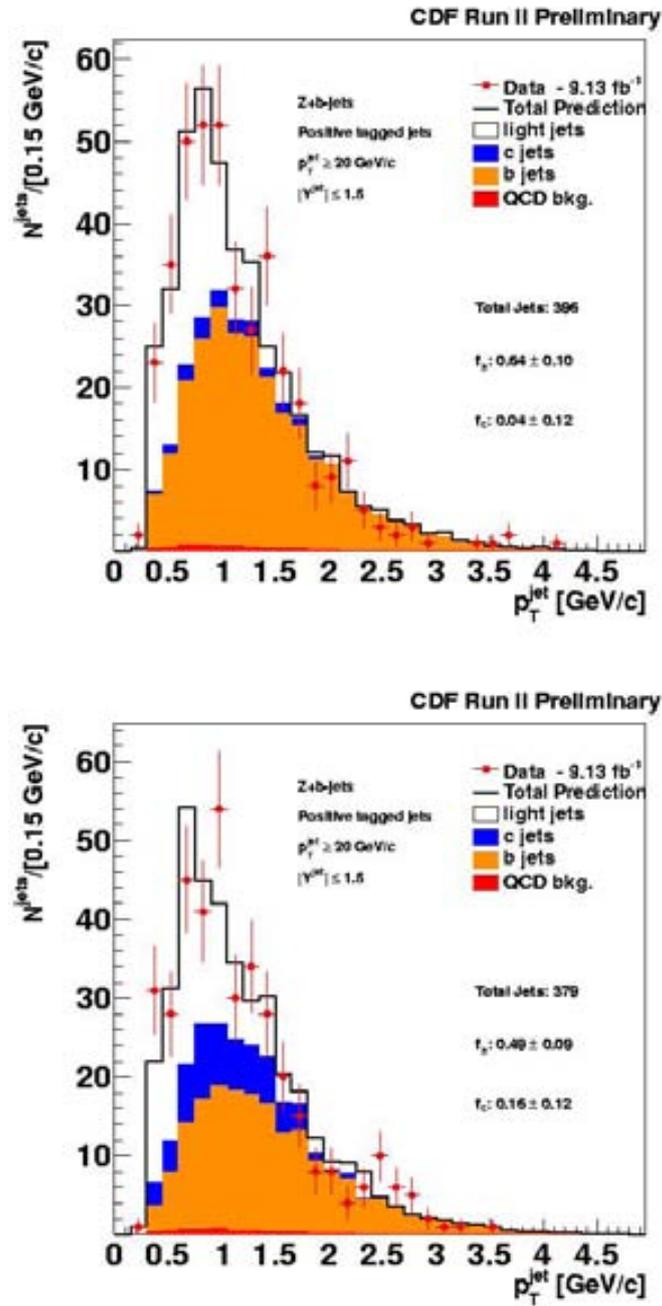
are presented in Figures 7.1, 7.2, 7.3 and for different rapidity bin Figure 7.4, 7.5 and are summarized in Table 7.2.

Fit result	
p_T bins	
20 - 27 GeV/c	252 ± 41
27 - 35 GeV/c	184 ± 32
35 - 45 GeV/c	198 ± 35
45 - 60 GeV/c	194 ± 22
60 - 100 GeV/c	129 ± 37
$ Y^{jet} $ bins	
0.0 - 0.3	238 ± 38
0.3 - 0.6	250 ± 38
0.6 - 1.0	296 ± 43
1.0 - 1.5	188 ± 28

Table 7.2: Number of fitted b tagged jets.

7.3 Unfolding

Once the fit is performed and background jets subtracted, the results are unfolded back to hadron level. Correction factors are estimated with Monte Carlo with the Alpgen+Pythia

Figure 7.1: Secondary vertex mass fit results for first two p_T bins.

7 $Z/\gamma^* + b$ -jet Differential Cross Sections

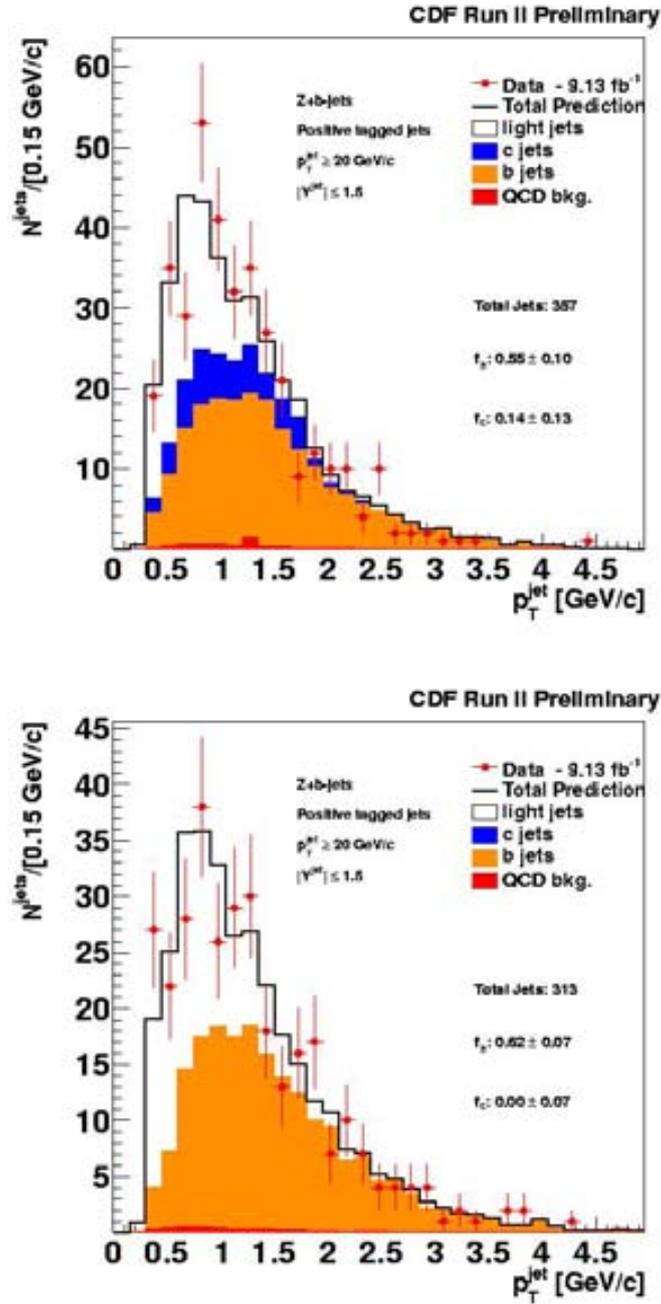


Figure 7.2: Secondary vertex mass fit results for third and fourth p_T bins.

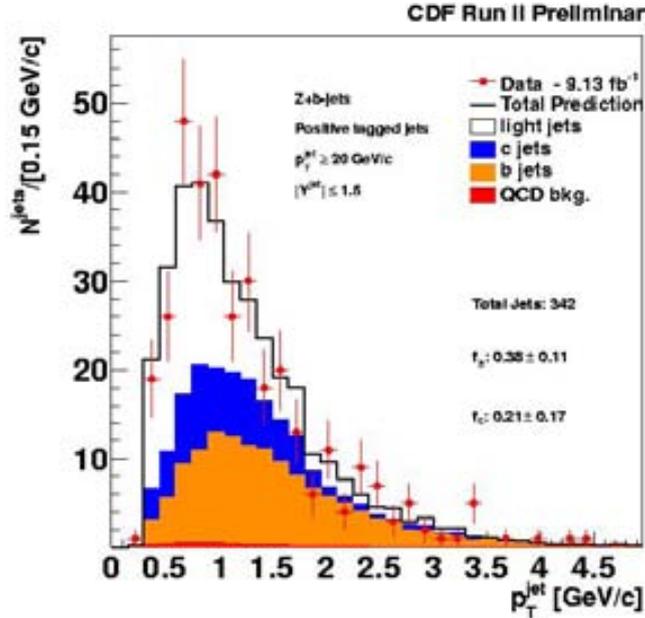


Figure 7.3: Secondary vertex mass fit results for the fifth p_T bins.

samples and defined as:

$$U_{had}^{cal}(\alpha) = \frac{N [Z/\gamma^* [particle\ level] + b - jet [particle\ level]]}{N [Z/\gamma^* [reconstructed] + b - tagged - jet [calorimeter]]} |_{MC} \quad (7.2)$$

Distributions are obtained at detector and hadron level and the ratio performed bin-by-bin. The Z/γ^* selection kinematic cuts are the same for both levels and are described in Chapter 5. Jets are matched to b hadrons and divided in p_T ($|Y|$) bins, following the threshold in Table 7.1. The same jet p_T (Y) cuts are applied to detector and hadron level. Therefore, unfolding factor corrects for both the small migration between bins due to experimental resolution and mainly for detector acceptance.

