Università di Pisa

Corso di Laurea Magistrale in Fisica

Anno Accademico 2011/2012

Tesi di Laurea Magistrale

Measurement of the ratio $\mathcal{B}(t \to W b)/\mathcal{B}(t \to W q)$ in $t\bar{t}$ dilepton channel at CDF

Candidato:
Camilla Galloni

Relatore:
Prof. Giorgio Chiarelli

Contents

Introduction

In the Standard Model of fundamental interactions, the top quark together with the bottom quark, constitute the third quark family. Top quark was postulated in 1973 by M. Kobayashi and T. Maskawa [1], but it was discovered during Tevatron Run I by both CDF and DØ only in 1995 [2]. The top quark is the most massive elementary particle known up to now with a mass of 173.20 ± 0.87 (stat+syst) [3], very close to the scale of the electroweak simmetry breaking. Produced at Tevatron in proton-antiproton collision through strong interactions, due to its very short life time ($\sim 5 \times 10^{-25}$s), about 20 times shorter than the strong interaction timescale, the top quark decays weakly in a W boson and a down-type quark before the adronization, giving the chanche to study the properties of a *bare* quark. In the Standard Model picture the top quark decay rates in down-type quaks (d,s,b) are proportional to the $\mid V_{tq} \mid^2$ the Cabibbo-Kobayashi-Maskawa (CMK) matrix element that relates the top and the down-type quark. In the hypothesis of three generations of quarks, the unitarity of the CKM leads to $\mid V_{tb} \mid^2 = 0.99915^{+0.00002}_{-0.00005}$, so the top quark decays primarily and almost exclusively to Wb. If this hypothesis is denied, a fourth quark generation would remove the constrains on $\mid V_{tb} \mid$ and lower values would be possible, giving rise to effects in the top properties (cross section, lifetime), but also in CP violation and B mixing. $\mid V_{tb} \mid$ could be estimated directly measuring the cross section of single top production, or it can be extracted from the top decay rate in the $t\bar{t}$ sample. In fact the ratio (R) between the branching fraction of top quark decaying to b quark and the branching fractions of the top quark decaying to any kind of down quark is related to $\mid V_{tb} \mid$.

$$R = \frac{\mathcal{B}(t \to Wb)}{\mathcal{B}(t \to Wq)} = \frac{\mid V_{tb} \mid^2}{\mid V_{td} \mid^2 + \mid V_{ts} \mid^2 + \mid V_{tb} \mid^2} \tag{1}$$

If the previous constrains are assumed, R is expected to be $0.99830^{+0.00006}_{-0.00009}$. Durig both Run I and Run II, CDF performed several measurements of R combining the lepton plus jets ($l + jets$) decay channel with the dilepton one, where both the W bosons decay into a lepton and a neutrino. The most recent combined measurement was performed with a luminosity of 162 pb^{-1} and found a value of R= $1.12^{+0.21}_{-0.19}$(stat)$^{+0.17}_{-0.13}$(syst) extracting R> 0.61 at 95% of CL [4] . Also the DØ collaboration measured R [5] using 5.4 fb^{-1} using the $l+jets$ and the dilepton channels, obtaining R =0.90±0.04 (stat+syst) with R> 0.79 at 95% of CL. Since the old measurement of R was dominated in the error by the low statitics, recently CDF updated the result in the l+jets channel using the complete dataset, up to a luminosity of 8.7 fb^{-1} [8]. The analysis performed a simoultaneous fit over R and the top pair production cross section, finding R = 0.94 ± 0.09 (stat+syst) and $\sigma_{p\bar{p} \to t\bar{t}} = 7.5 \pm 1.0$ (stat+syst). Thus, assuming the CKM Matrix unitarity and the three quark generations, a $\mid V_{tb} \mid$ = 0.97 ± 0.05 is found in agreement with the Standard Model prediction. In order to have a complete information from the CDF data, we decided to perform a new measurement of R in the dilepton sample ($t\bar{t} \to W^+ q W^- \bar{q} \to q\bar{q}\bar{\ell}\ell\bar{\nu}\nu$) using the whole dataset.

My analysis is based on the number of b-jets found in $t\bar{t}$ events using the dilepton sample with at least 2 jets in the final state. The charged leptons could be either electrons or muons. Tau leptons are not included. We use SecVtx algorithm, based on the reconstruction of a secondary vertex in the event, in order to identify a jet coming from b-quark fragmentation (b-tagging). Due to the high purity of the $t\bar{t}$ signal in dilepton events it is possible to perform a kinematic measurement of the $t\bar{t}$ cross section. Our strategy is to use this result to make prediction on the number of $t\bar{t}$ events. We divide our sample in subsets according to dilepton type (combination of the lepton type), number of jets in the final states and events with zero, one or two tags. The comparison between events and the prediction, given by the sum of the expected $t\bar{t}$ estimate and the background yield, in each subsample is made using a Likelihood function. Our measured value for R is the one which maximizes the Likelihood, i.e. gives the best match between our expectation and the observed data. We measure: $\sigma_{p\bar{p}\to t\bar{t}} = 7.05 \pm 0.53_{stat} \pm 0.42_{lumi}$, R= 0.86±0.06 (stat+syst) and, in the hypothesis of CKM matrix unitarity with three quark generations, $\mid V_{tb} \mid = 0.93 \pm 0.03$. Our analysis on the $\sigma_{p\bar{p}\to t\bar{t}}$ was performed independently of the official dilepton analisys on the $t\bar{t}$ production cross section. So it represents also a valuable crosscheck for the official analysis.

In chapter 1, a brief introduction to the theoretical framework is given. The standard model of elementary particles and the Quantum Cromodynamic theories are introduced. Then the top quark is presented, with a short descpription of its properties, as its mass, its production mode and its cross section. Some previous results on R are listed as well. Later we present the experiment that collected our data, both the collider (chapter 2) and the detector (CDF)(chapter 3). In chapter 4 we describe the physics object reconstruction, so how we collect the detector signals and translate those into physical particles traversing our detector. The event selection is described in chapter 5, where we report the complete list of the selection requirements and we estimate our sample composition. In chapters 6 and 7 we report our results for the $t\bar{t}$ production cross section and R. An indirect measurement for $|V_{tb}|$ is given as well.

Chapter 1

Standard Model and Top Quark Physics

1.1 The Standard Model

Our present understanding of the fundamental constituents of matter and of their interactions is expressed in a theoretical framework called the standard model (SM) of elementary particles and fundamental interactions. The SM was developed in the 1960's and 70's and has been extensively tested experimentally. Whenever a prediction for an experimental observable could be made by the Model, excellent agreement with experiment was found. The SM integrates two gauge theories: Quantum Chromodynamics (QCD), describing the strong interactions, and the electroweak (EW) theory of Glashow-Weinberg-Salam (GWS), which unifies the weak and the electromagnetic interactions. These are both quantum field theories, and therefore the Standard Model is consistent with both quantum mechanics and special relativity.

1.1.1 Particle Classification

In the SM there are two families of elementary particles, fermions (with spin 1/2) and bosons (with integer spin). There are 12 elementary fermions and their associated antiparticles. The 6 elementary fermions interacting by the EW force only are named leptons and the 6 ones interacting by both the EW and the strong force are named quarks.

Leptons The elementary leptons comprise three weak isospin doublets, the charged electron (of unit electric charge by definition) and its neutrino of zero charge, the charged muon and its neutrino, and the charged tau lepton and its neutrino. Being neutrinos electrically neutral, they only experience the weak force.

Quarks Also the elementary quarks are gathered in three weak isospin doublets, each with an " up " and a "down" element. Quarks are electrically charged. Within a doublet, the down-type quark has a unit of electric charge less then the up-type quark, the charges being $+2/3$ in the up and $-1/3$ in the down. The quark types are conventionally named " flavours " . The first doublet comprises the up and down quark flavours, the second comprises the charm and strange quarks, and the third the Top and beauty (or bottom) quarks (u,d,c,s,t,b). Quarks carry a strong interaction charge which is called "color". Quarks are also in the fundamental representation of the SU(3)-color symmetry. There are three states of color named red, green, blue and they combine in physical particles such that every composition of these colors is color-neutral. Antiquarks

carry anticolor charges. Quarks do not exist free in nature, but always in bound states such that their combination is color neutral. This feature is called asymptotic freedom and is peculiar of a non-abelian gauge theory, as the $SU(3)_c$ is. In fact the strong force strength is supposed to increase to infinity with increasing quark-quark distance, bounding quarks in color-neutral groups of three (called barions) or quark/antiquark pairs (called mesons); all barions and mesons are color composed so that to be globally colorless.
Gauge Bosons The other fundamental particles, the gauge bosons, are the interaction mediators. The carrier of the electromagnetic force is the photon γ, which is massless and chargeless. The weak force is mediated by three vector bosons: W^+ , W^- , Z^0 , the strong force is mediated by gluons (g), which are an octet in color space. Particles, elementary or not, may have a corresponding anti-particle (indicated in the following with an upper bar) with opposite electric charge and magnetic moment (examples are $\ell^+\ell^-$, *ttbar*).

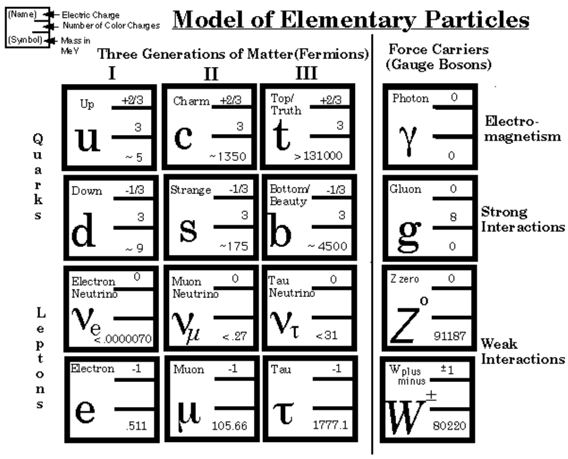


Figure 1.1: The three families for elementary particles in the Standard Model Theory.

1.1.2 The Electroweak Theory

The Electroweak (EW) theory unifies the weak isospin non-Abelian group SU(2) and the weak hypercharge (Abelian) group U(1) in SU(2)×U(1). The weak force distinguishes between left and right handed components of fermions and allows parity and charge conjugation violations in weak processes.
Before the EW unification in a single theory, there are a $U_Q(1)$ symmetry related to the electromagnetic field, where the electric charge is the coupling constant, and a $SU_L(2)$ simmetry related to charged and neutral current interaction[1]. In order to preserve the $SU_L(2)$ symmetry when constructing the isospin triplet of weak currents, it became necessary to modify the U_1 electromagnetic group generator to account for the right-handed interactions. The hypercharge Y was then introduced to replace the electric charge as a group generator with the definition $Y = 2(Q + T_3)$, where Q is the electric charge and T_3 the third component of the particle weak isospin. The EW theory is built around the conservation of the weak isospin and weak hypercharge making the

[1] L stands for left-hand and Q for electric charge.

Lagrangian invariant under local gauge $SU(2)_L$ $U(1)_Y$ transformations. The gauge invariance of the electroweak Lagrangian is complicated by the non-zero mass of particles and carriers of the weak force $W^\pm$ and Z^0 [10]. To generate those masses a weak-isospin doublet of fundamental scalar fields $\Phi = (\varphi^-, \varphi_0)$ was introduced in the Lagrangian with a potential like:

$$V(\Phi^\dagger\Phi) = \mu^2(\Phi^\dagger\Phi) + |\lambda|(\Phi^\dagger\Phi) \tag{1.1}$$

where λ is the self coupling of the scalar field. If μ^2 is chosen to be negative, the $SU(2)_L$ U_Y symmetry is spontaneously broken and just the $SU(1)_Q$ survives. The gauge bosons combine and acquire masses. This mechanism, called spontaneous symmetry breaking, was proposed by Peter Higgs et al. in 1964 [9] and permits to introduce the new field in the vacuum without loosing the local gauge invariance. In fact it provides masses to the quark [11] and W and Z bosons as well as to a massive and chargeless scalar gauge boson called Higgs Boson. The Higgs boson interaction is thus responsible for generating the large masses for the weak gauge bosons ($M_W = 80$ GeV , $M_Z = 91$ GeV) explaining the short range of the weak force, and for lepton and quark masses including the extremely large mass of the Top quark ($M_{top} = 170$ GeV). The very large value of the top quark mass, close to the value of the higgs potential at its minimum, justifies the suspect that this quark might play a special interface role between the SM and a more fundamental theory beyond it.

The Cabibbo-Kobayashi-Maskawa Matrix

Quark mass eigenstates are not exactly the same as the electroweak eigenstates. The quark eigenstates of electroweak charged current interactions are given by

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \quad \begin{pmatrix} c \\ s' \end{pmatrix}, \quad \begin{pmatrix} t \\ b' \end{pmatrix}$$

where the mixing is described by the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix, whose elements specify the strength of the coupling in each transition between quarks i and j.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

Here the different V_{ij} are constants to consider in calculation of Feynman vertices connecting the quark i with the quark j and are fundamental in calculating processes amplitude and cross section. Although there are 9 complex elements in the CKM matrix, imposing unitarity, they found that the CKM matrix is function of only 4 free parameters[2], which can be expressed as 3 angles and one CP (charge-parity) violating phase. This corresponds to the CKM form:

$$V = \begin{pmatrix} c_1 & -s_1c_3 & -s_1s_3 \\ s_1c_2 & c_1c_2c_3 - s_2s_3e^{i\delta} & c_1c_2s_3 + s_2c_3e^{i\delta} \\ s_1s_2 & c_1s_2c_3 - c_2s_3e^{i\delta} & c_1s_2s_3 + c_2c_3e^{i\delta} \end{pmatrix}$$

where s and c refer to sin and cos and their subscript to the angle θ_i and δ is then the CP violating phase. Thus there are four parameters which are θ_1 , θ_2 , θ_3 and δ. CP violation, though very small, is now well established, but CPT (charge, parity and time inversion tranformation) is

[2]Five parameters in the CKM matrix can be eliminated redefining quark fields up to a phase by $u_i \to e^{i\phi}u_i$.

believed to be the preserved underlying symmetry. In a similar way it is possible to parametrize the CKM matrix in terms of A, λ, η and ρ according to the **Wolfenstein** parametrization, which emphasizes the relative amplitudes of the elements of the matrix, showing that transitions across two quark generations, or more generally transitions that involve further off-diagonal elements of the CKM matrix elements are suppressed with respect to transitions involving diagonal elements. This parametrization maintains the unitarity to $O(\lambda^4)$:

$$V = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4).$$

In the SM the unitarity of the CKM matrix must hold:

$$\sum_{i=1}^{3} V_{ij}V_{ik}^{*} = \sum_{i=1}^{3} V_{ji}V_{ki}^{*} = \delta_{ij} \tag{1.2}$$

Expanding eq. 1.2 for any j and k yields nine equations, of which the six equations involving the off-diagonal elements of δ_{ij} describe triangles in the complex plane. These six triangles fall into two groups of three, differing only by their orientation in the complex plane: these are the so-called unitary triangles.

The free parameters of the CKM matrix are not specified by the theory and must be determined by experiments. The study of processes involving flavor-changing charged weak interaction (i.e. matter-antimatter oscillations of mesons or weak decays) allows the measurements of physical observables (oscillation frequencies, decay rates) that depend on real quantities such as the moduli of elements $|V_{ij}|$ in various combinations. These measurements can be converted into measurements of the length of the sides and interior angles of the unitary triangle.

1.1.3 Quantum Chromodynamics

Quantum chromodynamics (QCD) is the theory of the strong interactions between quarks and gluons. Quarks carry a single color charge while gluons, which are mediator of color flow, are an octect state of color (i.e.they are carrying color and anti-color states). In $SU(3)_C$ the three colors give nine total color states for the gluon: an octet and a color singlet. However, the singlet is colorless and so in nature there are only 8 possible colored gluons. Quarks only exist in colorless bound states with integer charge. Colored partons[3] will be confined in objects which are, as a whole, colorless. We can only observe color singlet quark-antiquarks bound states (*mesons*) or of three quarks or three antiquarks (*baryons*). Mesons and baryons are collectively denoted as hadrons and, being composed of quarks, are subject to strong interactions. The feature of the strong force is that the coupling becomes increasingly large with separation distance. The coupling constant of QCD (α_S) is a running function of the momentum transferred in the interaction, q^μ, which is approximately given by

$$\alpha_S(q^2) = \frac{12\pi}{(33-2n_f)\log(q^2/\Lambda^2_{QCD})} \tag{1.3}$$

where $\Lambda \sim 0.1$ GeV and n_f is the number of quark flavors whose mass is greater than the q^2 of interest [23].