Figure 7.6 shows the inverse of the unfolding factors as function of jet p_T and jet rapidity.

7.4 Systematic Uncertainties

As explained previously the main systematic uncertainty is due to M_{SecVtx} template shape mismodeling in MC. Also here the PE technique is used to evaluate these effects. Pseudodata is built using templates modified to include each systematic effect and the fraction of b -jets is obtained by fitting it with the nominal templates. The uncertainty is estimated by the shift with respect to the mean obtained using pseudodata simulated with nominal templates. These measurements greatly benefit by the development of the ANNs to identify leptons, since

7 $Z/\gamma^* + b$ -jet Differential Cross Sections

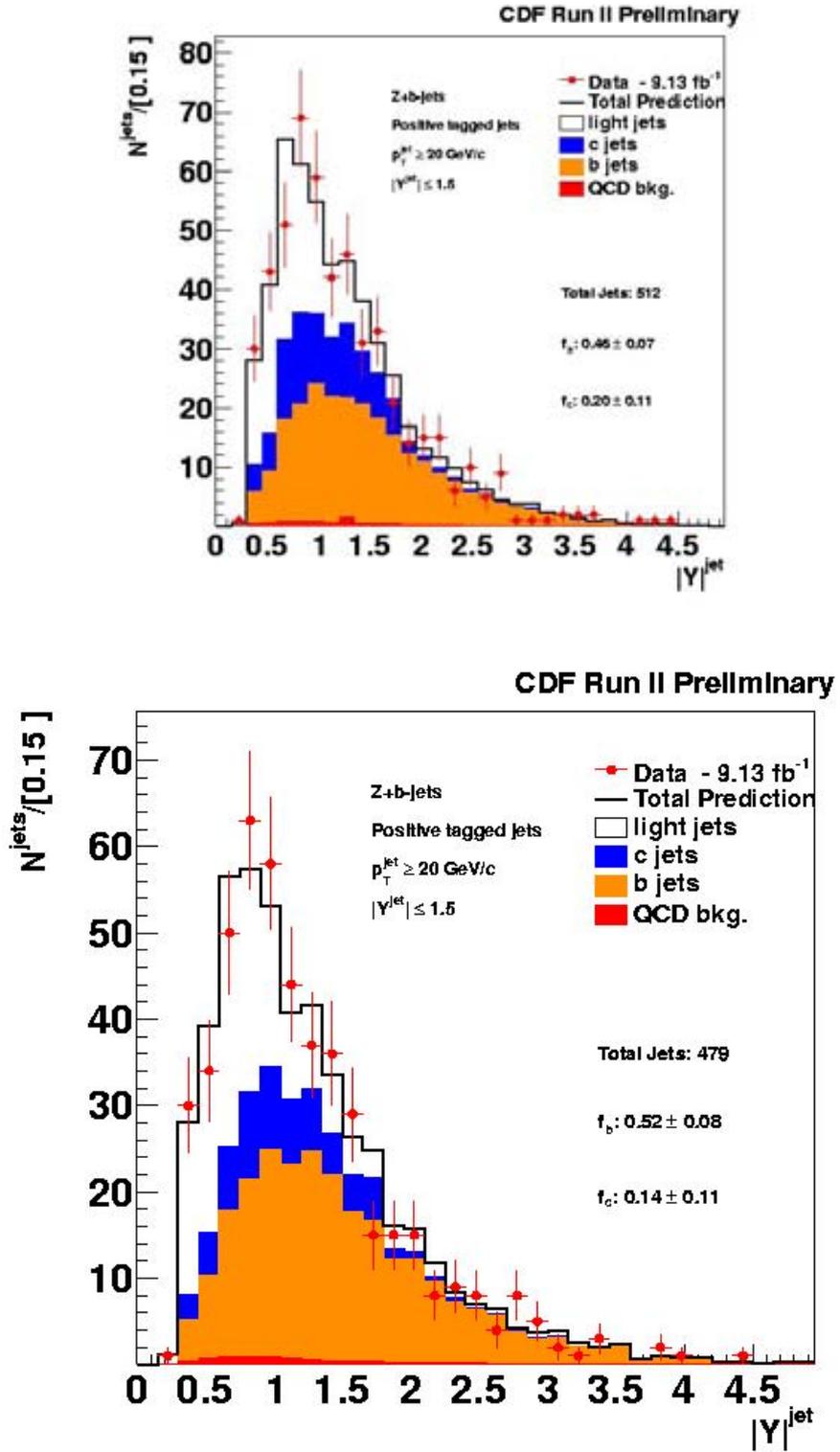


Figure 7.4: Secondary vertex mass fit results for first and second jet rapidity bins.

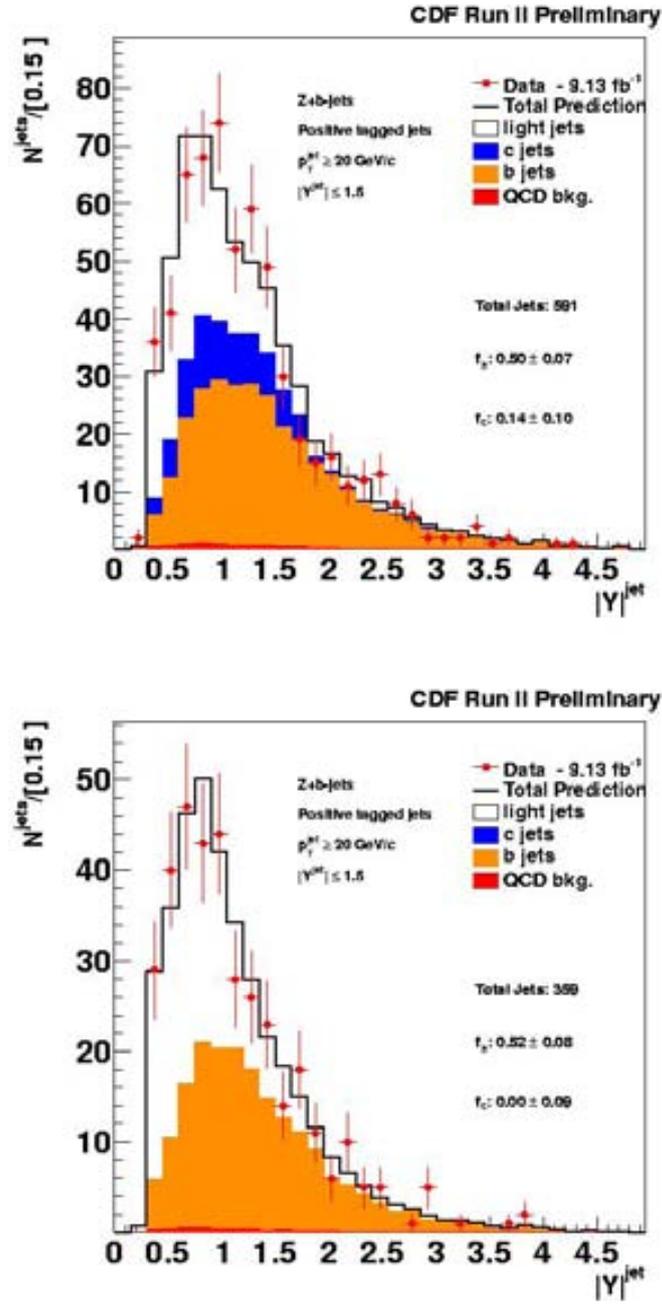


Figure 7.5: Secondary vertex mass fit results for third and fourth jet rapidity bins.

7 $Z/\gamma^* + b$ -jet Differential Cross Sections

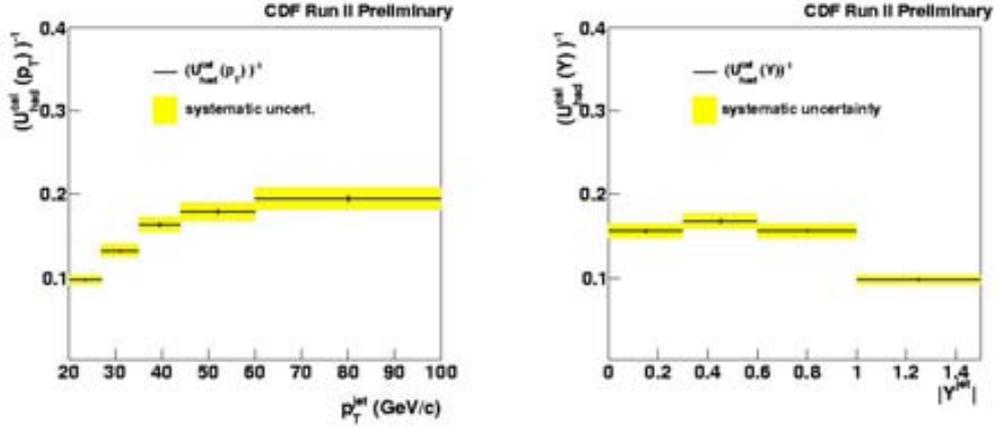


Figure 7.6: Inverse Unfolding factor as function of jet transverse momentum and jet rapidity.

the gain in statistics allows to perform a sensible fit per jet p_T and jet rapidity bin. In Figure 7.7 the fitted b jets fraction with the total systematics for jet p_T and rapidity, is presented.