[3] Quarks as well as gluons.

Thus α_S becomes increasingly small at very large q^2 (corresponding to very short approach distances). This phenomena is known as asymptotic freedom. This property allows, for high-q^2 interaction, perturbative expansion of QCD processes which remain finite. This is often the case of the collisions at Tevatron, where it is possible to calculate interaction cross sections as perturbative expansions.

When highly energetic quarks or gluons are produced in high energy physics experiment a process called *showering* takes place: after a parton is produced in an interaction, the potential with the parton system of origin, tries to keep it bound until the strength reaches a breaking point where $q\bar{q}$ pairs are created. The new partons are approximately collinear with the original parton and combine into meson or baryons (*hadronization*) in such a way that a spray of colorless particles is observed which move close to the same direction. The final state in which we observe the parton generated in the interaction is a collimated *jet*.

Most of the processes occurring in the fragmentation are non-perturbative and not completely theoretically under control, since during the fragmentation process the particle energies become successively smaller and perturbative QCD is no longer applicable. Phenomenological models are usually applied in order to describe completely jet features [24]. These models are relevant to Monte Carlo code that describes the jet formation and properties.

1.2 Top quark at Tevatron

The Top quark of charge +2/3 is the weak isospin partner of the charge -1/3 bottom quark in the third generation of elementary fermions. With its extraordinarely large mass of around 170 GeV, the Top quark is the heaviest known elementary particle. After a long hunting started in Europe at the $Sp\bar{p}S$ at CERN, it was discovered in 1995 by the CDF and DØ Collaborations at the Tevatron, Fermilab.

Top quark completes the third quark generation and interacts with other particles through electroweak and strong forces. The last combination of measurements performed at Tevatron yields to a top-quark mass of $m_t = 173.20 \pm 0.87(stat{+}syst)$ [3]. Its mass is about 35 times bigger than the bottom quark, and this result can be used, with a measurement of W mass, to set limits on Higgs boson mass. Given the large mass, the predicted lifetime it $\tau \approx 5 \times 10^{-25}$ s, one order of magnitude smaller than the time scale for hadronization. Therefore the top quark decays through weak interaction without forming bound states. This leads to the possibility to study a bare quark with well defined properties, not disturbed by the hadronization process. In the Standard Model the CKM matrix forces the top quark to decay with a very high probability into a b quark and a W boson, while the decays in light flavor quarks (down or strange) are suppressed by small values of $|V_{td}|$ and $|V_{ts}|$.

1.2.1 Top-pair Production

The Standard Model theory can be used to calculate cross sections and decay rates for elementary particles, evaluating the matrix element M of the process and integrating it on all possible final states. At the Tevatron, the collisions are between composite particles, protons and antiprotons, made up respectively of two u and one d and their antiparticles. These quarks are called *valence quarks* and are bound by virtual gluons that can split into quark-antiquarks pairs, called *sea quarks*. In this scheme the momentum P of the proton (antiproton) is shared between all the constituents, called partons. The i-th parton carries a fraction of the total momentum that is $P_i = x_i \cdot P$ and its distribution is described by Parton Distribution Functions (PDFs) $f_i(x_i, \mu^2)$, corresponding to the probability to find the given parton inside the hadron with momentum fraction x_i, when probed at an energy scale μ^2. In the top-quark production the typical energy scale of the interaction is usually set to the order of the top mass $\mu = m_t$. Since the Tevatron is a symmetric $p\bar{p}$ collider, the square of the center of mass energy of two interacting partons is defined by

$$\hat{s} = (x_1 P_p + x_2 P_{\bar{p}})^2 = 4x_1 x_2 s \tag{1.4}$$

where s is the square of the center of mass energy of the colliding protons and antiprotons ($\sqrt{s} =$ 1.96 TeV at the Tevatron). The minimum square center of mass energy of the interacting partons requested to produce a couple $t\bar{t}$ is $\hat{s}_{min} = 4m_t^2$. This value leads to momentum fractions greater than

$$x_1 x_2 \geq \frac{\hat{s}_{min}}{s} \approx 0.032 \tag{1.5}$$

For this fraction values the partons responsible of the production of the top-pairs are mostly valence quarks, as can be seen from Figure 1.2.

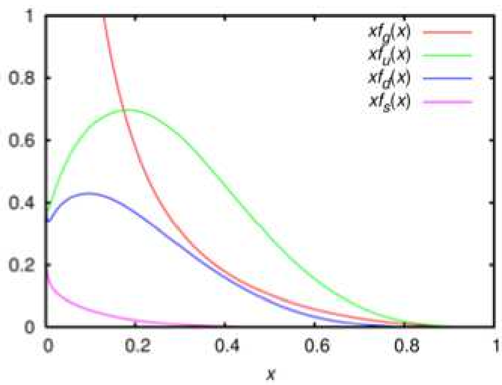


Figure 1.2: Parton distribution function for $\mu = m_t$ [25]

At the Tevatron energy scale, it is calculated that roughly the 85% of production cross section is due to quark-antiquark annihilation while the remaining fraction is due to gluon-gluon interactions. The $t\bar{t}$ cross section, corresponding to the hard process $(\hat{\sigma}(\hat{s}, m_t))$ can be written as a sum over the PDFs and on all the interacting partons, in the following way:

$$\sigma_{p\bar{p}\to t\bar{t}}(\sqrt{s},\, m_t) = \sum_{i,\, j=q,\, \bar{q},\, g} \int \hat{\sigma}_{ij\to t\bar{t}}(\hat{s},\, m_t) f_i(x_i,\, m_t) f_j(x_j,\, m_t) dx_i dx_j \tag{1.6}$$

The parton-parton cross section can be calculated as a perturbation series in the strong coupling constant $\alpha_s(m_t^2)$. The LO Feynman diagrams of $t\bar{t}$ production are shown in Figure 1.3 and in Figure 1.4. At the NLO contributions due to initial and final state radiation, gluon splitting and *bremsstrahlung* are added to the leading order calculation. The current NNLLO approximated top pair production cross section at Tevatron is $7.22^{+0.31}_{-0.47}\,{}^{+0.71}_{-0.55}$ [29] [4] for a top quark with mass $m_t = 173.3$ GeV/c^2.

[4]The first error set denotes the total theoretical uncertainty, the second the PDF+α_S error.

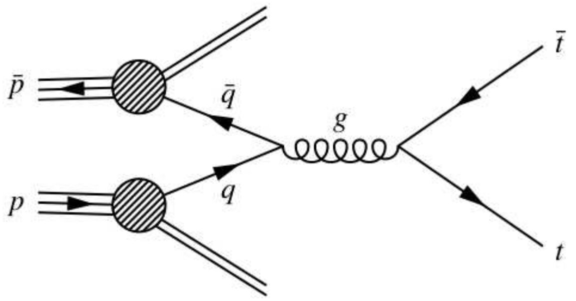


Figure 1.3: Leading order Feynman diagram of the $t\bar{t}$ production via quark annihilation

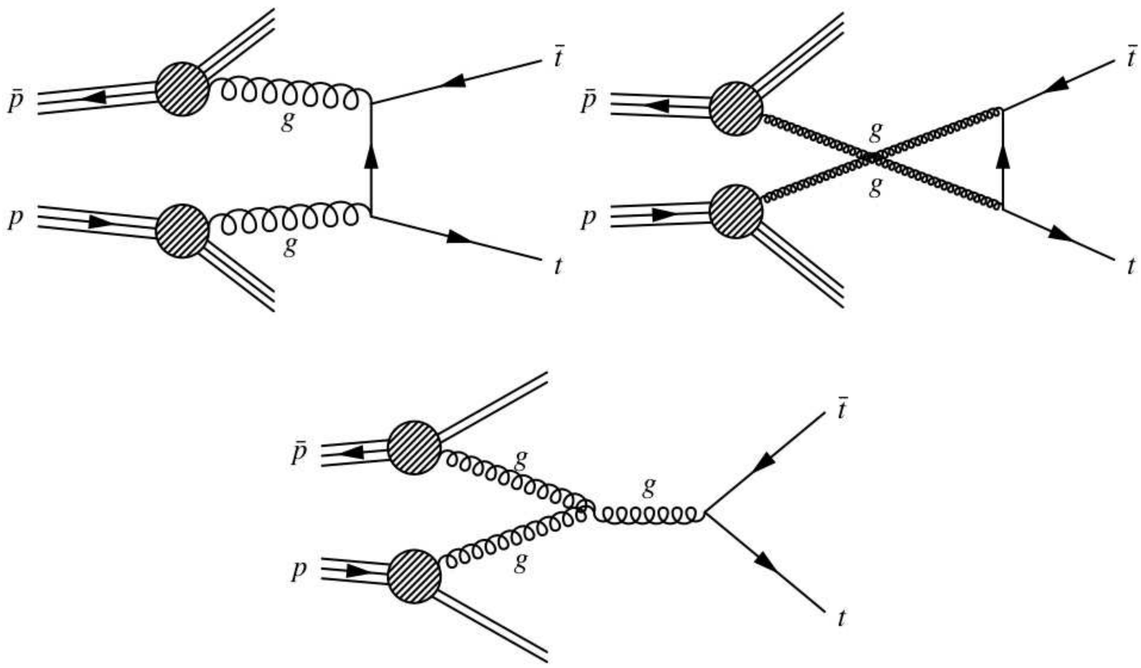


Figure 1.4: Leading order Feynman diagrams of the $t\bar{t}$ production via gluon fusion

1.2.2 $t\bar{t}$ Cross Section

The $t\bar{t}$ production cross section has been calculated to $O(\alpha_s^3)$ in perturbative QCD. An estimate at order $O(\alpha_s^4)$ was also made with soft gluon resummation techniques. Higher order corrections are expected to increase the tree-level cross section by about 30%. For a Top mass of 170 GeV the theory predicts a cross section of $\sigma_{t\bar{t}}$ 6 pb. Since the total $p\bar{p}$ inelastic cross section is about 80 mbarn, we expect one Top quark event every about 10^{10} interactions. The expected Top quark rate at the Tevatron is extremely weak. This huge background of undesired events is a real challenge in the search for top events.

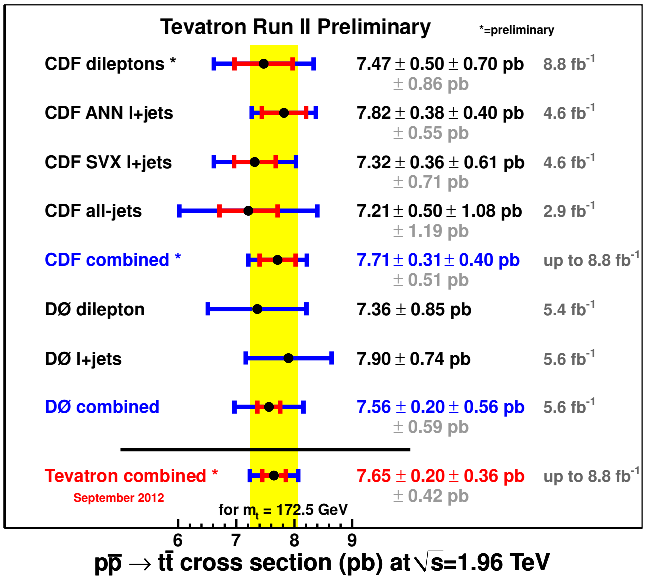


Figure 1.5: Cross section measurements by CDF compared with theoretical prediction. This plot is updated to September 2012.

1.2.3 Top Quark Decay

The Top quark decay is mediated by the weak force. The dominant Top quark decay branching fraction is B(t→Wb) with a minimal contribution by t→Wd, t→Ws. These additional decay channels are allowed by SM, but are highly suppressed due to the very small values of the off-diagonal elements in the quark flavor mixing into weak eigenstates (the Cabibbo-Kobayashi-Maskawa CKM matrix). The object of this thesis is the measurement of the ratio between B(t→W b) and B(t→Wq), where q= b, s, d, as an experimental test of the SM prediction.
The final states of the $t\bar{t}$ system are determined by the independent decays of the two W bosons produced in the t and $\bar{t}$ decays:

$$t\bar{t} = (bW^+)(\bar{b}W^-) \tag{1.7}$$

each of the two produced W bosons can decay either leptonically in a lepton and in a neutrino of the same family (with the same branching ratio) or hadronically into a pair of the two lower mass quark doublets. A schematic representation of the $t\bar{t}$ decay is reproduced in Fig.1.6. Therefore we can have a fully-leptonic, fully-hadronic or semileptonic $t\bar{t}$ final state, depending of the W's decay paths.

Different channels produce different experimental signatures in the detector. The leptonic channels involving τ's are difficult to isolate because of the τ signature. The fully hadronic channels suffer from a large QCD background of multijet states. The single-lepton final states involving electrons or muons, which can be fully reconstructed from the experimental observables are best suited for measurements such as the Top mass measurement. Due to the presence of a lepton in

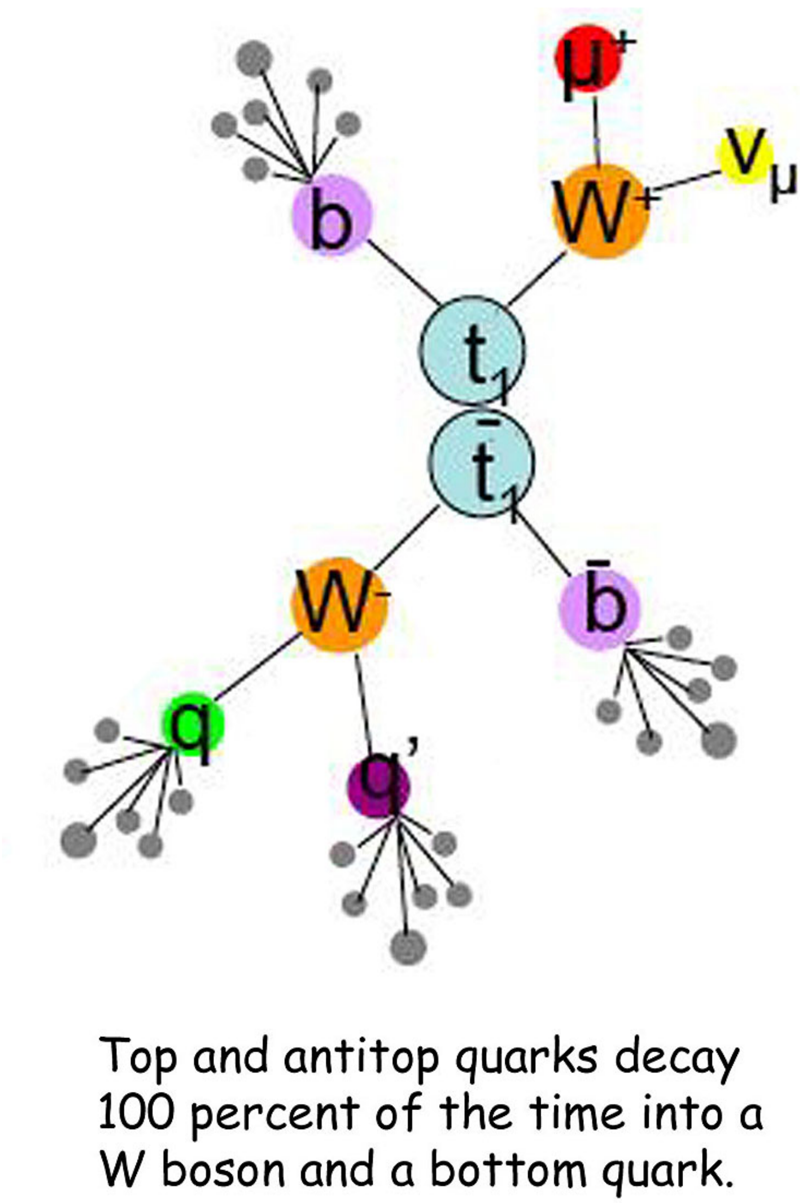


Figure 1.6: A schematic view of a possible $t\bar{t}$ decay: a W boson decays leptonically in a lepton and in a neutrino, while hadronically into a pair of the two lower mass quark doublets.

the final state, QCD background affects less this channel if compared to the full hadronic one, still retaining a relatively large number of signal events. All these characteristics make this channel the best one also for measurement of important paramenters, such as the R ratio. Measurement of the Top properties in the dilepton channel is complicated by the two non-observable neutrinos in the final state. The request of two leptons makes this channel suffer from low statistics, but make the signal over background ratio very large. In measurements where the purity of the selected data sample is important, as the one that is the object of this thesis, dilepton channel gives a fundamental contribution in reaching a better precision.