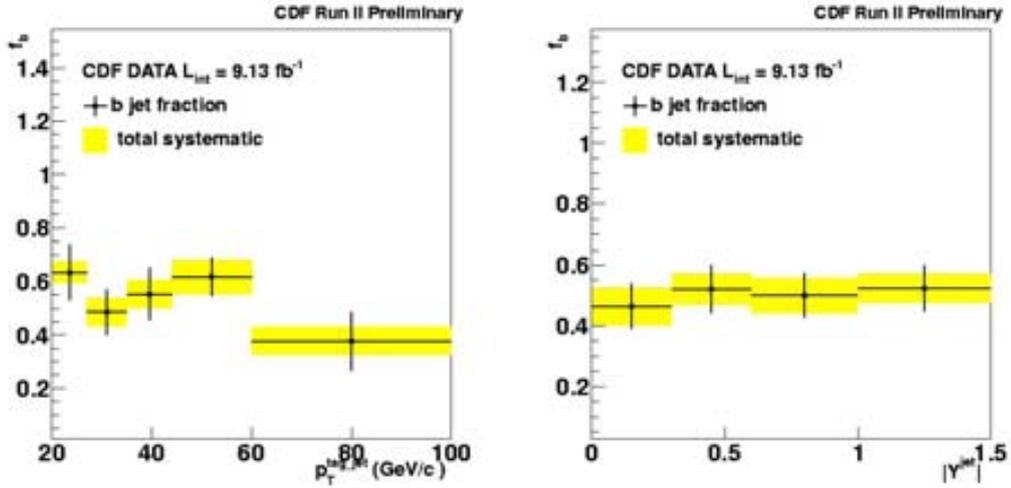


Figure 7.7: Secondary vertex mass fit results as function of jet transverse momentum and jet rapidity.

The systematics uncertainties due to jet energy scale and b -tagging uncertainty affect the unfolding factors. They are evaluated varying the uncertainty for each source of errors by $\pm 1\sigma$ and quoting half of the difference as systematic. Their contributions for each bin are summarized in Table 7.3.

7.4 Systematic Uncertainties

	abs	mpi	eta	tag	tot
20 - 27 GeV/c	2.0	0.2	1.4	5.2	5.8
27 - 35 GeV/c	1.2	1.1	1.5	5.2	5.6
35 - 45 GeV/c	1.8	2.4	0.6	5.2	6.1
45 - 60 GeV/c	2.5	1.8	0.7	5.2	6.1
60 - 100 GeV/c	3.4	1.7	1.5	5.2	6.7
0.0 - 0.3	0.9	1.2	0.5	5.2	5.4
0.3 - 0.6	1.0	1.5	0.3	5.2	5.6
0.6 - 1.0	1.1	1.5	0.7	5.2	5.7
1.0 - 1.5	1.2	1.5	1.0	5.2	5.7

Table 7.3: Systematic uncertainties (%) for each component of the JES and b -tagging efficiency as function of jet p_T and jet rapidity.

8 Results

In this chapter the integrated and differential cross section for the $Z/\gamma^* + b$ -jet ratio with respect to the inclusive Z/γ^* and $Z/\gamma^* + \text{jets}$ production cross section are presented. The measurements are compared to next-to-leading order predictions corrected for non perturbative QCD effects and to leading-order ME+PS Monte Carlo.

8.1 Theory predictions

8.1.1 pQCD calculation

The NLO pQCD predictions are performed with the MCFM program according to the calculus explained in Chapter 3, i.e. using the prediction for Z and one b -jet in the 5FNS scheme. As seen the main contribution is coming from $gb \rightarrow Zb$. Other processes, that contribute at the same order in α_s , are those having Zbg and $Zb\bar{b}$ in the final states. $Zb\bar{b}$ can be evaluated by MCFM only at leading order and therefore is more sensitive to scale variation. b quark is treated as massless except for $Zb\bar{b}$ processes where the quark mass is required to rend the calculation finite.

The pQCD theoretical calculation is performed setting the PDF to MSTW.2008 NLO and using several factorization and renormalization scales such as $\mu_0 = \mu_F = \mu_R = \sqrt{M_Z^2 + p_T^Z}$, $\mu_0 = \mu_F = \mu_R = \frac{1}{2}\hat{H}_T$ ¹ or $\mu_0 = \mu_F = \mu_R = \sqrt{\sum_i p_{T,i}^{jet}}$. Jets are reconstructed with the MidPoint algorithm with a cone size of $R = 0.7$ and $R_{sep} = 1.3$. The calculation is performed in the same phase space as for the measurement, i.e. Z/γ^* mass range of [66, 116] GeV/c² with a central ($|Y^{jet}| \leq 1.5$) high- p_T ($p_T \geq 20$ GeV/c) jet.

8.1.2 Non pQCD corrections

Measurements are presented at hadron level to remove detector effects, thus to compare with theory, pQCD predictions have to be corrected for non perturbative effects such as underlying event (UE) and the fragmentation of the partons into hadrons (hadronization).

The underlying event is expected to add extra energy inside the cone of the jet due to the soft interaction between the $p\bar{p}$ remnants, while the fragmentation could cause that some particles originated from the same parton end up out of the jet cone.

The correction factors are obtained from MC and applied bin-by-bin for each distribution. Alpgen+Pythia samples with MSTW2008 NLO PDFs and Tune Perugia 2011 is used to described the corrections as it has been shown to describe well jet shapes[38] and jet distributions in $Z + \text{jets}$ data [103]. They are estimated by comparing the hadron level cross section

¹ $\hat{H}_T = \sum_i^n p_T^j + p_T^{l^+} + p_T^{l^-}$, $j = \text{partons}$

with UE on and the parton level cross section with UE off:

$$C_{HAD}(\alpha) = \frac{\frac{d\sigma_{HAD}^{UE}}{d\alpha}}{\frac{d\sigma_{parton}^{noUE}}{d\alpha}}, \quad \alpha = p_T^{jet}, Y^{jet} \quad (8.1)$$

Figures 8.1.3 and 8.2 show the C_{HAD} factors as function of jet p_T and rapidity. Though they do not factorize, corrections due only to UE or to the hadronization are also presented for comparison. As expected, the underlying event correction represents a positive factor (extra energy added to jet cones), while the hadronization is giving a negative one. The effects are greater at low jet p_T since the more energetic is a jet, the more collimated is and so less sensitive to soft radiation. The corrections are almost flat as function of jet rapidity.

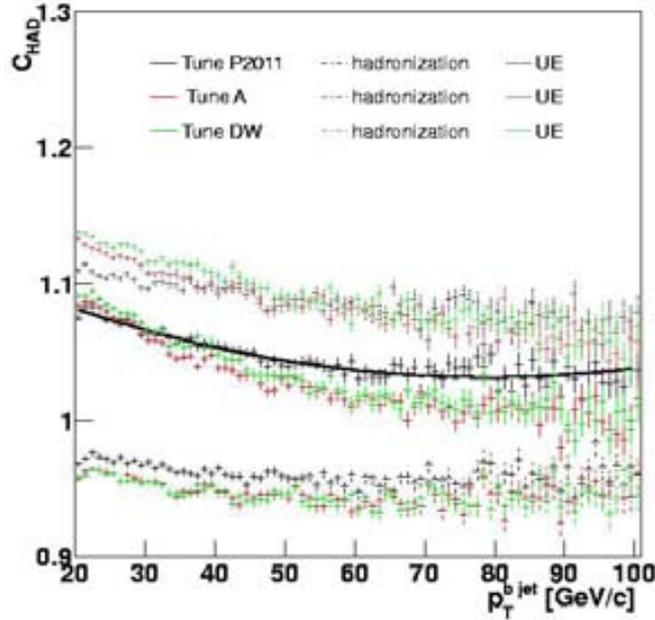


Figure 8.1: Non pQCD correction as function of jet p_T evaluated with Alpgen+Pythia MC samples. Contributions from underlying event and hadronization are also shown. Results using different Pythia tunes are presented.

8.1.3 Theoretical Uncertainties

Theoretical uncertainties include the renormalization and factorization scale variation and PDF errors.

The scale uncertainties are obtained by varying up and down the nominal renormalization and factorization scale (μ_0) by a factor of 2, so evaluating the prediction at $\mu_0/2$ and at $2\mu_0$. This represents the main uncertainty, that is approximately 20 %, and it is mainly due to

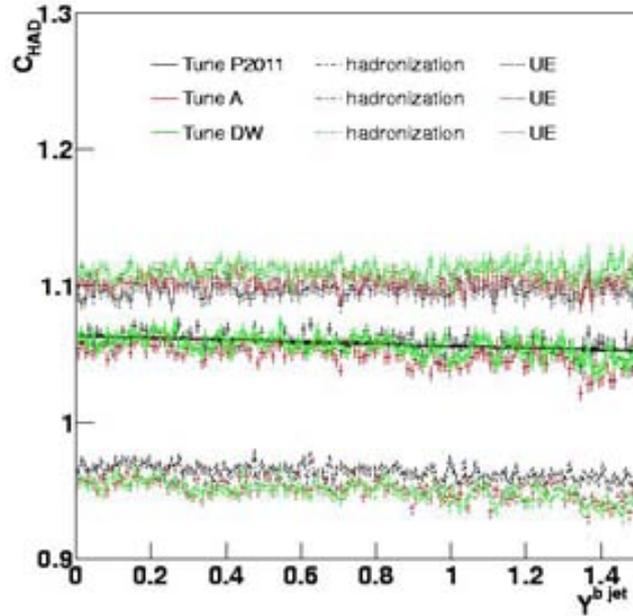


Figure 8.2: Non pQCD correction in function of jet rapidity evaluated with ALPGEN+PYTHIA. Here is possible to see the two contributions coming from underlying events and from hadronization. Different tunes for PYTHIA are shown.

the LO nature of the calculation of the $Zb\bar{b}$ term.

The PDF uncertainty is of the order of 2 % and it is evaluated using the Hessian method explained in Chapter 2 through up and down variations along the 20 eigenvectors of MSTW2008. Asymmetric uncertainties are obtained by summing in quadrature the maximal deviations in each direction associate to each of the 20 eigenvectors. The dependence of the prediction on the PDF set (MSTW2008, CTEQ6.6, NNPDF) was studied and found to be negligible. Other uncertainties could come from the modelling used for non pQCD corrections. Since UE is the main component, corrections were obtained using different Pythia Tunes such as Tune A and Tune DW. Differences are found to be small and between 2-3% (Figure .

8.2 Comparison with theoretical predictions

Results on the ratios of the b -jet cross section with respect to the inclusive Z/γ^* and $Z/\gamma^*+\text{jets}$ cross sections are presented. These measurements are defined for events with a Z/γ^* boson (in the mass range $66 \leq M_{l+l-} \leq 116 \text{ GeV}/c^2$) and jets of $p_T \geq 20 \text{ GeV}/c$ and $|Y| \leq 1.5$.

	NLO $Q^2 = m_{Z/\gamma^*}^2 + p_{T,Z}^2$	NLO $Q^2 = \langle p_{T,jet}^2 \rangle$
$\frac{\sigma(Z/\gamma^*+b)}{\sigma(Z)}$	2.3×10^{-3}	2.9×10^{-3}
$\frac{\sigma(Z/\gamma^*+b)}{\sigma(Z+jet)}$	1.8×10^{-2}	2.2×10^{-2}

Table 8.1: NLO MCFM theoretical predictions corrected for non pQCD effects.

8.2.1 Integrated $Z/\gamma^* + b$ -jet production cross section

The ratio of the integrated $Z/\gamma^* + b$ -jet cross section with respect to the Z/γ^* inclusive cross section was found to be:

$$\frac{\sigma_{Z/\gamma^*+b-jet}}{\sigma_{Z/\gamma^*}} = 0.256 \pm 0.020(stat) \pm 0.029(syst)\%$$

and with respect to the Z +jets inclusive cross section :

$$\frac{\sigma_{Z/\gamma^*+b-jet}}{\sigma_{Z/\gamma^*+jet}} = 2.05 \pm 0.16(stat) \pm 0.27(syst)\%$$

Theory predictions calculated using MCFM and corrected by non perturbative effects are shown in Table 8.1.

Predictions are found in agreement with data. However there is a large dependence on the scale suggesting high-order terms are needed.

8.2.2 Differential cross section

Differential cross section as function of jet p_T and rapidity are normalized to the measured Z/γ^* inclusive cross section and shown in Figures 8.3 and 8.4 where the data are compared with NLO pQCD corrected by non perturbative effects. Theory predictions are obtained for different renormalization/factorization scales. A general good agreement is obtained between data and theory. Experimental uncertainties are comparable to those of the theory.

8.3 Comparison with LO ME+PS predictions

We have also performed the calculation with Alpgen v2.10+Pythia 6.325 tune BW with CTEQQ5L PDF at $\mu_F = \mu_R = \sqrt{M_Z^2 + p_{T,Z}^2}$. In order to evaluate the MC prediction at hadron level, b -jets are defined as jets that match within $\Delta R \leq 0.4$ an outgoing b hadron. The b -jet cross section is therefore defined as:

$$\sigma_{Z/\gamma^*+b-jet}^{Alpgen+Pythia} = \frac{\sigma_{event}}{N_{gen}} \cdot N^{bhad}$$

where:

8.3 Comparison with LO ME+PS predictions

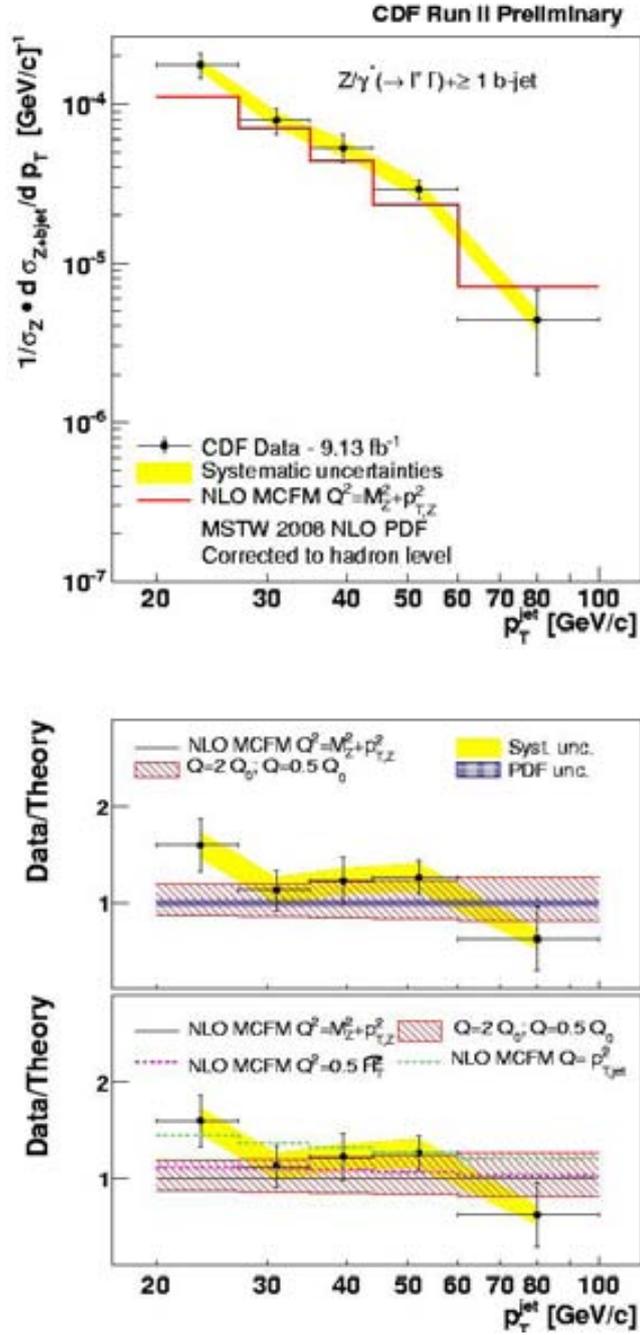


Figure 8.3: Differential cross section as function of jet p_T . The prediction is obtained with NLO MCFM at $Q^2 = M_Z^2 + p_{T,Z}^2$. It has also been studied the dependence on different PDF sets but the effect is negligible compared with the uncertainties.