1.3 Previous measurements of the top quark branching fraction ratio

Measurements of R were performed in CDF, first in RunI and then in RunII. The first measurement was performed using only 80pb^{-1} in the lepton plus four jet samples. The idea behind the measurement technique was based on the calculation of the efficiency to tag a b in an event and on the determination of the number of selected events with a model in which R was one of the free parameters. According to the number of tags requested (zero, one or two) data was organized in bins. The background was computed using a kinematical-based analysis and by comparing observed and expected number of events in each bin R was fit. This analysis was performed again with the final Run I CDF data set (109 pb^{-1}) using also data collected in the $t\bar{t} \rightarrow ll\nu\nu$ channel. The final result, $R = 0.94^{+0.31}_{-0.24}(stat + syst)$ or $R > 0.56$ at 95 % C.L. was the first measurement of this quantity [28].

CDF performed a new analysis at the beginning of Run II, using 162 pb^{-1} [27]. Several improvements in the estimate of the backgrounds in different tag bins were applied. The result $R > 0.61$ and $|V_{tb}| > 0.78$ at 95 % CL (assuming three generation of quarks and the unitarity of the CKM matrix) was again largely dominated by the statistics of the sample. In Fig.1.7 we show this result.

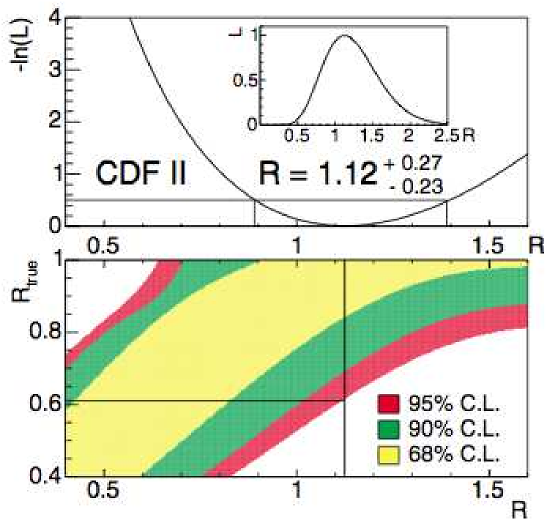


Figure 1.7: The upper plot shows the negative likelihood function at the minimum and the likelihood as a function of R in the inset. The lower plot shows 95% (red), 90% (green) and 68% (yellow) confidence level bands. The measurement of $R = 1.12$ implies $R > 0.61$ at the 95% C.L.

At the Tevatron, the most precise R measurement to date was performed by the DØ Collaboration using 5.4 fb^{-1} [5]. It was obtained using both the lepton plus jet and the dilepton channels.

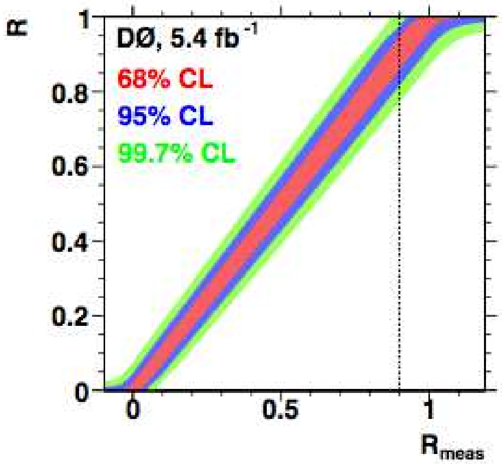


Figure 1.8: Confidence Level bands calculated by Feldman-Cousins technique. The dotted line is in correspondance of the measured value.

Table 1.1: Summary of Dzero measurement results.

Parameter	Value	Lower Limit
$\sigma_{p\bar{p}\to t\bar{t}}$	$7.74^{+0.65}_{-0.57}$(stat.+syst.) (pb)	-
R	0.90 ± 0.04(stat.+syst.)	> 0.92
$\|V_{tb}\|$	0.95 ± 0.02(stat.+syst.)	> 0.96
Parameter (dilepton channel)	**Value**	**Lower Limit**
$\sigma_{p\bar{p}\to t\bar{t}}$	$8.19^{+1.06}_{-0.92}$	
R	0.86 ± 0.05	
$\|V_{tb}\|$	0.94 ± 0.04	

This analysis makes use of a simultaneous fit of R and of the top pair production cross section $\sigma_{p\bar{p}\to t\bar{t}}$. A new method to calculate the number of $t\bar{t}$ was introduced in the analysis: two new PYTHIA Monte Carlo samples were generated forcing $t\bar{t} \to WbWq_l$ and $t\bar{t} \to Wq_lWq_l$, where q_l stands for d or s quarks, for a top quark of mass $m_t = 172.5GeV$. The number of $t\bar{t}$ events in the i-th tag bin was described as a function of R and of the number of $t\bar{t}$ events with zero, one or two light quark jets. Using a maximum likelihood technique, contents of each jet bin in data are compared to expectations and $\sigma_{t\bar{t}}$ and R are extracted. The results are summarized in Fig. 1.8 showing the confidence level intervals. The measured value for the branching fraction ratio, cross section and $|V_{tb}|$ are summarized in Tab. 1.1.

The last CDF measurement of R was performed in the lepton+jets channel using the full CDF dataset collected until the end of the Tevatron run in September 2011, correspondig to 8.7 fb^{-1}. The analysis strategy was similar to the one adopted by DØ the comparison between the total prediction and the observed data in each subsample is performed using a Likelihood function and R

obtained by maximizing this likelihood while leaving the top pair cross section as a free parameter in a recursive procedure. In Table 1.9 the measured value for the branching fraction ratio, cross section and $|V_{tb}|$ are shown.

Lepton+Jets	CDF Preliminary 8.7 fb^{-1}
$\sigma_{p\bar{p}\to t\bar{t}}$ (pb)	7.5 ± 0.3 (stat) ± 0.9 (syst)
R	0.94 ± 0.04 (stat) ± 0.09 (syst)
$\|V_{tb}\|$	0.97 ± 0.05 (stat+syst)

Figure 1.9: Summary of CDF measurement results.

Recently, also the CMS Collaboration measured R $= 1.023^{+0.036}_{-0.034}$ with a data semple corresponding to 17 fb^{-1}, obtaining $|V_{tb}| = 1.011^{+0.018}_{-0.017}$, in agreement with the Standard Model.[6]

Chapter 2

The Tevatron Collider

2.1 The Tevatron Collider

The Tevatron is a proton-antiproton collider located at the "Fermi National Accelerator Laboratory" (FNAL) near Chicago (USA) that produced $p\bar{p}$ collisions at a center-of-mass energy of $\sqrt{s} = 1.96\,TeV$. In the following we will use the present tense although the Tevatron terminated its operation in September 2011. Proton and antiprotons collide at two interaction points, where the "Collider Detector at Fermilab" (CDF) and DØ detectors are installed. The proton and antiproton beams are the result of a complex apparatus, which involves proton and antiproton production, antiproton storage, an intermediate acceleration chain up to the injection into the Tevatron ring. The Tevatron acceleration complex is shown in Fig. 2.1.

2.1.1 The Proton Source

The process leading to the $\mathrm{p\bar{p}}$ collisions starts with a Cockroft-Walton electrostatic preaccelerator and its Proton Source, which contains an hydrogen ions source. Hydrogen gas is ionized into electrons and H^+ ions. The latter are accelerated and strike on a cathode surface made of vapor coated cesium, characterized by a low work function. The H^+ in the interaction can absorb two electrons and transform into a negatively ionized hydrogen atoms, H^-, which are then accelerated to a kinetic energy of 750 KeV by a preaccelerator. Preaccelerated H^- ions are focused and injected into the LINAC, where they first reach an energy of 116 MeV and later they are accelerated to an energy of 400 MeV and collected into bunches by traveling through a 150 m long chain of 15 Hz rate radio-frequency (RF) accelerating cavities. Prior to being injected into the *Booster*, a synchrotron of 75 m radius, the H^- ions pass through a carbon foil which strips their electrons off. In the *Booster* the protons are accelerated in "batch″ of 84 bunches, each one made of $\sim 6\,10^6$ protons time-spaced of about 19 ns, to 8 GeV by a number of RF cavities. Later they are transferred to another synchrotron, called *Main Injector* [1], which brings their energy up to 150 GeV. This is the final step before protons are injected into the Tevatron.

[1]Completed in 1999 for Run II, it is located in a 3 km circumference tunnel, which houses also the antiproton *Recycler* and is approximately tangent to the Tevatron.

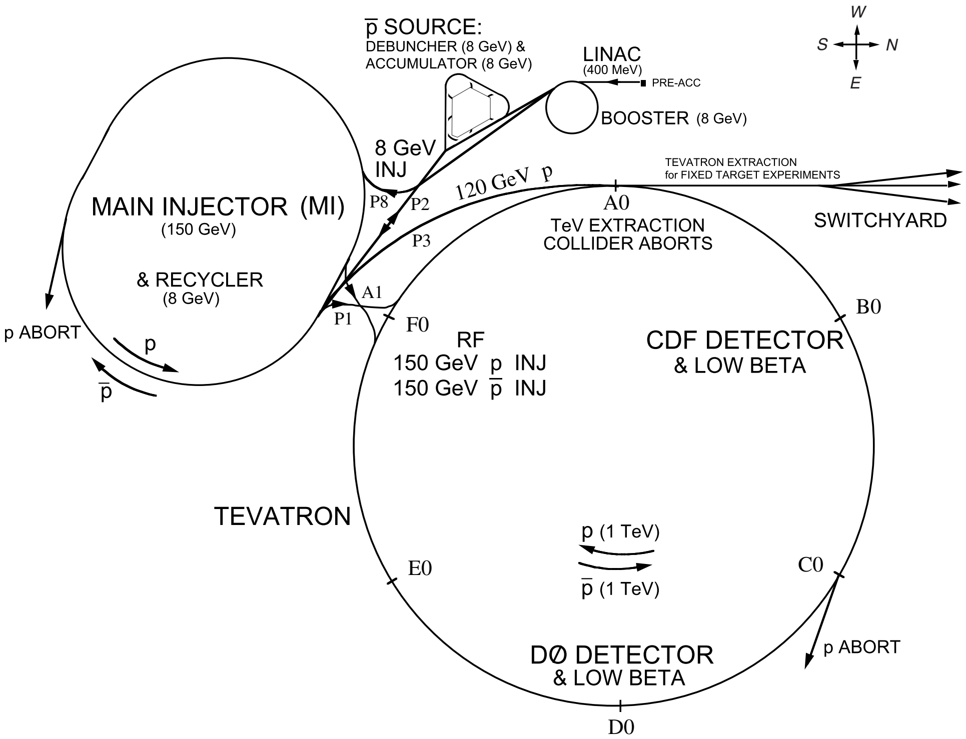


Figure 2.1: The accelerator system operating at FNAL.

2.1.2 The Main Injector

The Main Injector [7] is an oval synchrotron with a mean radius of 0.5 km, that accelerates particles from 8 Gev up to 150 GeV using radio frequency systems. It also accelerates protons up to 120 GeV for the antiproton production and accepts antiprotons back from the Recycler and accelerates them to 150 GeV for the injection in the Tevatron. One of the main purpose of the Main Injector is providing bunches for Tevatron more intense than the ones coming form the booster. From the booster a short batch made of 8 bunches is extracted and then accelerated to 150 GeV. At this energy the bunches are *coalesced*, i.e. pushed together to from one intense bunch that is injected into Tevatron. This is repeated 36 times. The same is also done for antiprotons that are coalesced in groups of 4 bunches, for 9 times, and injected into Tevraton in the opposite direction with respect to the protons, until 36 bunches are delivered.

2.1.3 The Antiproton Source

The production of the antiproton beam is significantly more complicated. The cycle starts with the extraction from the *Main Injector* of a 120 GeV proton beam, which is compacted by quadrupole magnets and directed onto a Nickel alloy target, wich is kept into rotation to reduce asymmetrical depletion. The collisions creates a variety of different particles, among which are $\bar{p}$, that are produced with an efficiency of about $18\,\bar{p}/10^6\,p$. The particles, coming off the target at different angles, are focused into a beam line by means of a magnetic lithium collection lens. In order to select only the antiprotons, the beam is sent through a pulsed magnet which acts as a charge-mass spectrometer. The emerging antiprotons, which have a bunch structure similar to that of the incident protons and a large momentum spread, are stored in the *Debuncher*, a storage ring where the $\bar{p}$ momentum spread is reduced via bunch rotation, adiabatic debunching and stochastic cooling [2] [33].

At the end of the debunching process, the bunch structure is destroyed resulting in a continuous beam of 8 GeV antiprotons which are successively transferred to the *Accumulator*. This machine is a triangle-shaped storage ring, housed in the same tunnel with the debuncher, where the antiprotons are further cooled down and stored until all the debuncher cycles are completed. When the collected antiprotons saturate the accumulator acceptance ($\sim 6 \times 10^{11}$ antiprotons), they are transferred to the *Recycler* [3], a 8 GeV fixed energy storage ring with a larger acceptance, made of permanent magnets and placed in the Main Injector enclosure. In the Recycler the size and spread of the antiproton beam is further shrunk by the electron cooling process: in one of the sections of the recycler a beam of electrons travels close to the antiprotons at the same velocity, absorbing energy from the antiprotons. When a current sufficient to create 36 bunches with the required density is available, the $\bar{p}$ are injected into the main injector where they are accelerated to 150 GeV. At this point also the antiprotons are ready to be injected into the Tevatron, in opposite direction with respect to the proton beam.

2.1.4 The Tevatron Ring

The Tevatron is 1 *km*-radius superconducting synchrotron, that accelerates particles from 150 GeV to 980 GeV. The proton and antiproton beams circulate in opposite directions in the same beam pipe. Electrostatic separators produce a strong electric field that keeps the two beams away form each other except at the collision points. The beam is steered by 774 super-conducting dipole magnets and focused by 240 quadrupole magnets with a maximum magnetic field of 4.2 Tesla. A cryogenic system based on a liquid nitrogen followed by a liquid helium state cools down the Tevatron magnets to 4.2 K, at which temperature the niobium-titanium alloy of the magnet coils becomes superconducting. The process of injecting particles into the machine, accelerating them, and initiating collisions is referred to as a "shot". It starts with the injection from the main injector of 150 GeV protons, two bunches at a time. Once the proton beam is in the machine, groups of four antiprotons bunches are mined from the recycler, accelerated to 150 GeV in the main injector and injected into the Tevatron. The beams contain both 36 bunches, grouped in three trains, separated by an abort gap of 2.6 μs. These gaps are necessary to keep a proper and safe environment for the collision and for the removal of the beam at run end or in case of malfunction. The bunches in every single train have a time spacing of 396 ns. Each proton bunch contains $\sim 10^{12}$ particles, while the antiproton ones are made up of $\sim 10^{11}$ particles. The RF cavities accelerate the beams to 980 GeV, and then some electrostatic separators switch polarity to cause the beams to collide at two points. Each interaction point lies at the center of a particle detector: DØ named after its location in the Tevatron optics, and CDF located at BØ. Successively, beams are scraped with remotely-operated collimators to remove the beam halo and, as soon as the beam conditions are stable, the experiments begin to take data. A continuous period of collider operation with the

[2] Stochastic cooling is a technique used to reduce the transverse momentum and energy spread of a particle beam without any accompanying beam-loss. This is achieved by applying iteratively a feedback mechanism that senses the beam deviation from the ideal orbit with a set of electrostatic plates, processes and amplifies the signal, and transmits an adequately-sized synchronized correction pulse to another set of plates downstream.