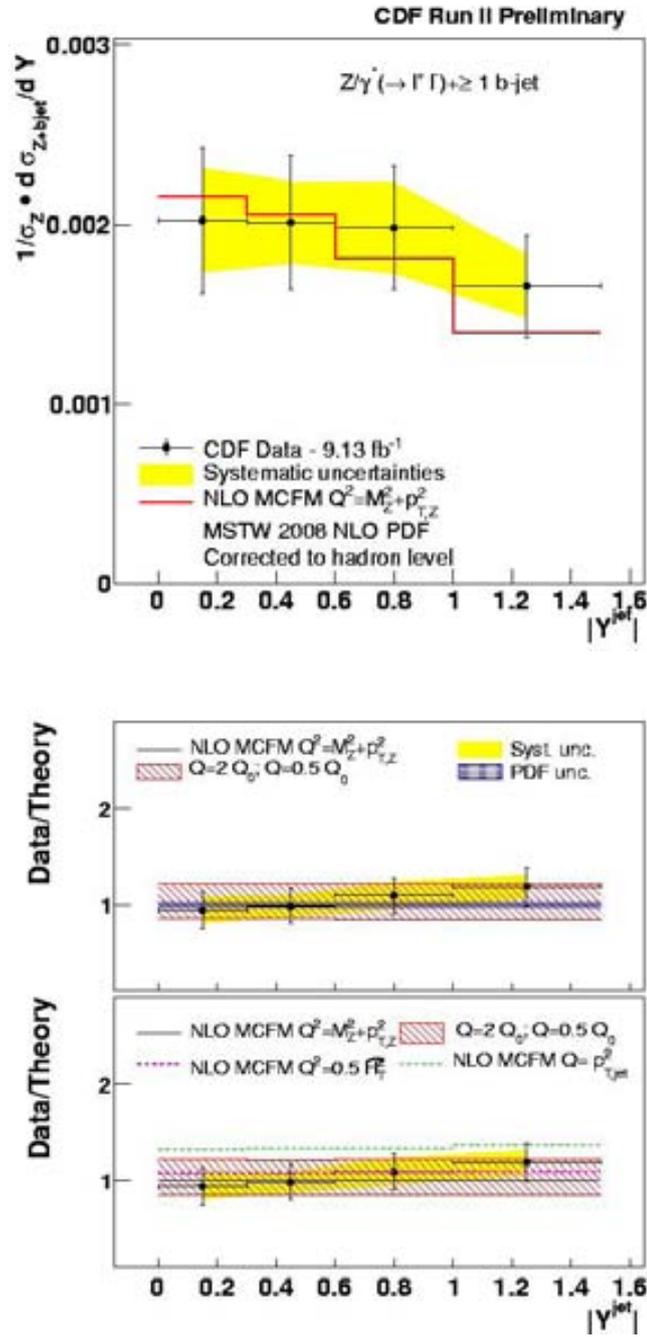


Figure 8.4: Differential cross section as function of jet rapidity . The prediction is calculated with NLO MCFM at $Q^2 = M_Z^2 + p_{T,Z}^2$. It is also studied the sensibility at different PDF sets but the uncertainty is too big to appreciate this effect.

8.3 Comparison with LO ME+PS predictions

- σ_{event} is the generator cross section times leptonic Z/γ^* branching ratio in Alpgen for a given sample
- N^{gen} is the number of generated events that have passed $|z - vertex| < 60 \text{ cm}$ requirements
- N^{bhad} is the number of b -matched hadronic jets with $p_T \geq 20 \text{ GeV}/c$ and $|Y^{jet}| \leq 1.5$.

The prediction for each process in Alpgen+Pythia MC is reported in Table 8.2.

Process	Predictions (pb)	
	$Z/\gamma^* \rightarrow e^+e^-$	$Z/\gamma^* \rightarrow \mu^+\mu^-$
z+bb+0p	0.1565	0.1566
z+bb+1p	0.0499	0.0500
z+bb+2p	0.0145	0.0145
z+0p	0.0043	0.0047
z+1p	0.0453	0.0437
z+2p	0.0158	0.0165
z+3p	0.0043	0.0044
z+4p	0.0014	0.0014
z+cc+0p	0.0001	0.0001
z+cc+1p	0.0016	0.0016
z+cc+2p	0.0015	0.0015

Table 8.2: Alpgen+Pythia prediction for $Z/\gamma^* \rightarrow \mu^+\mu^-$ and for $Z/\gamma^* \rightarrow e^+e^-$

Summing over all processes we obtained the overall prediction that is:

$$\sigma = \sum_i^{process} \frac{\sigma_i \cdot N_i^{bhad}}{N_i^{gen}}$$

$$\sigma_{Z/\gamma^* \rightarrow \mu^+\mu^- + b-jet} = 0.294 \text{ pb} \quad \sigma_{Z/\gamma^* \rightarrow e^+e^- + b-jet} = 0.295 \text{ pb}$$

As expected, the predictions from electron and muon channels are in agreement. In the same way the Z/γ^* +jet cross section is calculated; thus is possible to estimate the prediction for $Z/\gamma^* + b$ -jet cross section ratio respect to σ_{Z/γ^*+jet} :

$$R_{electron}^{Alpgen+Pythia} = 0.0143 \quad R_{muon}^{Alpgen+Pythia} = 0.0142$$

The measured value, as shown in previous section is

$$R = 0.0205 \pm 0.0016(stat) \pm 0.0027(syst)$$

8 Results

This result is quite different to that from Alpgen+Pythia, larger by a factor of 1.45, though it is in agreement with MCFM.

This is an important comparison as can be used as input for background estimation for analysis such as the search for the Higgs produced in the ZH channel.

9 Conclusions

Processes involving bottom quarks have a key role in hadron colliders. Being among the heaviest quarks they are expected to interact strongly with the electroweak symmetry breaking sector in the Standard Model (SM) and in many models beyond the SM. For example, a light Higgs boson decaying into a pair of bottom and antibottom quarks constitutes one of the main channels for the search of associated production (WH/ZH) at the Tevatron. Recent results have excluded at 95% of C.L. the high mass range for the SM Higgs. The low mass region is preferred and the ZH channel is one of the principal contributor in this case. Therefore the understanding of the $Z + b$ -jet process is crucial. Accurate theoretical predictions are needed and the measurements described in this thesis provide a testing ground for Standard Model perturbative QCD predictions and the simulation tools developed for the description of this process.

In this thesis I present a new measurement of the production cross section for b jets in events with a Z/γ^* boson decaying into a pair of electrons or muons. The data corresponds to the complete dataset collected at CDF II from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. The per jet cross section measurement is done for a phase space characterized by a Z/γ^* within $66 \leq M_{ll} \leq 116$ GeV/c² ($l = e, \mu$) and high- p_T central jets ($p_T \geq 20$ GeV/c $|Y| \leq 1.5$). Jets are reconstructed using the MidPoint algorithm in a cone size of $R=0.7$. The measurement is defined at hadron level and compared to a LO event generator matched to parton showers (ME+PS) and to NLO pQCD predictions computed with the MCFM program and corrected for non perturbative effects such as underlying event and hadronization. The comparison is performed for different values of renormalization and factorization scales and using several PDF sets.

The ratio of the integrated $Z/\gamma^* + b$ -jet cross section with respect to the Z/γ^* inclusive cross section is measured to be:

$$\frac{\sigma_{Z/\gamma^*+b-jet}}{\sigma_{Z/\gamma^*}} = 0.256 \pm 0.020(stat) \pm 0.029(syst)\%$$

and with respect to the Z/γ^*+jets inclusive cross section :

$$\frac{\sigma_{Z/\gamma^*+b-jet}}{\sigma_{Z/\gamma^*+jet}} = 2.05 \pm 0.16(stat) \pm 0.27(syst)\%$$

The latter is found to be a factor 1.4 larger than the LO ME+PS prediction but it is in agreement with MCFM. In fact, MCFM predictions are consistent with differential and integrated cross section measurements, though large variations are seen for different scales. New theoretical developments with improved accuracy, merging NLO with parton showers, have been recently released which can be tested against the measurements presented in this thesis.

List of Figures

2.1	Asymptotic freedom and confinement	4
2.2	Color Flow in a QCD vertex	4
2.3	Lattice QCD scheme	5
2.4	Feynman diagram at LO and NLO for Z+jet process.	6
2.5	Running α_s coupling constant	8
2.6	Typical hadron collision illustration	9
2.7	Deep Inelastic Scattering scheme and HERA experimental layout.	10
2.8	Experimental structure function by ZEUS	11
2.9	Scale rule break scheme explication	12
2.10	Initial State radiation scheme	12
2.11	Gluon PDFs	14
2.12	Kinematic regions and data sets typically used in PDF fits.	14
2.13	PDF gluon distribution for MSTW, CTEQ, NNPDF	15
2.14	Hessian Formalism in a geometric point of view	16
2.15	Uncertainties of gluon and u quark PDF	17
2.16	$e^+e^- \rightarrow Z/\gamma^* \rightarrow q\bar{q}$ LO diagram	17
2.17	Color flow in a DIS process	19
2.18	Invariant mass distribution of cluster of color singlet quarks after non perturbative gluon splitting.	20
2.19	QCD potential and string model scheme.	21
2.20	Lund Model	22
2.21	Illustration of string and cluster fragmentation.	22
2.22	$Z + Np$ contributions	23
2.23	Feynman diagrams at LO and NLO for Z+jets process.	24
2.24	$2 \rightarrow n$ factorization PS scheme.	26
2.25	Shower cascade	27
2.26	Example of double counting for MEPS MC.	28
2.27	MLM mechanism.	29
2.28	Jets at different orders of perturbation theory and at a different points in the analysis.	31
2.29	Collinear unsafety	31
2.30	Infrared unsafety	31
2.31	MidPoint Infrared unsafety example	33
2.32	Computation needs for jet algorithms	33
2.33	Jet algorithms Shapes	34
3.2	LO diagrams for $qq \rightarrow Zbb$	37