[3] Antiproton availability is the most limiting factor at the Tevatron for attaining high luminosities: keeping a large antiproton beam inside the Recycler has been one of the most significant engineering challenges and the excellent performance of the recycler has been an achievement of prime importance for the good operation of the accelerator.

same protons and antiprotons beams is called a “store”.

Parameter	Run II value
number of bunches (N_b)	36
revolution frequency [MHz] (f_{bc})	1.7
bunch rms [m] σ_l	0.37
bunch spacing [ns]	396
protons/bunch (N_p)	2.7×10^{11}
antiprotons/bunch ($N_{\bar{p}}$)	3.0×10^{10}
total antiprotons	1.1×10^{12}
β^* [cm]	35

Table 2.1: Accelerator nominal parameters for Run II configuration.

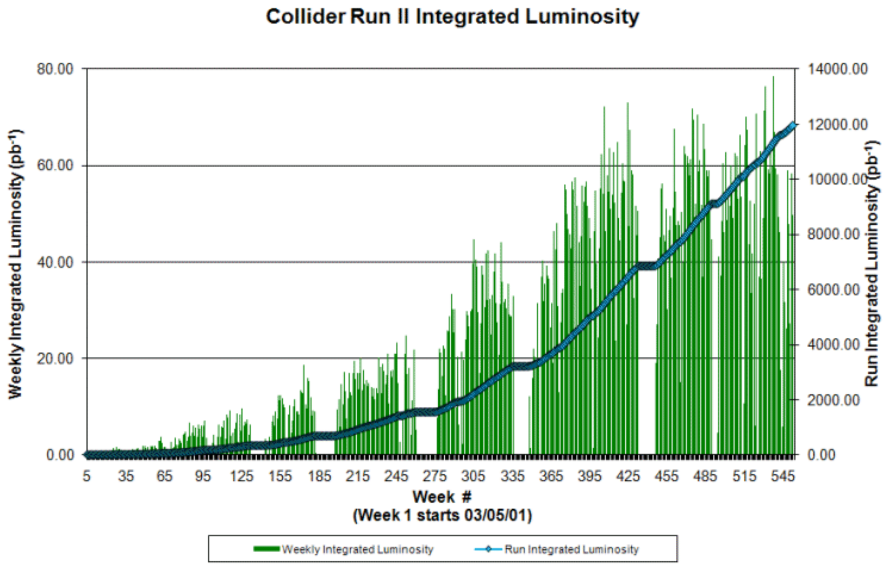
2.1.5 Luminosity and Tevatron Performance

The performance of a collider is evaluated in terms of two key 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 final state particles. The latter is defined as the interaction rate per unit cross section of the colliding beams. In the absence of a crossing angle or position offset of the beams, the luminosity at CDF or DØ is given by the expression:

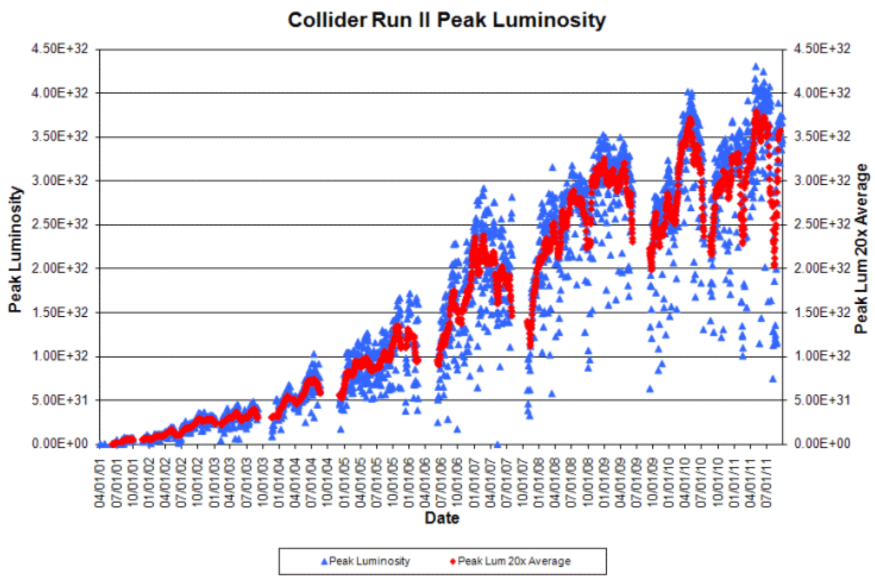
$$\mathcal{L} = \frac{f_{bc} N_b N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} F\left(\frac{\sigma_l}{\beta^*}\right), \tag{2.1}$$

where f_{bc} is the revolution frequency, N_b is the number of bunches, $N_{p(\bar{p})}$ is the number of protons (antiprotons) per bunch, and $\sigma_{p(\bar{p})}$ is the rms gaussian fit to proton (antiproton) beam size at the interaction point. F is a form factor with a complicated dependence on the beta function (β^*) value at the interaction point [4], β^*, and the bunch length, σ_l. Tab. 2.1 shows the design Run II accelerator parameters while Fig. 2.2 shows the evolution of the integrated luminosity, defined as $L = \int \mathcal{L} dt$, and the instantaneous luminosity at the start of Tevatron stores during the Run II. The steady increase of the integrated luminosity and the continuous improvement of the instantaneous luminosity up to a valueof $\mathcal{L} = 4.31 \cdot 10^{32}$ cm^2 s^{-1} on 4^{th} of May 2011, prove the outstanding performance of the accelerator. The Tevatron program was terminated on September 30, 2011. During the Run II the Tevatron delivered 12 fb^{-1} of data per experiment, $\sim$ 10 of which were collected by the CDF and DØ detectors.

[4] The suitable normalized beta function (β^*) represents a measure of the transverse beam size along the accelerator ring.



(a)



(b)

Figure 2.2: Integrated luminosity as a function of the Run II weeks (2.2a) and Tevatron peak luminosity as a function of the calendar date (2.2b). Empty periods correspond to Tevatron shut-down periods.

Chapter 3

The CDF II Detector

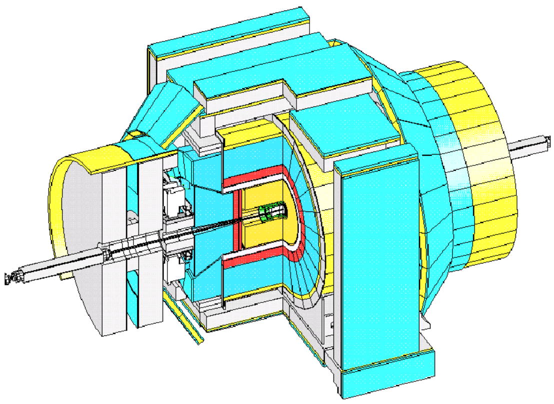


Figure 3.1: Isometric view og the CDF II Detector.

The CDF II detector, in operation from 2001 to 2011, is an azimuthally and forward-backward symmetric apparatus designed to study the $p\bar{p}$ collisions at the BØ interaction point of the Tevatron. The basic goal for CDF is to measure the momentum, energy and, if possible, the identity of particles produced in $p\bar{p}$ collisions over as large a fraction of the solid angle as possible. Thus CDF is a general purpose, cylindrical-shaped detector (Fig. 3.1) , which consists of:

- a **tracking system**, which comprises three silicon microstrip trackers (Layer 00, `SVXII` and `ISL`) and an open-cell drift chamber (`COT`) inside a superconducting solenoid, that provides a constant 1.4 T magnetic field parallel to the beam direction, with the purpose of bending into helices the trajectories of charge particles to allow the determination of their momentum and charge;

- a **Time of Flight** system (TOF), located outside the `COT`, for measuring the mass of charged particles with momenta up to 2 GeV/c;

- a **calorimeter system**, with the purpose of measuring the energy of charged and neutral particles;

- **muon chambers and scintillators**, used to track and identify muons, that pass through the calorimeters interacting as minimum-ionizing-particles (m.i.p.);
- **luminosity monitors**, for the instantaneous luminosity measurement, necessary to derive cross section from event yields.

In the following we will briefly describe the parts of the detector most relevant to this measurement.

3.0.6 Coordinates System and Standard Definitions at CDF

CDF adopts a left handed Cartesian coordinate system with origin at the nominal BØ interaction point, coincident with the center of the drift chamber. The positive z-axis lies along the nominal beam-line and has the direction of the proton beam (eastwards). The x-y plane is therefore perpendicular to the beam-line, with the y-axis pointing upwards and the x-axis in the horizontal plane, pointing radially outward with respect the center of the accelerator ring. Since the colliding beams of the Tevatron are unpolarized, the resulting physical observations are invariant under rotations around the beam line axis. Thus, a cylindrical (r, ϕ, z) coordinate system is particularly convenient to describe the detector geometry, where

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \phi = \tan^{-1} \frac{y}{x} \ . \tag{3.1}$$

A momentum-dependent particle coordinate, named *rapidity*, is also commonly used in particle physics for its transformation properties under Lorentz boost in z direction. The rapidity is defined as

$$Y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \ , \tag{3.2}$$

where E is the energy and p_z is the z component of the particle momentum. Rapidity intervals turn out to be Lorentz invariant. In the relativistic limit, or when the mass of the particle is negligible, rapidity depends only upon the production angle of the particle with respect to the beam axis, $\theta = \tan^{-1} \frac{\sqrt{x^2+y^2}}{z}$. This approximation is called *pseudorapidity* η and is defined as

$$Y \overset{p \gg m}{\rightarrow} \eta = -\ln \left(\tan \frac{\theta}{2} \right) \ . \tag{3.3}$$

A value of $\theta = 90°$, perpendicular to the beam axis, corresponds to $\eta = 0$. Since the event-by-event longitudinal position of the interaction is distributed around the nominal interaction point with a $30\, cm\, rms$ width, sometimes a distinction between the detector pseudorapidity (usually indicated with η_{det}), measured with respect to the (0,0,0) nominal interaction point, and the event pseudorapidity (η), which is measured with respect to the z position of the actual event vertex, is considered. The spatial separation between particles in the detector is commonly given in terms of a Lorentz invariant variable defined as:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \ . \tag{3.4}$$

This parameter is used to define a cone around a single particle and it is of great importance in the jet reconstruction algorithm. Since the quadrimomentum balance can be applied only on the transvers plane, due to the unknown longitudinal fraction of quadrimomentum carried by the interaction partons in the hard process, it is very useful to introduce a set of variables which are the transverse momentum and the transverse energy, defined as $p_T = p \sin\theta$ and $E_T = E \sin\theta$, respectively.

3.1 The Tracking System

A three-dimensional tracking of charged particles is achieved through an integrated system consisting of three inner silicon subdetectors, just outside the beampipe, and a large outer drift-chamber, all immersed in the 1.4 T magnetic field of a superconducting solenoid . The silicon detectors provide a precise determination of the track impact parameter, the azimuthal angle and the z coordinate at production, whereas the drift chamber has excellent resolution on the transverse momentum, ϕ and η. The combined information of the tracking detectors provides very accurate measurements of the helical paths of charged particles inside the detector. We will describe this system starting from the devices closest to the beam and moving outwards (see Fig. 3.4a).

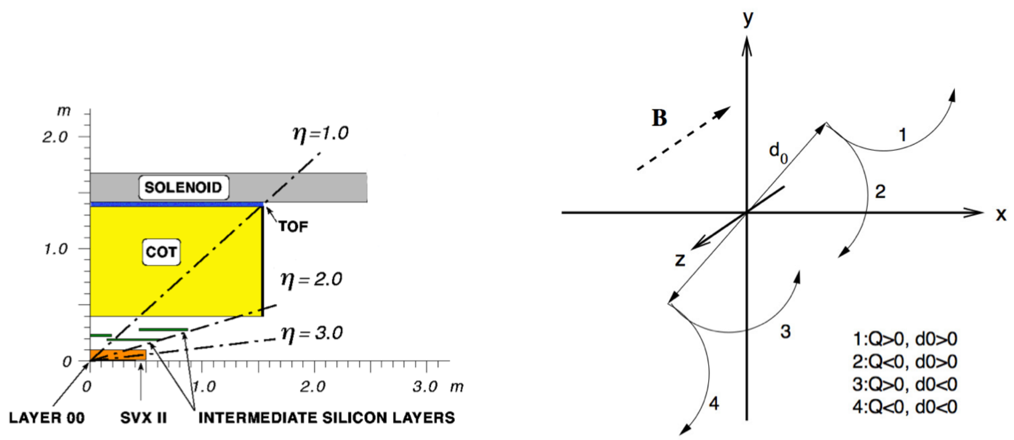


Figure 3.2: (a) The CDF II tracker layout showing the different subdetector systems. (b) Schematic drawing of the impact parameter d_0. The sign of the impact parameter is defined as positive or negative with reference to the direction of the track momentum vector (in the X,Y quadrant in the drawing).

3.1.1 The Silicon Tracker

The full CDF silicon detector is composed of three approximately cylindrical coaxial subsystems: the Layer 00 (L00), the *Silicon Vertex* detector (SVX) and the Intermediate *Silicon Layers* (ISL). Silicon sensors operate as reverse-biased p-n junctions. By segmenting the p or n side of the junctions into “strips” and reading out the charge deposition separately on every strip, the location where a charged particle traverses the sensor, is measured. At CDF the distance between two strips changes depending upon the cosidered detector $60 \div 110\mu$ m. There are two types of microstrip detectors: single- and double-sided. In single-sided detectors only the *junction*, the p side of the detector is segmented into strips, while double-sided detectors have both sides segmented into strips. In general single-sided sensors have strips parallel to the z direction and provide only r-ϕ position measurements, while double-sided detectors have strips at an angle (stereo angle) with respect to the z direction on one side and, therefore, provide also information on the particle position along z.

`L00` is a 90 cm-long, radiation hard, assembly of single-sided silicon detectors, structured in longitudinal strips. It is mounted directly on the beam pipe on a carbon fiber support with integrated cooling system. To avoid gaps, `L00` was built in 2 overlapping layers, alternating in ϕ, respectively at 1.35 and 1.62 cm from the beam axis. The detector covers a region

$\eta \leq 4$. Being so close to the beam, L00 allows to reach a resolution of $\sim 25-30\mu$m on the impact parameter of tracks of moderate p_T, providing a powerful tool to identify long-lived hadrons containing a b quark.

SVX is composed of three 29 cm-long cylindrical barrels, radially organized in five layers of double-sided silicon sensors extending from 2.5 cm (see Fig. 3.3a). Each barrel is segmented into 12 wedges, each covering $\sim 30°$in ϕ. The double-side structure of the wafers allows a three dimensional position measurement: one side of the sensor has axial strips (parallel to the beam), the other one, in three layers, has 90°strips (perpendicular to the beam), while in the remaining two has 1.2°stereo strips (at small angle with respect to the beam). This detector provides position information with a 12 μm resolution on the single hit and some dE/dx ionization information.