List of Figures

3.3	Diagrams contributing to the associated production of a Z boson and two high- p_T jets	38
3.4	Real NLO corrections for $gb \rightarrow Zbg$	38
3.5	Diagrams of some real corrections for the two processes.	39
4.1	Illustration of the complete accelerator chain at Fermilab	41
4.2	Antiproton production.	43
4.3	Integrated luminosity as a function of time for Run II data-taking.	45
4.4	CDF II detector sketch in three dimensions.	45
4.5	Particles identification.	46
4.6	CDFII Tracking system.	47
4.7	Silicon tracking sub-detectors	48
4.8	Schematic illustration of the three instrumented mechanical barrels of SVXII detector	49
4.9	Intermediate Silicon Layer cartoon	50
4.10	Scheme of COT	53
4.11	Track Reconstruction coordinates.	54
4.12	View of the Time of Flight	55
4.13	Schematic illustration of an azimuthal sector of the central electromagnetic calorimeter.	56
4.14	Muon sub-detectors	59
4.15	CDFII trigger system	60
4.16	CDFII trigger block diagram.	61
5.1	Scheme of a single calorimeter tower.	70
5.2	Dijet p_T balance as a function of η_{det} before the corrections	71
5.3	Dijet p_T balance as a function of η_{det} after the corrections	72
5.4	$\langle E_T \rangle$ versus the number of vertexes for jets of cone size $R = 0.7$	73
5.5	p_T difference between particle and calorimeter jet for different p_T range.	74
5.6	Absolute corrections for different cone sizes as a function of calorimeter jet p_T	75
5.7	b -tag scheme.	76
5.8	Efficiency for SecVtx tagger algorithm.	77
5.9	Mistag for SecVtx algorithm as function of jet p_T and jet rapidity	78
6.1	Comparison between probe and fake muons for some kinematical distributions.	84
6.2	Muon ϵ_{ID} and background reduction rate as function of Same Charge ANN output. The cut on this ANN output in the analysis is 0.875.	85
6.3	Muon ϵ_{ID} and background reduction rate as function of Decay In Flight ANN output. The cut applied in the analysis on this ANN output is 0.725.	86
6.4	SC and DIF ANNs output.	86
6.5	$Z/\gamma^* \rightarrow \mu^+\mu^-$ inclusive significance as a function of the output of the SC and DIF ANNs.	87
6.6	Comparison between fake and probe central electron distributions for the variables used in the central ANN.	90
6.7	Comparison between fake and probe plug electron distributions for the variables used in the plug ANN.	91
6.8	Output of the central and plug ANNs	92

6.9	Electron ϵ_{ID} and background reduction rate as a function of central ANN . . .	92
6.10	Electron ϵ_{ID} and background reduction rate as a function of plug ANN . . .	93
6.11	Scale factors for central and plug electron fakes	95
6.12	Fake rate for central and plug electrons.	96
6.13	Data-MC comparison for M_{Z/γ^*} , p_T and rapidity of jets for electron and muon channel in the pretag sample.	97
6.14	Secondary vertex mass templates for Tight SECVTX tagged b, c, LF jets. . .	98
6.15	Secondary vertex mass templates for Tight SECVTX tagged b, c, LF jets. Plots show the contribution from each process and its corresponding weight. .	99
6.16	Secondary vertex mass templates for Tight SECVTX tagged b, c, LF jets for different processes, normalized to 1.	100
6.17	Comparison between electron and muon channel SecVtx templates for the b, c and LF jet flavors	101
6.18	b SecVtx Mass template comparison between MC signal and background . . .	102
6.19	SecVtx Mass distribution for $Z/\gamma^*x \rightarrow l^+l^-$	103
6.20	Pseudo-Experiments study to check possibly bias	104
6.21	Pseudo-Experiments with several input scenarios	105
6.22	Track reconstruction inefficiency SecVtx Mass templates.	108
6.23	Comparison between SecVtx Mass templates for single/double b/c hadron in tagged jets.	109
6.24	Comparison of SecVtx Mass templates constructed using LF MC jets and Negative tags from data	110
6.25	Example of the template systematic uncertainty evaluation	111
7.1	Secondary vertex mass fit results for first two p_T bins.	115
7.2	Secondary vertex mass fit results for third and fourth p_T bins.	116
7.3	Secondary vertex mass fit results for the fifth p_T bins.	117
7.4	Secondary vertex mass fit results for first and second jet rapidity bins.	118
7.5	Secondary vertex mass fit results for third and fourth jet rapidity bins.	119
7.6	Inverse Unfolding factor as function of jet transverse momentum and jet rapidity.	120
7.7	Secondary vertex mass fit results as function of jet transverse momentum and jet rapidity.	120
8.1	Non pQCD correction as function of jet p_T evaluated with Alpgen+Pythia MC samples.	124
8.2	Non pQCD correction in function of jet rapidity evaluated with ALPGEN+PYTHIA.	125
8.3	Differential cross section as function of jet p_T	127
8.4	Differential cross section as function of jet rapidity	128

List of Tables

2.1	Comparison between the main important PDF distribution	15
4.1	Summary of the main Tevatron characteristics.	42
4.2	Characteristics of SVX II layers	50
4.3	Summary of the technical details of the tracker sub-detectors.	52
4.4	<i>Track resolutions using COT only or Silicon information plus COT.</i>	54
4.5	The CDF II calorimeters with their acronym, η region, segmentation, thickness and energy resolution. X_0 represents the shower length and λ_I is the pion nuclear absorption length in $g\text{ cm}^{-2}$	56
5.1	Track quality cuts for the primary vertex	66
5.2	SecVtx Tight parameters for Pass 1 and for Pass 2.	77
6.1	Observed events for the electron and muon channels at different stages of the event selection.	94
6.2	MC samples used to estimated the expected background in the analysis. . . .	94
6.3	The background contributions for $Z/\gamma^* + b$ tagged jets in muon and electron channel.	105
6.4	Events observed in data and expected backgrounds for all processes in the electron and muon channels	105
6.5	Acceptance table for electron and muon channel for different processes. . . .	107
6.6	Systematic uncertainties for the acceptances due to jet energy scale (JES) and tag efficiency (tag)	112
6.7	Summary of the systematics that affects the cross section measurement. . . .	112
7.1	Number of observed tagged jets and expected background contributions. The muon fake error since this background is negligible is quoted to be 100 %. For the electron fake the uncertainty is coming from the fit of fake rate.	114
7.2	Number of fitted b tagged jets.	114
7.3	Systematic uncertainties (%) for each component of the JES and b -tagging efficiency as function of jet p_T and jet rapidity.	121
8.1	NLO MCFM theoretical predictions corrected for non pQCD effects.	126
8.2	Alpgen+Pythia prediction for $Z/\gamma^* \rightarrow \mu^+\mu^-$ and for $Z/\gamma^* \rightarrow e^+e^-$	129

Bibliography

- [1] P. Higgs, *Broken Symmetries, Massless Particles and Gauge Fields*, [Phys. Lett. 12](#), (1964) 132;
F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, [Phys. Rev. Lett. 13](#) (1964) 321;
- [2] Atlas collaboration, *An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to 4.9 fb¹ of pp collision data at $s = 7$ TeV*, [ATLAS-CONF-2012-019](#), (2011);
CMS collaboration, *Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV*, CMS-HIG-11-032, CERN-PH-EP-2012-023, (2011), [arXiv:1202.1488](#) [hep-ex];
- [3] M. Drees, R. Godbole and P. Roy, *Theory and Phenomenology of Sparticles*, World Scientific, (2004);
- [4] D. Acosta (CDF collaboration), *Measurement of the b jet cross section in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, [Phys Lett B 706, 295-313](#), (2008);
- [5] V. M. Abazov et al. (D0 coll.), *Measurement of the ratio of inclusive cross sections $\sigma(p\bar{p} \rightarrow Z + b\text{-quark jet})/\sigma(p\bar{p} \rightarrow Z + \text{jet})$ at $\sqrt{s} = 1.96$ TeV* [Phys Rev D 83, 031105](#), (2011)
- [6] Atlas Collaboration, *Measurement of the cross-section for b-jets produced in association with a Z boson at $\sqrt{s} = 7$ TeV with the ATLAS detector* [Phys Lett. B 706, 295-313](#), (2012), [arXiv:1109.1403](#) [hep-ex];
- [7] CMS collaboration, *Measurement of the $Z/\gamma^* + bb\text{-jets}$ cross section in pp collisions at $\sqrt{s} = 7$ TeV with the CMS detector*, [CMS-PAS-SMP-12-003](#), (2012);
- [8] Bethke, *The 2009 World Average of α_s* , (2009), [arXiv:0908.1135](#) [hep-ph];
- [9] Necco and Sommer, *The $N_f=0$ heavy quark potential from short to intermediate distances*, [Nucl. Phys B622 328,346](#), (2002), [hep-lat/0108008v1](#);
- [10] R.K. Ellis, W. J. Stirling and B.R. Webber, *QCD and Collider Physics* Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, vol. 8, (1996);
G. Dissertori, I. G. Knowles and M. Schmelling, *High Energy Experiments and Theory*, Clarendon, Oxford, (2003);
- [11] G. Salam, *Elements of QCD for hadron colliders*, CERN Yellow Report CERN-2010-002, pp. 45-100, (2011), [arXiv:1011.5131](#) [hep-ph];