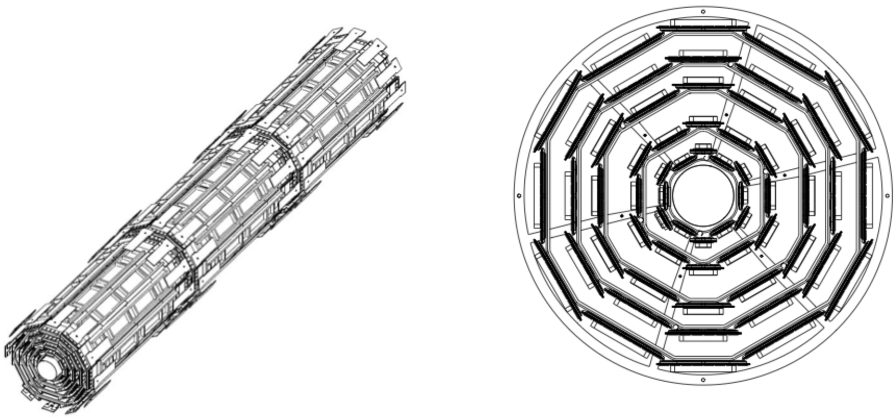


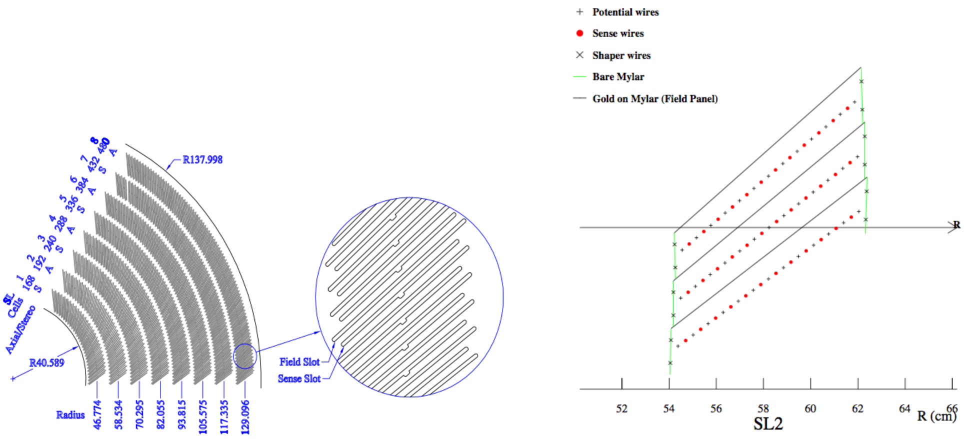
Figure 3.3: The SVX silicon detector: on the left, a three-dimensional view of the detector allows to see the barrel structure along the beam axes; on the right, the transverse plane section shows in detail the layer sequence.

ISL consists of two layers of double sided silicon wafers, similar to those of SVX, one of which is assembled in a twofold telescopes with planes at a radial distance of 22 cm and 28 cm from the beam-line and covering $1 < |\eta| < 2$. One single central layer is located at $r = 22$ cm, covering $|\eta| < 1$. The two ISL layers are important to increase the tracking coverage in the forward region, where the COT coverage is limited, and to improve the matching between SVX and COT tracks.

The combined resolution of the CDF inner trackers for high momentum tracks is $\sim 40\,\mu$m in impact parameter and $\sim 70\,\mu$m along the z direction. All silicon detectors are used in the off-line track reconstruction algorithms, while SVX plays a crucial role also in the on-line track reconstruction of the trigger system. The CDF trigger employs an innovative processor, the Silicon Vertex Trigger (SVT) [34, 35], which uses the SVX information to measure the track impact parameter on-line with a precision that allows to resolve the secondary vertices, displaced from the primary interaction point, such as those produced in B hadron decays.

3.1.2 Central Outer Tracker

The Central Outer Tracker surrounds the silicon detector (COT) [36]. It is a 3.1 m long cylindrical drift chamber, coaxial with the beam, which covers the radial range from 43 to 133 cm for $|\eta| < 1$,



(a) COTsector

(b) COT Cell

Figure 3.4: A 1/6 section of the COT end-plate (3.4a): for each super-layer the total number of cells, the wire orientation (axial or stereo), and the average radius in cm are given. The enlargement shows in detail the slot where the wire planes (sense and field) are installed. Fig. 3.4 represents the cross-section of three axial cells in super-layer 2, the arrow indicates the radial direction.

coverage decreases up to $|\eta| \simeq 2$. The COT contains 96 sense wire layers, which are radially grouped into 8 superlayers, as inferred from the end plate slot structure shown in Fig. 3.4. Each superlayer is divided in ϕ into supercells, and each supercell has 12 alternated sense and field shaping wires. So within the supercell width the trajectory of a particle is sampled 12 times. The maximum drift distance is approximately the same for all superlayers. Therefore, the number of supercells in a given superlayer scales approximately with the radius of the superlayer. The entire COT contains 30240 sense wires. Approximately half the wires run along the z direction (*axial* wires), the other half are strung with a small (2°) stereo angle with respect to the z direction (*stereo* wires). The combination of the axial and stereo information allows to measure the z positions and a three-dimensional reconstruction of tracks. Particles originated from the interaction point, which have $|\eta| < 1$, pass through all the 8 COT superlayers. The COT is filled with an argon/ethane/CF4 mixture (50:35:15). The mixture is chosen in order to have a constant drift velocity, approximately 100 μm/ns, across the cell width. The maximum electron drift time is approximately 100 ns. Due to the magnetic field that the COT is immersed in, electrons drift at a Lorentz angle of 35°. The supercells are tilted by 35°with respect to the radial direction to compensate for this effect and make the drift path perpendicular to the radial direction.

The hit position resolution in the r-ϕ plane is about 140 μm. Tracking algorithms are used to reconstruct particle trajectories (helices) that best fit to the observed hits. The reconstructed trajectories are referred to as "tracks". Particle momentum and charge are determined from the bending of tracks in the magnetic field. The COT hits are also processed on-line by the XFT (eXtremely Fast Tracker), which reconstructs the tracks used in the trigger system, (Sec. 3.5.4). The transverse momentum resolution of off-line tracks, estimated using cosmic ray events, is:

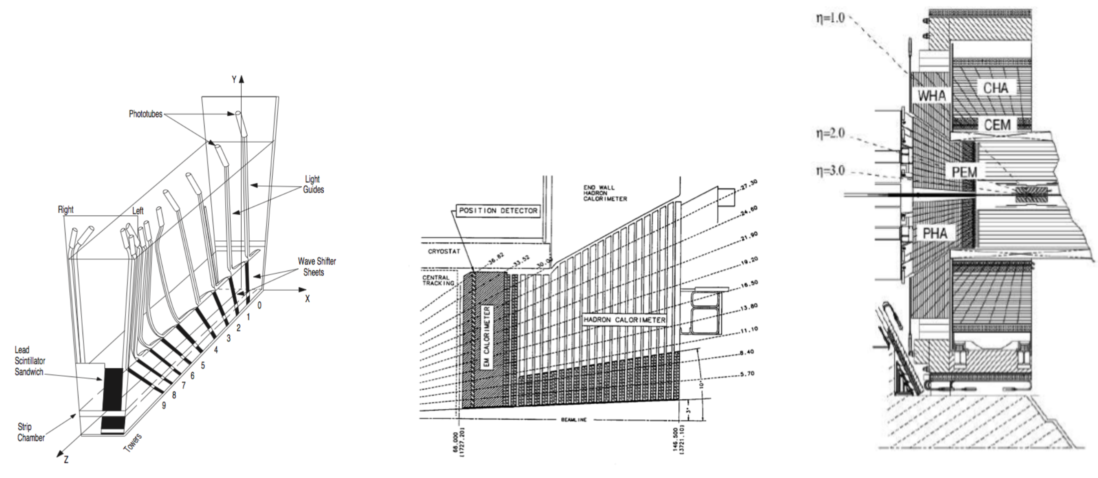
$$\frac{\sigma_{p_T}}{p_T^2} \simeq [0.015][GeV/c]^{-1} \tag{3.5}$$

for tracks with $p_T > 2$ GeV/c [37].

3.2 Calorimeters

The calorimeter system, together with the muon and tracking systems, represents one of the main sub-detector of `CDF II`. A detailed description of this system can be found in the CDF II Technical Design Report [38]. The CDF II calorimetry system has been designed to measure energy and direction of neutral and charged particles leaving the tracking region. In particular, it is devoted to jet reconstruction and it is also used to measure the missing energy associated to neutrinos. Particles hitting the calorimeter can be divided in two classes, according to their main interaction with matter: electromagnetically interacting particles, such as electrons and photons, and hadronically interacting particles, such as mesons or barions produced in hadronization processes. To detect these two classes of particles, two different calorimetric parts have been developed: an inner electromagnetic and an outer hadronic section, overall providing coverage up to $|\eta| < 3.64$. In order to supply information on particle position, the calorimeter is also segmented in towers, projected toward the geometrical center of the detector. Each tower consists of alternating layers of passive material and scintillator tiles. The signal is read out via wavelength shifters (WLS) embedded in the scintillator and light from WLS is then carried by light guides to photomultiplier tubes. The central sector of the calorimeter, covering the region $|\eta| < 1.1$, was recycled from Run I, while brand new calorimeters (called plug calorimeters) were built up to cover the forward and backward regions. Fig. 3.5b shows the plug calorimeter system while Fig. 3.5c shows an elevation view of the components of the `CDF` calorimeter: `CEM`, `CHA`, `WHA`, `PEM` and `PHA`.

3.2.1 The Central Calorimeter



(a) CEM Wedge (b) Plug calorimeters (c) Whole calorimeter system

Figure 3.5: One azimuthal electromagnetic calorimeter wedge 3.5a, the elevation view of one quarter of the plug calorimeter 3.5b. In 3.5c elevation view of the CDF detector showing the components of the CDF calorimeter: `CEM`, `CHA`, `WHA`, `PEM` and `PHA`.

Excluding upgrades on the readout electronics, needed to cope with the increased collision rate, the central calorimeter is almost the same as in Run I[1]. The Central Electro-Magnetic calorimeter (CEM) is segmented in $\Delta\eta \times \Delta\phi = 0.11 \times 15°$projective towers, called *wedges*, consisting of 31 alternate layers of lead and scintillator, for a total material depth of 19 X_0 [2]. The Central and End-Wall Hadronic calorimeters (CHA and WHA respectively), whose geometry tower segmentation matches the CEM one, use 32 steel layers sampled each 2.5 cm (5 cm in the wall) by 1 cm thick acrylic scintillator. The total thickness of the hadronic section is approximately constant and corresponds to 4.5 interaction lengths (λ_0) [3]. A perspective view of a central electromagnetic calorimeter module (*wedge*) is shown in Fig. 3.5a, where both the arrangement in projective towers and the light-collecting system are visible. A projective geometry was used in order to take advantage of the momentum conservation in the transverse plane: before the $p\bar{p}$ collision, the projection in the transverse plane w.r.t. the beam direction of the beam energy is zero, therefore this quantity must be the same after the collision. Thus, for each tower the transverse energy E_T is defined as $E_T = E \sin\theta$, where E is the energy detected by the tower and θ is the angle between the beam axis and the projective tower direction. Two position detectors are embedded in each wedge of CEM:

- The Central Electromagnetic Strip chamber (CES) is a two-dimensional stripwire chamber arranged in correspondence to maximum shower development ($\sim 5.9X_0$). It measures the charge deposit of the electromagnetic showers, providing information on their pulse-height and position with a finer azimuthal segmentation than calorimeter towers. This results in an increased purity on electromagnetic object reconstruction. The CES purpose is to measure the position and the shape of electromagnetic showers in both transverse plane and longitudinal direction, which is used to distinguish electrons and photons from hadrons.

- The Central Pre-Radiator (CPR) consists of two wire chamber modules placed immediately in front of the calorimeter. It acts as pre-shower detector by using the tracker and the solenoid coil material as radiators, resulting to be a very useful tool in rejection of electron and photon background.

Tab. 3.1 summarizes the basic quantities of calorimeter detectors. The energy resolution for each calorimeter section was measured in the test beam and, for a perpendicularly incident beam, can be parametrized as:

$$\frac{\sigma}{E} = \frac{\sigma_1}{\sqrt{E}} \oplus \sigma_2 \tag{3.6}$$

where the first term comes from the sampling fluctuations and the photostatistics of the PMTs, (stochastic term) and the second term comes from the non-uniform response of the calorimeter (constant term), and the symbol $\oplus$ indicates addition in quadrature.

[1] CDF, from 1992 to 2004, used gas proportional chambers in the central calorimeter to improve the identification of electrons and photons (*Central Preshower* (CPR) and *Central Crack* (CCR) detectors). Late in 2004 the **CDF Central Preshower and Crack Detector Upgrade** was installed. The CDF Central Preshower and Crack Detector Upgrade consist of scintillator tiles with embedded wavelength-shifting fibers, clear-fiber optical cables, and multi-anode photomultiplier readout.

[2] The radiation length X_0 describes the characteristic amount of matter transversed by high energy electrons to loose by bremsstrahlung all but $1/e$ of their energy, which is equivalent to 7/9 of the length of the mean free path for e^+e^- pair production of high energy photons. The average energy loss due to bremsstrahlung for an electron of energy E is related to the radiation length by $\left(\frac{dE}{dx}\right)_{brem} = -\frac{E}{X_0}$

[3] An interaction length is the average distance that a particle will travel before interacting with a nucleus: $\lambda = \frac{A}{\rho\sigma N_A}$, where A is the atomic weight, ρ is the material density, σ the cross section and N_A the Avogadro number.

3.2.2 The Plug Calorimeter

The plug calorimeters covers the $|\eta|$ region from 1.1 to 3.64 on both sides. Both electromagnetic and hadronic sectors are divided in 12 concentric η regions, with $\Delta\eta$ ranging from 0.10 to 0.64, according to increasing pseudorapidity. Each of them is segmented in 48 or 24 (for $|\eta| < 2.11$ or $|\eta| > 2.11$ respectively) projective towers. The actual size of these towers was chosen so that identification of electron in the narrow forward jets would be optimized. Projective towers consist of alternating layers of absorbing material (lead and iron for electromagnetic and hadronic sectors, respectively) and scintillator tiles. The first layer of the electromagnetic tiles is thicker ($10\,mm$ instead of $6\,mm$) and made of material with higher photon yield. It acts as a pre-shower detector.

Calorimeter	CEM	CHA	WHA	PEM	PHA
Absorber	Lead	Steel	Steel	Lead	Iron
Segmentation ($\eta \times \phi$)	0.1×15	0.1×15	0.1×15	$(0.1 \div 0.6)\times(7.5 \div 15)$	$(0.1 \div 0.6)\times(7.5 \div 15)$
Num. Towers ($\eta \times \phi$)	20×24	9×24	6×24	12×24(48)	11×24(48)
Thickness	$[19]X_0$,1 λ_0	4.7 λ_0	4.7 λ_0	$[23]X_0$,1 λ_0	6.8 λ_0
Resolution (%)	$14/\sqrt{E_T} \oplus 2$	$50/\sqrt{E_T} \oplus 3$	$75/\sqrt{E_T} \oplus 4$	$16/\sqrt{E} \oplus 1$	$80/\sqrt{E} \oplus 5$

Table 3.1: Summary of the main characteristics of the CDF II calorimeter system.

3.3 The Muon Chambers

Most particles produced in the primary interaction or in subsequent decays have a very high probability of being absorbed in the calorimeter system. Muons represent an exception. Muons do not interact via strong interaction with nuclei in matter either. Therefore, a muon with enough energy will pass through the calorimeter systems releasing only a small amount of its energy[4]. The muon system is the outermost layer of the CDF II detector and consists of four layers of drift cells and one layer of scintillation counters which are used to reconstruct track segments ("stubs") of minimum ionizing particles and to provide accurate timing. Stubs are matched with the `COT` information in order to reconstruct the full trajectory of the muons. Some additional steel shielding layers, in between the chambers and the calorimeters, reduce the probability for other particles to escape the calorimetric system. Four independent systems detect muons in the $|\eta| \lesssim 1.5$ pseudo-rapidity range reconstructing a small segment of their path (stub) sampled by the chambers, employing similar combinations of drift tubes, scintillation counters, and absorbers [31], [32]. The track momentum is found by pointing back the stub to the corresponding track bent in the `COT`. Scintillators serve as trigger and vetoes while the drift chambers measure the ϕ coordinate using the absolute difference of drift electrons arrival time between two cells, and the z coordinate by charge division. All types of muon detectors use single wire, rectangular drift chambers, arranged in arrays with fine azimuthal segmentation and coupled with scintillator counters. The four sub-detector systems are (see Fig. 3.6):

`CMU`: the `CMU` detector is located around the central hadronic calorimeter at a radius of $347\,cm$ from the beam-line with coverage $0.03 \lesssim |\eta| \lesssim 0.63$. It is segmented into 24 wedges of 15°. However, because of wedge effects only 12.6°of each wedge is active, resulting in an overall

[4] Muons are over 200 times more massive than electrons, so bremsstrahlung radiation, inversely proportional to the mass squared of the incident particle, is suppressed by a factor of $4 \cdot 10^4$ with respect to electrons.