Bibliography

- [12] ML Mangano, *QCD and the physics of hadronic collisions*, [Phys-Usp 53 109.](#), (2010);
- [13] R. Brock et al. [CTEQ Collaboration], *Handbook of perturbative QCD*, [Rev. Mod. Phys. 67 157](#), (1995);
x Field R D, *Application of Perturbative QCD*, Redwood City, Calif. Addison-Wesley, (1989);
Dokshitzer Yu L et al., *Basics of Perturbative QCD*, Ed J Tran Thanh Van, (1991);
- [14] J. Smit, *Introduction to quantum field on a lattice, A robust mate*, Cambridge Lect. Notes Phys. 15 (2002) 1.;
Perspectives in Lattice QCD, Proceeding of the Workshop, Nara International Seminar House, Nara, Japan, 2005, Ed.Y. Kurumashi, World Scientific, Singapore, (2007);
- [15] D.J. Gross and F. Wilizec, *Asymptotically Free Gauge Theories*, [Phys. Rev. D 8, 3633](#) (1973);
- [16] Collins, Soper, Sterman, *Perturbative QCD*, ed A. H. Mueller, World Scientific Singapore, (1989);
- [17] R. Devenish A. Cooper-Sarkar, *Deep Inelastic Scattering*, Oxford University Press, (2004);
- [18] 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);
- [19] 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);
- [20] Pumplin et al, *A new generation of Parton distribution functions with uncertainties from global QCD analysis*, [JHEP 0207:012](#), (2002), [hep-ph/0201195v3](#);
- [21] G. Watt, *MSTW PDFs and impact of PDFs on cross sections at TeVatron and LHC*, [Nucl.Phys.Proc.Suppl. 222-224 61-80](#), (2012), [arXiv:1201.1295](#) [hep-ph];
- [22] NNPDF Coll., *A determination of parton distribution with faithful uncertainty estimation*, [Nucl. Phys B, 809, 1-63](#) (2009), [arXiv:0808.1231](#) [hep-ph];
- [23] J. Pumplin et al, *Uncertainties of Predictions from Parton Distributions Functions II. The Hessian Method*, [Phys. Rev D 65, 01413](#) (2001), [hep-ph/0101032v3](#);
- [24] Dokshitzer, Khoze, Troyan, *On the concept of Local Paton-Hadron Duality*, [J Phys G Nucl. Part Phys 17 1585](#) (1991);

- [25] Winter, Krauss, Soff, *A modified cluster-hadronization model*, *Eur. Phys J* **C36:391-395**, (2004), [hep-ph/0311085v1](#);
- [26] B Andersson, *The Lund Model*, Cambridge Monogr. Part. Phys, Nucl, Phys. Cosmol., (1997);
- [27] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, *J. High Energy Phys.* **0307** (2003) **001**, [hep-ph/0206293](#).
- [28] F. Caravaglios, M.L. Mangano, M. Moretti, R. Pittau, *A New approach to multijet calculations in hadron collisions*, *Nucl. Phys.* **B 539** (1999) 215, [hep-ph/9807570](#).
- [29] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *Eur. Phys. J. C* **53** (2008) 473, [arXiv:0706.2569](#) [hep-ph].
- [30] J. Campbell, R. K. Ellis, F. Maltoni, S. Willenbrock, *Associated Production of a Z Boson and a Single Heavy-Quark Jet* *Phys.Rev.D* **69:074021**, (2004), [hep-ph/0312024v2](#)
- [31] J.C. Collins, *The Sudakov Form factors*, Ad.Ser Direct High Energy Physics, I:537-614, (1989);
- [32] T Sjostrand, *Monte Carlo Generators*, CERN-CLGAPP-2006-66, (2006), [hep-ph/0611247v1](#);
- [33] T Sjostrand et al, *High energy Physics event Generation with Pythia 6.1*, Comput. Phys Commun. 135, 238 (2001) [hep-ph/0603175](#);
- [34] Rick Field, R. Craig Group, *Pythia Tune A, Herwig, and Jimmy in Run 2 at CDF*, [hep-ph/0510198](#).
- [35] R. Field, *CDF Run 2 Monte-Carlo Tunes*, [FERMILAB-PUB-06-408-E](#).
- [36] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, *Phys. Rev. D* **82**, **074018** (2010), [arXiv:1005.3457](#) [hep-ph].
- [37] T. Sjöstrand, P. Skands, *Transverse-Momentum-Ordered Showers and Interleaved Multiple Interactions*, *Eur. Phys. J. C* **39** (2005) 129, [hep-ph/0408302](#).
- [38] B.Cooper, J.Katzy, M.L.Mangano, A.Messina, L.Mijovic, P.Skands, *Monte Carlo tuning in the presence of Matching*, CERN-PH-TH/2011-228, DESY 11-124, [arXiv:1109.5295](#) [hep-ph].
- [39] Gavin P Salam, *Towards Jetography*, (2010) [arXiv:0906.1833v2](#) [hep-ph];
- [40] J.E. Hutcj et al, *Toward a Standardization of Jet Definitions*, in Snowmass Summer study, p 134, World Scientific (1992), FERMILAB-CONF-90-249-E;

Bibliography

- [41] Blazey,Dittman,Ellis et alt. *Run II Jet Physics*,(2010) [hep-ex/0005012v2](#));
J. Huston and M. Toennesmann, *On Building Better Cone Jet Algorithms*, eConf C010630 P513 (2001),[hep-ph/0111434](#);
- [42] TeV4LHCn QCD Working group *Tevatron for LHC Report of QCD working group* ,
Fermilab-Conf-06-359, 1 (2006), [hep-ph/061005](#);
- [43] Beate Heinemann, A. Mehta, *Theoretical Predictions on the Z+b-jet Production Cross Section and their Uncertainties*, CDF Note 9595, (2008);
- [44] J. Campbell, Ellis, Maltoni, Willenbrock, *Associated Production of a Z Boson and a Single Heavy-Quark Jet*, FERMILAB-Pub-03/378-T (2003), [hep-ph/0312024v2](#);
- [45] Maltoni, McElmurry, Willenbrock, *Inclusive Production of a Higgs or a Z boson in association with heavy Quarks*,[Phys.Rev.D72:074024](#), (2005), [hep-ph/0505014](#) ;
- [46] Aivazis, Collins, Olness, Tung, *Leptoproduction of heavy quarks. II. A unified QCD formulation of charged and neutral current processes from fixed-target to collider energies* [Phys. rev. D 50, 3102](#) (1994);
J. Collins, *Hard-scattering factorization with heavy quarks: A general treatment*, [Phys. Rev. D 58 094002](#) (1998)
- [47] Campbell, Ellis, Maltoni, Willenbrock, *Higgs-Boson Production in association with a single bottom quark*,[Phys.Rev.D67:095002](#), (2003), [hep-ph/0204093](#) ;
- [48] J. Campbell, Ellis, Maltoni, Willenbrock, *Production of a Z boson and two Jets with One Heavy-Quark Tag*, [Phys.Rev.D73:054007](#),(2006),[hep-ph/0510362](#);
- [49] J. Campbell and Ellis, *Radiative corrections to $Zb\bar{b}$ production*, [Phys.Rev.D62:114012](#), (2000), [hep-ph/0006304](#);
- [50] Cordero, Reina, Wackerroth, *NLO QCD corrections to $Zb\bar{b}$ production with massive bottom quarks at the Fermilab TeVatron*, [Phys.Rev.D78:074014](#), (2008), [arXiv:0806.0808v2](#) [hep-ph];
Cordero, Reina, Wackerroth, *Associated production of a W or a Z boson with bottom quarks at the TeVatron and the LHC*, PoS(RADCOR2009)055, (2010), [arXiv:001.3362v1](#) [hep-ph];
- [51] Frederix, Frixione, Hirschi, Maltoni, Pittau, Torrielli, *W and Z/γ^* boson production in association with a bottom-antibottom pair*, [JHEP 1109:061](#), (2011), [arXiv:1106.6019v2](#) [hep-ph];
- [52] Hirshi, Frederix, Frixione, Garzelli, Maltoni, Pittau, *Automation of one-loop QCD corrections*, CERN-PH-TH/2011-031 (2011), [arXiv:1103.0621v1](#) [hep-ph];
- [53] Campbell, *Overview of the theory of W/Z+ jets and heavy flavor*, Hadron Collider Physics Symposium (HCP 2008),[arXiv:0808.3517v1](#) [hep-ph] ;
- [54] D.P. McGinnis, *Fermilab Tevatron operational status*. In C. Horak, editor, proceedings of the Particle Accelerator conference (PAC 05), 2005. In addition, detailed and

updated information on the Tevatron is available in the following websites: <http://www-bd.fnal.gov/runII/index.html>, <http://www-bdnew.fnal.gov/operations/rookiebooks/rbooks.html>;