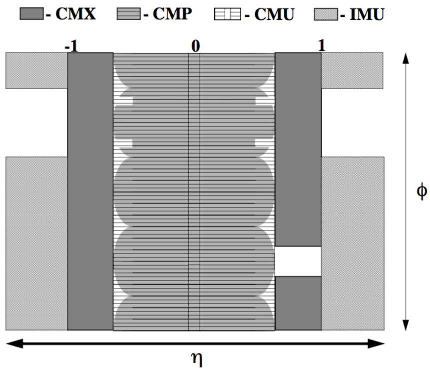


Figure 3.6: Muon detectors coverage in the η-ϕ plane.

azimuthal acceptance of 84%. Each wedge is further segmented into three 4.2°modules each containing four layers of four drift cells. Each cell operates in proportional mode, with a maximum drift time of 0.8 μs and a longitudinal resolution of $\delta\ z \simeq 10$ cm.

`CMP`: the `CMP` is a second set of muon drift chambers outside of `CMU` behind an additional 60 cm -thick steel absorber. The material further reduces the probability of hadronic punch-through to the `CMP`. Muons need a transverse momentum of about 2.2 GeV to reach the `CMP`. The `CMP` system is arranged in a box shape of similar acceptance as the `CMU` and serves as a confirmation of `CMU` for higher momentum muons. A layer of scintillation counters (CSP) is mounted on the outer surfaces of the `CMP`. The `CMP` and `CMU` have a large overlap in coverage and are often used together in indicating a muon track. `CMP` helps to cover `CMU` ϕ gaps and the `CMU` covers the `CMP` η gaps. Muon candidates which have both `CMU` and `CMP` stubs are the least contaminated by fake muons. Also `CMP` works in proportional mode, with a maximum drift time of 1.4 μs and a multiple scattering resolution of $15/(p[\mathrm{GeV}])$ cm.

`CMX`: the `CMX` consists of drift tubes and scintillation counters (CSX) assembled in conically arranged sections. The `CMX` extends the pseudo-rapidity coverage to $0.6 \lesssim |\eta| \lesssim 1$. There are 8 layers of drift chambers in total with a small stereo angle between layers. The `CMX` has a multiple scattering resolution of $13/(p[\mathrm{GeV}])$ cm and a longitudinal resolution of $\delta\ z \simeq 14$ cm. There are layers of scintillators that provide time information.

`IMU`: the `IMU` extends the pseudo-rapidity coverage even further to $1.0 \lesssim |\eta| \lesssim 1.5$. The `IMU` is mounted on the toroid magnets which provide shielding. `IMU` has a maximum drift time of 0.8 μs and a multiple scattering resolution of $(13-25)/(p[\mathrm{GeV}])$ cm.

Tab. 3.2 summarizes a few of the relevant design parameters of the detectors.

3.4 CLC detector

Absolute measurements of the instantaneous luminosity by the machine, based on beam parameters measurements, have uncertainties of the order of 15-20%. For this reason in CDF the collider luminosity is determined using gas Cherenkov counters (CLC)[30] located in the pseudorapidity region $3.7 < |\eta| < 4.7$, which measure the average number of inelastic interactions per bunch crossing. Each module consists of 48 thin, gas-filled, Cherenkov counters. The counters are arranged

Parameter	CMU	CMP	CMX	IMU
Pseudorapidity range	$\|\eta\| < 0.6$	$\|\eta\| < 0.6$	$0.6 < \|\eta\| < 1.0$	$1.0 < \|\eta\| < 1.5$
Azimuthal coverage [°]	360	360	360	270
Maximum drift time [ns]	800	1400	1400	800
Drift tube cross section[cm]	2.68 × 6.35	2.5 × 15	2.5 ×15	2.5 × 8.4
Pion interaction length	5.5	7.8	6.2	62 - 20.0
Minimum $p_T(\mu)$ [GeV/c]	1.4	2.2	1.4	1.4-2.0

Table 3.2: Design parameters of the muon detectors. Assembled from Ref. [31], [32]

around the beam pipe in three concentric layers, with 16 counters each pointing to the center of the interaction region. The cones in the two outer layers are about 180 cm long and the inner layer counters, closer to the beam pipe, have a length of 110 cm. This geometry allows detecting only particles produced at the collision point. The total signal of the counters allows separating forward inelastic interactions from background and deriving the collider luminosity from their observed rate (see Appendix A).

3.5 The CDF Trigger sistem

The collisions at Tevatron happen with a frequency of $2.5\,MHz$ (i.e every $396\,ns$). The bunch-bunch luminosity and the interaction cross section are such that in average one or a few interactions take place at each bunch crossing. With an average event size of $\sim 250\,Kb$ this represents a huge amount of data which would flow through the CDF data acquisition system (`DAQ`). The CDF `DAQ` can sustain only a small fraction of this data flow, since the maximum rate for storing data to disk is $\sim 200\,Hz$. The trigger is the system devoted to perform a quick online selection and keep only the events interesting for physics. A rejection factor of 10000 is needed to match the `DAQ` capabilities. As shown in Fig. 3.7, the CDF trigger is implemented in three levels of successively tighter and more sophisticated event selections.

The first level is hardware based, while the second is a mixture of hardware and software and the third is purely software, implemented in an on-line computer cluster.

3.5.1 `LEVEL 1`

At `LEVEL 1` the decision logic is implemented in hardware, that is the selection algorithms are hard-coded into the electronic circuits of the trigger boards. In a synchronous pipeline up to 42 subsequent events can be stored for $\sim 5.5\,\mu$ s while the hardware is taking a decision. If no acceptance decision is made within that time the event is lost. L1 decisions are made on average in about 4 μs: no dead time is expected from this level. Level 1 rejects 97% of the events, by reducing the event rates from 2.53 MHz to less than 40 kHz. The L1 decision is generated using

- XFT (extremely fast tracker), which reconstructs approximate tracks ($p_T > 1.5$ GeV) in the transverse plane by exploiting information from `COT` superlayers. These tracks are extrapolated to the calorimeters and muon chambers parts to contribute to all trigger levels.
- the calorimeter towers, which carry information on the electromagnetic and hadronic energy deposits (these can be seed for electrons or jets identification).
- the muon "stubs" (segment of tracks reconstructed in the muon chambers), which are matched to the `XFT` tracks.

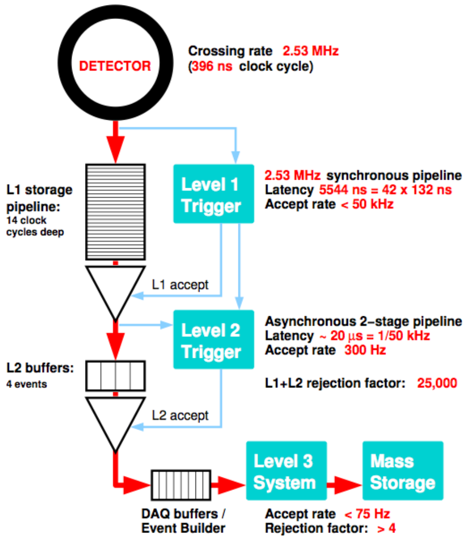


Figure 3.7: Functional block diagram of the CDF II trigger and data acquisition systems.

The XFT is a custom processor used to reconstruct two-dimensional tracks in the (r, ϕ) plane in the COT. The XFT is capable of reconstructing tracks with $p_T \gtrsim 1.5\,\mathrm{GeV}/c$ with an efficiency of about 95% and a fake rate of a few percent. The XFT has an angular segmentation of 1.25°, and an angular resolution of 0.3°. The momentum resolution is $\sigma_{p_T}/p_T^2 \approx 0.017\,[\mathrm{GeV}/c]^{-1}$. XFT sends the tracks to an extrapolation unit (XTRP) feeding three L1 elements: L1 CAL, L1 TRACK, and L1 MUON. L1 CAL and L1 MUON use extrapolated tracks and information from the calorimetry and muon systems respectively to search for possible electron, photon, jets and muon candidates. A decision stage combines the information from these low-resolution physics objects, called “primitives”, into more sophisticated objects.

3.5.2 LEVEL 2

LEVEL 2 is an asynchronous system which processes events that have received a L1 accept in a FIFO (First In, First Out) manner. It is structured as a two stage pipeline with data buffering at the input of each stage. The first stage is based on dedicated hardware processors which assemble information from a particular section of the detector. The second stage consists of a computer which uses the list of objects generated by the first stage and implements in software the event selection. Each of the L2 stages is expected to take approximately 10 μs with a latency of approximately 20 μs. The input buffers can store up to four events. After the Level 2, the event rate is reduced to about 1 KHz (rejection factor $\sim$ 40). L2 purposes are:

- to cluster the enery deposited in the towers around L1 seeds, as an approximate measure of an electron photon or jet energy.

- to use calorimeter and CES chamber information to improve separation of $e^{\pm}$ from γ.

- to improve the matching between XFT tracks and muon stubs in order to have a better muon signature.

- to provide a measurement of the track impact parameters by means of the Silicon Vertex Trigger element (SVT), which allows to select events with secondary vertexes from decay of long-lived heavy flavour hadrons.

The SVT uses SVX $r - \phi$ hits to extend XFT track primitives inside the SVX volume, closer to beam-line. The SVT improves the XFT ϕ_0 and p_T resolutions and allows measurement of the impact parameter d_0 (original XFT track primitives are beam-line constrained).

3.5.3 LEVEL 3

LEVEL 3 is a software trigger. It runs a simplified version of the offline and is operated on a cluster of $\sim$ 300 processors which reconstruct the entire event with the same accuracy as in the off-line analysis. The final decision to accept an event is made on the basis of a list of observables indicating candidate events of physical interest (top quark production events, W/Z events, Drell-Yan events, etc.). Events that satisfy the Level 3 trigger requirements are transferred onward to the Consumer Server/Data Logger (CSL) system for storage first on disk and later on tape. The average processing time per event in Level 3 is of the order of one second. The Level 3 leads to a further reduction in the output rate, with an accepted maximum of about $200\,Hz$.

3.5.4 Trigger Paths

A set of requirements that an event has to fulfill at Level 1, Level 2 and Level 3 constitutes a trigger path. The CDF II trigger system implements about 150 different trigger paths, which are periodically adjusted depending on machine luminosity and physics needs. An event will be accepted if it passes the requirements of any one of these paths. The trigger system described above exploits the information of all detector subsystems. Combining the measurements of the various subsystems it is possible to efficiently record, at the same time, events characterized by different signatures. Triggers that use a bandwidth fraction larger than the assigned one are prescaled. A trigger path is said to be prescaled by a factor N if it is configured to accept only one event out of N recorded events. Prescaling is dynamically implemented by luminosity-dependent factors during data taking. This is important in order to ensure that no trigger path reaches rates so high as to create unacceptable dead time to triggers on rare events of primary importance. During data taking the luminosity decreases with time, and consequently a number of prescale factors can be relaxed. The prescale factors decrease proportionally to the rate of triggered events, so as the number of recorded events is constant. Using dynamic prescaling ensures that optimal use for physics is made of the available luminosity.

The accepted events are recorded to tape and organized in “data sets” according to the trigger path which they satisfy.

Chapter 4

Physics Object Reconstruction

Outgoing particles from $p\overline{p}$ interactions are identified using the information provided by the CDF sub-detectors described in the previous chapter. The raw outputs from the CDF detector are electronic signal recorded by the hardware components which must be converted into physical information. From the raw data, high level objects (such as tracks, vertices, calorimeter clusters) are reconstructed and combined to identify physics objects (electrons, muons, neutrinos and jets) of interest to the analysis. Their identification will be described in the next section. The ability to detect and reconstruct charged particle trajectories is essential for particle identification and momentum measurement.

4.1 Track Reconstruction

The ability to detect and reconstruct charged particle trajectories is crucial for particle identification and momentum reconstruction. Charged particles traveling through a homogeneous solenoidal magnetic field along the z direction follow helical trajectories. Knowing that the projection of the helix on the x-y plane is a circle, to uniquely describe the motion, only five parameters are needed:

$\boldsymbol{z_0}$ – the z coordinate of the point of closest approach.

$\boldsymbol{C}$ – signed helix (half)-curvature, defined as $C = q/2R$, where R is the radius of the helix and q is the particle charge. This is directly related to the transverse momentum. When the magnetic field (B) is measured in Tesla, C in m^{-1} and p_T in GeV/c: $p_T = 0.15\ qB/|C|$.

$\boldsymbol{\phi_0}$ – ϕ azimuthal angle of the particle trajectory at the point of closest approach to the z-axis.

$\boldsymbol{d_0}$ – signed impact parameter, i.e. the radial distance of closest approach to the z-axis in the transverse plane. defined as $d_0 = q(\sqrt{x_0^2 + y_0^2} - R)$, where (x_0, y_0) are the coordinates of the center. This is schematically drawn in Fig. 4.1.

$\boldsymbol{\lambda}$ – helix pitch, i.e. $\cot(\theta)$, where θ is the polar angle 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: $p_z = p_T \cot\theta$.

Another useful quantity is the displacement of the secondary vertices of decaying particles in the transverse plane, L_{xy}:

$$L_{xy} = \frac{\widehat{x}_V \cdot \vec{p}_T}{|p_T|} \tag{4.1}$$

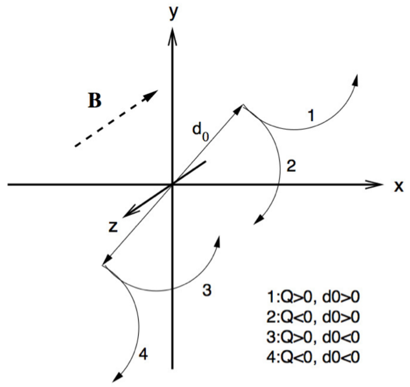


Figure 4.1: Schematic drawing of the impact parameter d_0.

where $\widehat{x}_V$ is the decay vertex position in the transverse plane.

The trajectory of a charged particle satisfies the following equations

$$\begin{aligned} x &= r\sin\phi - (r - d_0)\sin\phi_0 \\ y &= -r\cos\phi + (r + d_0)\cos\phi_0 \\ z &= z_0 + s\lambda \end{aligned} \tag{4.2}$$

where s is the projected length along the track, $r = 1/2C$ and $\phi = 2Cs + \phi_0$. 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 by the detector material. All tracks are first fit in the COT and then extrapolated inward to the silicon. This approach guarantees fast and efficient tracking with very good purity (i.e. very little contamination by fake tracks). The greater radial distance of the COT with respect to the silicon tracker results in a lower track density and consequent fewer accidental combination of hits in the track reconstruction. A brief overview of the tracking algorithms is given in Appendix B. For more details see Ref. [39], [40].

4.1.1 Primary Vertex Reconstruction

The primary vertex position for a given event is found by fitting high quality tracks to a common point of origin. At high luminosities, multiple collisions occur in a given bunch crossing. For a luminosity of $10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, there is an average of 2.3 interactions per bunch crossing. Typically, since the luminous region is sufficiently long (with $\sigma_z = 30\,\mathrm{cm}$), the primary vertices associated to the collisions are well separated in z. An iterative algorithm is used to find the vertex associated to the hardest collision: the first estimate of its position (x_V, y_V, z_V) is binned in the z coordinate, then the z position of each vertex is calculated from the weighted average of the z coordinates of all tracks within 1 cm of the first iteration vertex, with a typical resolution of $100\,\mu$m; finally the vertex associated with the highest sum of the tracks p_T is defined as the primary vertex of the event.

The envelope of all primary vertices defines the beam-line, the position of the luminous region of the beam-beam collisions through the detector. The beam-line is used as a constraint to refine the knowledge of the primary vertex in a given event. Typically the beam transverse section is circular with a width of $\sim 30\,\mu m$ at $z = 0$, rising to $\sim 50 - 60\,\mu m$ at $|z| = 40\,cm$. The beam at the interaction point is not necessarily parallel to the detector axis nor centered in the detector and moves up to fractions of millimeter as a function of time. Correction to the nominal beam line are applied before track reconstuction.

4.2 Lepton Reconstruction and Identification

Lepton reconstruction depends on the type of lepton and its direction inside the detector. In this analysis we are interested in the identification of electrons and muons. In the following we will present a brief overview of the identification criteria used in this work.

4.2.1 Electron Identification

Electrons are identified by requiring a track matched to an energy cluster in the calorimeter with an appropriate shower profile. The cluster reconstruction starts with an energetic tower in the electromagnetic calorimeter, the *seed tower*. Electrons are assumed to be massless and, in order to reconstruct the four momentum, $(E, \vec{p})$, track information is used to determine the three dimensional direction, $\vec{p}/|\vec{p}|$, while the calorimetric energy measurement gives the magnitude $E \equiv |\vec{p}|$. Also, as we are interested in leptons that come from W decay, we apply a topological isolation requirement. In this work we used the following electron categories.

Tight Central Electrons (CEM) Central electrons ($|\eta| < 1.0$) with high p_T are expected to traverse the silicon and COT detectors, leaving behind a track. Then they enter the EM calorimeter where they will produce an electromagnetic shower and deposit most of their energy. An EM cluster is found if E_{had} /$\mathrm{E}_{em} < 0.125$, where E_{em} is the electromagnetic energy of the cluster and E_{had} is the hadronic energy in the corresponding towers of hadronic calorimeter. The selection criteria applied to identify the *tight central electrons* are reported in Tab. 4.1. The observables used are:

- track p_T: the transverse momentum of the track associated to the EM cluster;
- track z_0: the position along the longitudinal direction at the point of closest approach to the beam-line;
- *Axial and Stereo Superlayers*: (SL) are the numbers of axial and stereo superlayers in the COT having at least 5 hits associated to the track in question;
- $Q \times \Delta x_{CES}$: the distance in the $r - \phi$ plane between the extrapolated track and the nearest cluster reconstructed in the CES detector multiplied by the charge of the track (to account for asymmetric tails originated from bremsstrahlung);
- E_{had}/E_{EM}: the ratio between the energy deposited in the hadronic calorimeter and the energy deposit in the electromagnetic one.
- E/P: the ratio of the EM cluster transverse energy to the track transverse momentum as measured by the tracking system; expected to be ~ 1 for electrons, it can be up two if the object is very energetic;
- *CalIso* is the calorimetric isolation, defined as

$$CalIso \equiv \frac{E_T^{\Delta R = 0.4} - E_T^e}{E_T^e} \tag{4.3}$$

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

- L_{shr}, the lateral shower profile, compares the energy distributions of the EM towers adjacent to the seed tower to the distribution derived from electron test-beam data;

- Δz_{CES} is the distance in the $r-z$ plane between the extrapolated track and the best matching `CES` cluster.

	`Tight(CEM)`
Region	central ($\|\eta < 1.0\|$)
Fiducial	track fiducial to `CEM`
Track E_T	$E_T \geq 20$ GeV
Track p_T	$\geq$ 10GeV)
Track $\|z_0\|$	≤ 60 cm
# Ax SL (5hits)	≥ 3
# St SL (5hits)	≥ 2
Conversion	false
E_{had}/E_{EM}	$\leq 0.055 + 0.00045 \cdot E$ [GeV]
CalIso	≤ 0.1
L_{shr}	≤ 0.2
E/P	< 2 unless $p_T \geq 50$ GeV/c
$\Delta x \cdot Q$	$-3 \leq q \cdot \Delta x \leq 1.5$ cm
Δz_{CES}	$\leq$ 3cm
Track	Beam constrained

Table 4.1: Definition of fully identified Tight Central Electron (`TCE`).

Forward Electrons(`PHX`) Electron candidate clusters in the plug calorimeter are built starting from a seed tower and adding neighboring towers within two towers in η_{det} and ϕ from the seed. The hadronic energy of the cluster is required to be less than 0.05 times the electromagnetic energy. Plug electrons have to be reconstructed in a well-instrumented region of the detector, defined as $1.2 \leq |\eta| \leq 2.0$ as measured by the `PES` sub-system (η_{det}^{PES}). To improve the quality of the track-cluster matching, we require at least three silicon hits in the road defined by the E_T of the cluster and z_0 of the event. This combination of electromagnetic calorimeter cluster and silicon hits is called "PHX". Additional requirements are summarized in Table 4.2. Some of the observables used to define the forward electrons are

- `PEM` 3×3 χ^2 is the χ^2 of a fit to the signal shape in electron test beam data.

- `PES` 5×9 U/V are the ratios of the charge collected in the central 5 over a set of 9 `PES` strips in the horizontal and vertical planes.

- ΔR(`PES`,`PEM`) is the angular distance between the `PEM` and the `PES` clusters.

4.2.2 Muon Identification

Muons traverse the entire CDF detector, with an energy deposition of a m.i.p. in the calorimeters. They leave hits in the outer muon chambers. A muon is reconstructed starting from a track in

	PHX
Region	Plug
$\lvert\eta_{det}^{PES}\rvert$	$1.2 \leq \lvert\eta_{det}^{PHX}\rvert \leq 2.8$
E_{had}/E_{EM}	≤ 0.05
PEM $3 \times 3\,\chi^2$	≤ 10
PES5 $\times$ 9U	≥ 0.65
PES5 $\times$ 9V	≥ 0.65
CalIso	≤ 0.1
ΔR(PES, PEM)	≤ 3.0
NSiHits	≥ 3
Track z_0	$\leq 60\,cm$

Table 4.2: Definition of Forward Electrons (PHX)

the COT and adding track segments (stubs) formed with hits in the muon drift chambers. The track origin in the $x-y$ plane is constrained to the beam position. The muon four momentum $(E, \vec{p})$ is determined by measuring the track p and assuming a massless particle: $E \equiv |\vec{p}|$. Many reconstruction categories exist. Some of them require the muons to have hits in one muon detector subsystem, and we call them stubbed muons. As for the electrons, as we expect leptons from W decays to be isolated, we will apply a topological requirement of isolation to our μ candidates.

Stubbed Muon Stubbed muon candidates are required to have a reconstructed track with a fit $\chi^2/n.d.f. < 3$. The track is required to have at least three Axial and two Stereo COT super-layers with at least 5 hits. The track $|z_0|$ has to be less than 60 cm. For all muons, the impact parameter (d_0), has to be less than 0.2 cm, and it is tightened to be less than 0.02 cm if the track has also hits in the Silicon detectors, giving a much precise measurement of the impact parameter. To reject background we also require the track to be isolated: *TrkIso* < 0.1, where *TrkIso* is defined as

$$TrkIso \equiv \frac{\sum_i p_T^{i,\Delta R=0.4} - p_T}{p_T} < 0.1 \tag{4.4}$$

where p_T is the track transverse momentum. The reconstructed track is then required to be compatible with a minimum ionizing particle (*m.i.p.*) by cutting on the energy deposited in the EM and HAD towers hit by the extrapolated track. We also require calorimetric isolation

$$CalIso = \frac{E_T^{\Delta R=0.4} - E_T}{p_T} < 0.1 \tag{4.5}$$

where $E_T^{\Delta R=0.4}$ and p_T have been defined above. Stubbed muons are divided in four categories, depending on the region of the detector that the extrapolated track is pointing to, CMUP, CMU, CMP and CMX. We measure the location of an extrapolated muon track candidate with respect to the drift direction (local x) and wire axis (local z) of a given chamber. We do not take into account possible multiple scattering in the extrapolation. We refer to these requirements as fiduciality of the track to the given muon detector. In this analysis we used two categories of muons: CMUP, muons with a stub in both the CMU and CMP systems with $|\eta| < 0.7$, and CMX, muons with a stub in the CMX system with $0.7 < \eta < 1.1$. The cuts applied to each category are summarized in Tables 4.3 - 4.4.

	CMUP
Region	Central
Track $\chi^2/n.d.f.$	≤ 3
$NAxL$(5 hits)	≥ 3
$NStL$(5 hits)	≥ 2
Track z_0	$\leq 60\,cm$
Track d_0	$\leq 0.2\,cm(\leq 0.02\,cm$ if N $SiHits > 0)$
$TrkIso$	≤ 0.1
$CalIso$	≤ 0.1
E_{EM}	$\leq 2 + \max(0,$ (p - 100) $\cdot$ 0.0115) GeV
E_{had}	$\leq 6 + \max(0,$ (p - 100) $\cdot$ 0.028) GeV
$\Delta x_{\texttt{CMU}}$	$\leq 7\,cm$
$\Delta x_{\texttt{CMP}}$	$\leq \max(6.0,\ 150/p_T\ [GeV/c^2])\ cm$
Fiduciality	$x-$fid$_{\texttt{CMU}} < 0\ cm$ z-fid$_{\texttt{CMU}} < 0\ cm$
	$x-$fid$_{\texttt{CMP}} < 0\ cm$ z-fid$_{\texttt{CMP}} <$ -3 cm

Table 4.3: Definition of CMUP central muons.

	CMX
Region	Central
Track $\chi^2/n.d.f.$	≤ 3
$NAxL$(5 hits)	≥ 3
$NStL$(5 hits)	≥ 2
Track z_0	$\leq 60\,cm$
Track d_0	$\leq 0.2\,cm(\leq 0.02\,cm$ if Number of $SiHits > 0)$
$TrkIso$	≤ 0.1
$CalIso$	≤ 0.1
E_{EM}	$\leq 2 + \max(0,$ (p - 100) $\cdot$ 0.0115) GeV
E_{had}	$\leq 6 + \max(0,$ (p - 100) $\cdot$ 0.028) GeV
$\rho_{\texttt{COT}}$	$>$ 140cm
$\Delta x_{\texttt{CMX}}$	$\leq \max(6.0,\ 125/p_T\ [GeV/c])\ cm$
Good Trigger	Run $\geq$ 150145

Table 4.4: Definition of CMX central muons. The exit radius, $\rho_{\texttt{COT}}$ is defined as $\rho_{\texttt{COT}} = (z_{\texttt{COT}} - z_0) \cdot \tan\theta$

4.3 Jet Reconstruction

The color-carrying quarks and gluons, created in the scattering process, undergo the hadronization process which produces collimated bunches of colorless hadrons (jets) which keep track of the energy and the direction of the originating parton. From the experimental point of view, the resulting shower of particles appears as a large energy deposit in a localized area of the detector (Fig. 4.2). The challenge is to recover from the detector information the initial energy, the momentum and, possibly, the nature of the particle produced in the original interaction. To reconstruct jets, CDF developed several different jet reconstruction algorithms [42].

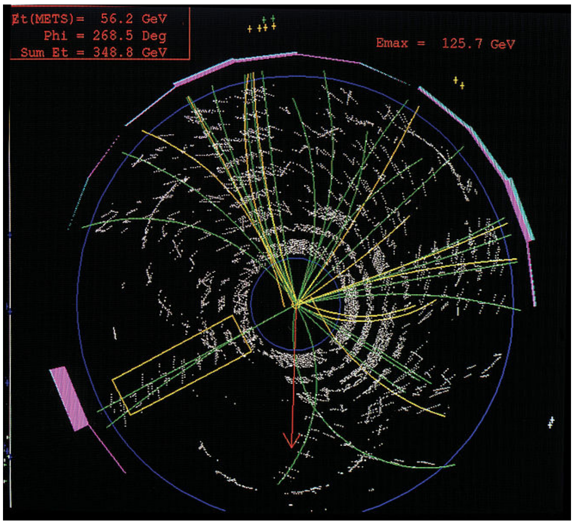


Figure 4.2: Top quark and anti top quark pair decaying into jets, visible as collimated collections of particle tracks, and other fermions in the CDF detector.

4.3.1 JETCLU algorithm

The most common jet reconstruction algorithm used at CDF is a "cone algorithm" named JETCLU [43], which consists of three steps. In the first step preclusters are built from adjacent "seed towers"

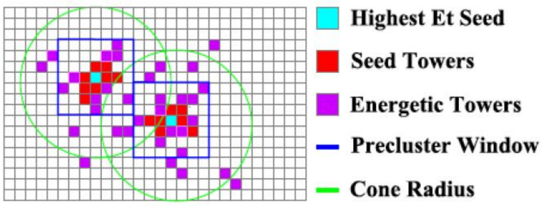


Figure 4.3: Jet reconstruction by the JETCLU algorithm.

(calorimeter towers with $E_T > 1\,GeV$). The size of these preclusters is limited to $2R_{\text{cone}} \times 2R_{\text{cone}}$ in the η-ϕ plane, where R_{cone} is the parameter of the jet algorithm which controls the size of the jets.

After that, for each precluster a cone is defined by all seed towers inside the precluster and all towers with $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < R_{\text{cone}}$ with respect to the highest E_T tower. The centroids of the cones are calculated. The identification of the members of the cones and the calculation of their centroids is repeated until the old centroids (the cone axes) agree with the new ones.

In the last step overlapping stable cones have to be reconsidered because each calorimeter tower may only belong to one jet. A pair of overlapping cones is merged if more than 75% of the transverse energy of one of the cones is shared by the other one. Otherwise they are separated using an iterative algorithm. The towers are redistributed to the cone whose centroid is closer and the centroids are recalculated until a stable configuration is reached.

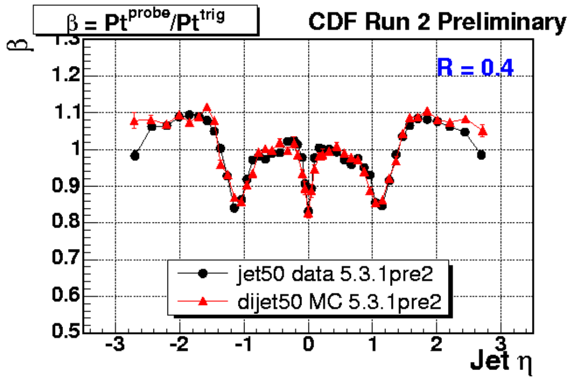


Figure 4.4: η-dependent energy scale correction factor for JETCLU with radius 0.4; a sample of events with at least one trigger tower above 50 GeV is used.

The transverse energy and the position of the reconstructed jet are then given by:

$$E_T^{\text{jet}} = \sum_i E_T^i \tag{4.6}$$

$$\eta = \frac{1}{E_T^{\text{jet}}} \sum_i E_T^i \eta^i \tag{4.7}$$

$$\phi = \frac{1}{E_T^{\text{jet}}} \sum_i E_T^i \phi^i \tag{4.8}$$

where E_T^i, η_i and ϕ_i are the energy and the position of the i-th tower.

4.3.2 Jet Energy Corrections

The four-momentum assigned to a jet must be corrected to account for detector defects and reconstruction algorithm imperfections. In order to convert the measured transverse jet energy into the transverse energy of the partons, a set of corrections to the measured jet energy ("raw energy") has been developed.
The corrections, developed using data and simulation of the CDF detector, address the response inhomogeneity in η, the contributions from multiple interactions, the non-linearity of the calorimeter response, the contribution by the underlying event (particles emitted in the event but not belonging to the hard interaction) and the jet energy flow out of the jet cone.
Each of those corrections has a fractional uncertainty, $\sigma_{JES}(p_T)$ which can be parameterized as a function of the corrected transverse momentum of the jet p_T.
They are applied in a sequence of levels (of "L-levels'") in order to correct for each bias independently [44].
The correction can be parameterized as follows

$$p_T^{\text{parton}} = (p_T^{jet} \cdot C_\eta - C_{MI}) \cdot C_{Abs} - C_{UE} + C_{OOC} = p_T^{\text{particle}} - C_{UE} + C_{OOC},$$

where the terms are described below. In this analysis we correct the energy up to level L5, so up to the "Absolute energy scale corrections".