- [55] D. Mohl et al., *Physics and Techniques of Stochastic Cooling*, Phys. Rept., 58, (1980);
- [56] S. Nagaitsev et al., *Experimental Demonstration of Relativistic Electron Cooling*, Phys. Rev. Lett., 96, (2006);
- [57] R. Blair and others (CDF II Collaboration), *The CDF II Detector*, Technical Design Report, FERMILAB-Pub-96390-E CDF, (1996);
- [58] A. Sill, *CDF Run II silicon tracking projects.*, Nucl. Instrum. Methods, A447, (2000);
- [59] T. K. Nelson (for the CDF II Collaboration), *The CDF Layer 00 Detector*, FERMILAB-CONF-01/357-E; (2001), [CDF Note 5780](#);
- [60] C. S. Hill, *Initial experience with the CDF layer 00 silicon detector*, [Nucl. Instr. Methods Phys. Revs. sect A, 511, 118-120](#), (2004);
- [61] A. Affolder et al., *Status report of the intermediate silicon layers detector at CDF II* , [Nucl. Instrum, Methods, A485](#), (2002);
- [62] A. Affolder et al., Nucl. Instrum. Methods Phys, Rev Sect A, 461 (2001), 216-218;
- [63] A. Affolder et al., *CDF Central Outer Tracker*, [Nucl. Instrum, Methods, A526](#), (2004);
- [64] J. Marriner, *Secondary vertex fit with mass and pointing constraints (CTVMFT)*, CDF Internal Note, (1996) unpublished;
- [65] CDF Collaboration, *SVX II simulation study and upgrade proposal*, Public CDF Note 1922 (1992), [CDF Note 1922](#);
- [66] CDF Collaboration, *The CDF upgrade*, Public CDF Note 3171 (1995), [CDF Note 3171](#);
- [67] S. Menzemer, *TrackingCal - A tracking and alignment software package of the CDF II silicon detector* Master thesis, University of Karlsruhe, (2002), CDF Internal Note 5968, unpublished;
- [68] C. S. Hill et al., *Operational experience and performance of the CDF II silicon detector*, [Nucl. Instrum. Methods, A530](#), (2004);
- [69] D. Acosta et al. (CDF collaboration), *A time of flight detector in CDF II*, [Nucl. Instrum. Methods, A518](#), (2004);
- [70] L. Balka et al., *The CDF central electromagnetic calorimeter*, Nucl. Instr. Methods, A267, (1988);
- [71] R.S.Hahn et al., *Calibration system for the CDF central electromagnetic calorimeter*, Nucl. Instrum. Methods, A267, (1988);
- [72] M. Gallinaro, *A new Scintillator Tile/Fiber Preshower Detector for the CDF Central Calorimeter*, IEEE Trans. Nucl. Sci., 52, (2005);

Bibliography

- [73] S. Bertolucci et al., *The CDF central and endwall hadron calorimeter*, [Nucl. Instrum. Methods, A267](#),(1988);
- [74] M. Albrow et al., *The CDF plug upgrade electromagnetic calorimeter: test beam results.*, [Nucl. Instrum. Methods, A480](#), (2002);
- [75] M. Albrow et al., *A preshower detector for the CDF Plug Upgrade: test beam results*, [Nucl. Instrum. Methods, A412](#), (1998);
- [76] G. Apollinari et al, *Shower Maximum detector for the CDF plug upgrade calorimeter*,[Nucl. Instrum. Methods, A412](#), (1998);
- [77] G. Ascoli et al., *CDF central muon detector*, [Nucl Instrum. Methods, A268](#), (1988);
- [78] C. M. Ginsburg et al., *CDF un 2 Muon System*,[Eur. Phys. J., 33 \(S1002\)](#), (2004);
- [79] F. Abe et al. *Measurement of the antiproton-proton total cross section at $\sqrt{s} = 546$ and 1800 GeV*, [Phys. Rev. D 50, 5550](#) (1993).
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).
- [80] D. Acosta et al. (CDF collaboration), *The performance of the CDF luminosity monitor*, [Nucl. Instrum. Methods, A494](#), (2002);
- [81] E. J. Thomson et al, *Online Track Processor for the CDF Upgrade*, [IEEE Trans Nucl. Sci, 49](#), (2002);
- [82] B. Ashmanskas et al., *The CDF Silicon Vertex Trigger*, [Nucl. Instrum. Methods, A518](#), (2004);
- [83] M. Dell’Orso , *The CDF Silicon Vertex Trigger*, [Nucl. Phys. \(Proc. Suppl\), B156](#), (2006);
- [84] S. Belforte et al, *Silicon Vertex Tracker*, Technical Design Report, Public CDF Note 3108 (1995), [CDF Note 3108](#);
- [85] J. Adelman et al., *The Silicon Vertex Trigger upgrade at CDF*, [Nucl. Instr. Methods, Phys Res Sect A 572,361-364](#), (2007) ;
- [86] G. Gomez-Caballlos et al., *Event Builder and Level 3 at the CDF experiment*, [Nucl. Instrum. Methods, A518](#), (2004);
- [87] J. Antos et al., *Data proceeding model for the CDF experiment*,[IEEE T NUCL SCI , 53 \(5\) 2897 - 2906](#), (2006);
- [88] E. Gerchtein and M. Paulini, *CDF detector simulation framework and performance*, In eConf C0303241 TUMT005, (2003), [arXiv:0306031v1](#) [physics.comp-ph];
- [89] G.C. Blazey et al., *Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop*, [[hep-ex/0005012](#)].
S.D. Ellis, J. Huston, M. Toennesmann, *On Building Better Cone Jet Algorithms*, eConf C010630 (2001) P513, [[hep-ph/0111434](#)].

- [90] Bhatti et al, *Determination of the Jet Energy Scale at the collider Detector at Fermilab*, *Nucl.Instrum.Meth.A566:375-412*, (2006), [hep-ex/0510047v1](#);
- [91] CDF collaboration, *Measurement of the Cross Section for $t\bar{t}$ Production in $p\bar{p}$ Collisions using the Kinematics of Lepton+Jets Events*, *Phys. Rev. D 71, 052003*, (2005), [hep-ex/0504053v2](#);
- [92] Stefano Camarda, Monica D’Onofrio, Sebastian Grinstein, Mario Martinez, Lorenzo Ortolan, Veronica Sorin *Measurement of $Z/\gamma^* \rightarrow e^+e^- +jets$ Production Cross Section with $6 fb^{-1}$* [CDF Note 101342](#);
- [93] Stefano Camarda, Monica D’Onofrio, Sebastian Grinstein, Mario Martinez, Lorenzo Ortolan, Veronica Sorin *Measurement of $Z/\gamma^* \rightarrow \mu^+\mu^- +jets$ Production Cross Section with $6 fb^{-1}$* [CDF Note 10155](#);
- [94] R. Brun, et al., *GEANT: Simulation program for particle physics experiments. User guide and reference manual*, [CERN-DD-78-2-REV](#) (1978);
GEANT: Detector description and simulation tool, Cern Program Library Long Writeup W5013, (1993);
- [95] H.L. Lai et al, *Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions*, *Eur. Phys. J. C 12* (2000) 375, [[hep-ph/9903282](#)];
- [96] O. Saltó, *Measurement of Inclusive Jet Cross Sections in $Z/\gamma (\rightarrow e+e) + jets$ Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV with the CDF Detector*, Ph.D. Thesis, U.A.B., Barcelona (2008), [FERMILAB-THESIS-2008-17](#);
- [97] A. Bhatti et al, *Determination of the jet energy scale at the Collider Detector at Fermilab*, *Nucl. Instrum. Meth. A566* (2006) 375 [[hep-ex/0510047](#)];
- [98] Johnson, T. Junk, *SECVTX Mistag Matrices for $5.6 fb^{-1}$ of data up to p28*, [CDF Note 10179](#) (2010);
- [99] [CDF II b tag webpage](#);
- [100] A. Nappi *A pitfall in the use of extended likelihood for fitting fractions of pure samples in a mixed sample*, *Comput.Phys.Commun.180:269-275*, (2009), [arXiv:0803.2711v2](#) [[physics.data-an](#)];
- [101] Joint Physics SF Spreadsheet [JPScaleFactor.txt](#);
- [102] Joint Physics Working Group,
http://www-cdf.fnal.gov/internal/physics/joint_physics/;
- [103] CDF Collaboration, *W/Z+Jets and W/Z+HF Production at the Tevatron* CDF Public Note 10536, (2010) ,[CDF Note 10536](#);