C_η: pseudorapidity-dependent correction (L1)

The L1 corrects for non-uniformities in calorimeter response along η. It is obtained by studying the p_T balancing in dijet events, which are selected in order to have one jet ("trigger jet") in the $0.2 < |\eta| < 0.6$ region (far away from detector cracks). The other jet, called "probe jet", is free to span over the entire $|\eta| < 3$ region.
Since in a perfect detector the two jets must be balanced in p_T, a balancing fraction is formed

$$f_b = \frac{\Delta p_T}{p_T^{ave}} = 2 \cdot \frac{p_T^{probe} - p_T^{trigger}}{p_T^{probe} + p_T^{trigger}}.$$

The average of f_b in the analyzed η bin is used to define the β_{f_b} factor[1] (see Fig. 4.4)

$$\beta_{f_b} = \frac{2 + \langle f_b \rangle}{2 - \langle f_b \rangle}. \tag{4.9}$$

The final L1 correction is defined as $f_{L1}(\eta, E_T^{raw}, R) = 1/\beta_{f_b}$ and the uncertainty associated with this correction is estimated to be of the order of 1% for central jets and 7.5% for forward jets.

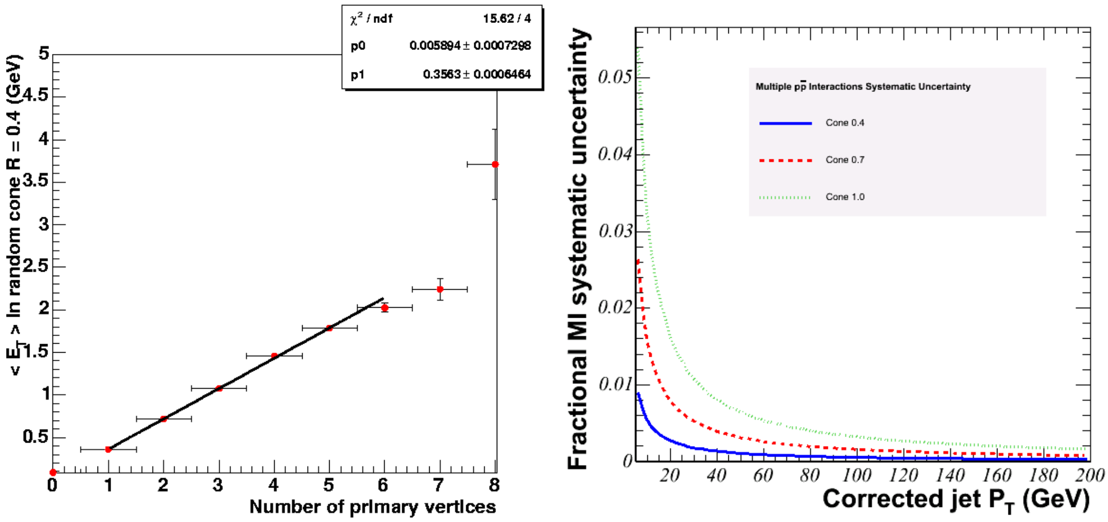


Figure 4.5: (a) E_T in R=0.4 cone as a function of the number of reconstructed primary vertexes in minimum bias events. (b) Fractional systematic uncertainties due to multiple interactions for different cone sizes as a function of jet transverse momentum.

C_{MI}: multiple interactions correction (L4)

The number of interactions occurring during beam bunch crossings follows a Poisson distribution whose mean increases with instantaneous luminosity.

[1]The definition of 4.9 is in average equal to $p_T^{\text{probe}}/p_T^{\text{trigger}}$ but it reduces the sensitivity to the presence of non-Gaussian tails which affect the $p_T^{\text{probe}}/p_T^{\text{trigger}}$ ratio.

These additional interactions, dominantly soft *minimum bias* events, cause extra unwanted energy to be deposited in the calorimeter.
L4[2] correction takes into account the number of reconstructed vertices to estimate the effect. The average energy flow in minimum bias events, which are triggered by the luminosity monitor CLC, is measured in the best-performing region ($0.2 < |\eta| < 0.6$) of the calorimeter as a function of the number of reconstructed vertices. The resulting plot is fit to a straight line (see Fig. 4.5). This is used to correct in average the energy of the jets. Because of the finite reconstruction efficiency of the vertices, this linear approximation works well for events with less than seven vertices.
This is not a serious limitation because in practice events with so many vertices are very rare.
The uncertainty on this correction is estimated to be of the order of 15%.

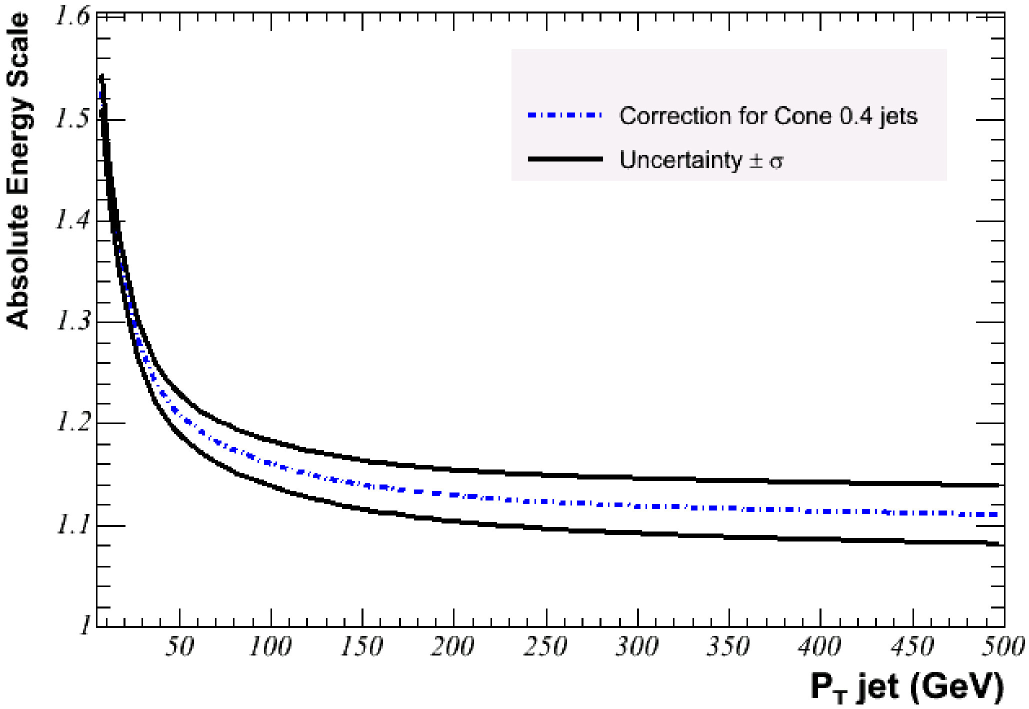


Figure 4.6: Absolute energy corrections for jets with cone size $\Delta R = 0.4$ as a function of jet p_T with uncertainty.

C_{Abs}: absolute energy scale corrections (L5)

While L1 and L4 are corrections at calorimeter level, L5 steps back to particle level. The procedure used to estimate the L5 correction factor is described accurately in [44]. It uses a MC sample of inclusive dijet events simulated with `PYTHIA` [45].
The correction is derived comparing particle jets at generator level (before they are passed through

[2]L2 and L3 have survived in the CDF jargon but are not used anymore.

the detector simulation), with calorimeter jets, as obtained from the detector simulation. These are required to be within 0.1 of each other in the $\eta - \phi$ plane to ensure that they are the same object. The probability at the distribution maximum of measuring a value of p_T^{jet} given p_T^{particle} is taken as a correction factor (see Fig. 4.6). The uncertainty on this correction is estimated to be of the order of 3.5% (15% near the edge of the calorimeter).

C_{UE} and C_{OOC}: underlying event (L6) and out-of-cone (L7) corrections

Reconstructed jet energies in hard $p\bar{p}$ interactions may contain contributions from particles created by interactions involving other partons in the colliding hadrons (spectator interactions) or by gluons from initial state radiation in the hard interaction. These contributions are called *underlying event*. On the other hand a fraction of the parton energy may be lost outside the jet cone because of final state gluon radiation, fragmentation at large angles relative to the jet axis or low p_T particles bending in the magnet field. This energy is modeled imperfectly in MC events, so a systematic uncertainty is assigned by examining photon + jets events in data and MC. A ring around the jet with a radius between 0.4 and 1.3 in the $\eta - \phi$ plane is examined, and the energy in this region is compared between data and MC simulation. The largest difference between MC events and data is taken as a systematic uncertainty.

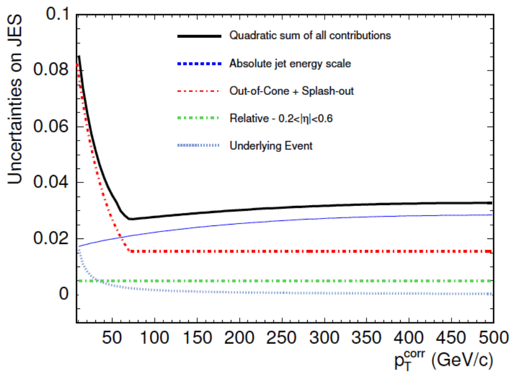


Figure 4.7: The fractional systematic uncertainty of the JES corrections as a function of the jet transverse momentum. The total uncertainty is taken as the sum in quadrature of all individual contributions.

Fig. 4.7 shows the individual fractional systematic uncertainties as a function of jet p_T in the central region, $0.2 < |\eta| < 0.6$, of the calorimeter. They are independent of each other and thus are added in quadrature to derive the total uncertainty.

4.4 Reconstruction of the Secondary Vertex: Bottom Jets Identification

The identification of B-hadrons in jets was fundamental for the discovery of the top quark in 1995 and was one of the crucial features of the searches for SM light Higgs boson at the Tevatron Collider. Jets carrying b-flavour (b-jets) are produced from b-quark hadronization in top or Higgs decays, while mostly light-quark jets are produced from the main background processes (e.g. W +jets), which contaminate the candidate event data sample. Several methods to identify b-jets exist at CDF. They take advantage of the fact that the weakly decaying B-hadrons have long lifetimes[3]. The tagging algorithms take advantage of this property and identify a point displaced from the primary vertex with tracks of large impact parameter d_0. The inner tracker, using the silicon detectors, can reconstruct those tracks with a $\sigma \simeq 50\ \mu$m. Unfortunately this algorithm is not able to distinguish between bottom and charm hadrons, that also can be tagged. Due to their shorter lifetime (D_0 lifetime = 0.6ps) and the lower track multiplicity in the second vertex the efficiency for charmed hadrons is low. We will describe the b-tagger used in this analysis, the Secondary Vertex Tagger *SecVtx*.

4.4.1 The SecVtx Algorithm

The SecVtx Algorithm is performed for all jets with $|\ \eta\ | \leq 2.4$ in the event. The algorithm searches for secondary vertices using the tracks inside the jet cone of radius $\triangle R = 0.4$. There are minimal requirements on the track quality to be used [12]:

- $p_T > 0.5$ GeV/c;
- $|\ d_0\ | \leq 0.15$ cm and $|\ d_0/\sigma_{d_0}\ | \geq 2.0$;
- $|z_0 - z_{PrimVtx}| \leq 5$ cm;
- have a minimal number of hits in the silicon detector, depending on the quality of the track and its position;
- be confirmed in the COT.

If there are at least 2 tracks with the above requirements the jet is called *taggable*.

The algorithm works following two steps and has two main operation modes: *tight* and *loose* that differs on the applied thresholds. In this work we will use the *tight* algorithm. The two steps are the following:

Pass 1 : at least 3 tracks are required and one is required to have $p_T > 1$ GeV/c. The tracks are combined in order to reconstruct a secondary vertex and a quality χ^2 is computed.

Pass 2 : begins if Pass 1 fails. The tracks requirements are tightened: $p_T > 1$ GeV/c and $|\ d_0/\sigma_{d_0}\ | \geq 3.0$, but the algorithm requires just 2 tracks, one of whom with $p_T > 1.5$ GeV/c.

If a secondary vertex is found the jets is tagged and a bi-dimensional decay length L_{xy} is calculated as the projection in the $r - \phi$ plane of the SecVtx vector (going from the primary vertex to the secondary one) into the jet axis (Fig 4.8).

The sign of L_{xy} is determined from the angle α between the jet axis and the SecVtx vector as shown in Fig.4.9. The secondary vertex coming from HF hadrons is expected to have positive and large L_{xy}, while mismeasured tracks are expected to have a symmetric distribution about zero. Thus, to reduce the background due to tracks mismeasuring $L_{xy}/\sigma_{L_{xy}} \geq 7.5$ is required.

The negative tags are however useful: in high statistics sample, we can evaluate from the total number of tagged jets, the number of false positive tag (*mistags*) taking as an estimate the number of negative tags.

[3]The B^0 has a proper life time of 1.6 ps and, usually, is boosted, so it can fly few millimeters before decaying($c\tau = 476\mu m$).

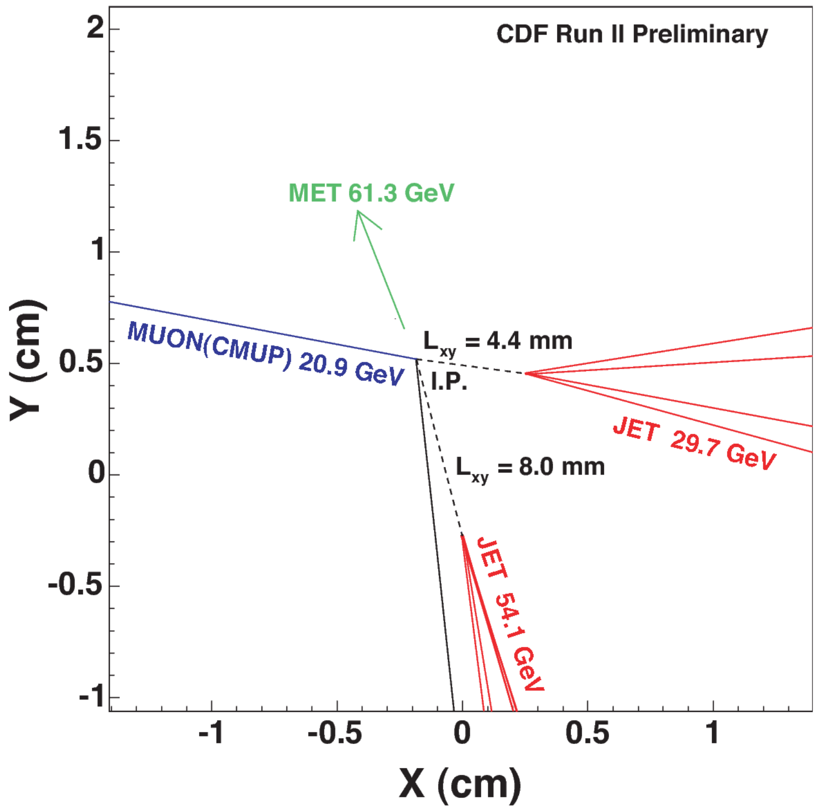


Figure 4.8: Example of a W+ 2 jets vertex display with two SecVtx tags. Muon, $\not\!\!E_T$, prompt track (black) and displaced tracks (red) are showned.

Tagging Scale Factor

The performance of the scale factor is evaluated on the basis of its $efficiency$ (ϵ), the number of correctly identified b-hadrons over the whole number of produced b-hadrons, and its *purity*, the rate of falsely identified b-hadrons in a sample with no true b-hadrons.

At CDF Monte Carlo samples are used to measure the tagger efficiency and then the result is compared to data, because Monte Carlo does not describe it perfectly. So a Scale Factor (SF) was introduced to correct for those differences.

$$SF = \epsilon_{data}/\epsilon_{MC}$$

The method used to calculate the SF is described in [15].

In this analisys we used the Tight SecVtx algorithm , which has a SF of 0.96 $\pm$ 0.05, updated and averaged on the whole Run II dataset [14].

4.4.2 Neutrino Identification

Neutrinos cannot be detected directly with the CDF detector. However, an indirect way to account for the escaping neutrinos is to measure the momentum unbalance in the detector.

The momentum unbalance can only be measured in the transverse plane. The total event vector momentum in the transverse plane is initially taken to be zero. The energy measured in the calorimeter cells is taken as a measure of the momentum carried by the interacting particles. We can thus define the transverse "missing energy" as

$$\vec{\not{E}}_T = -\sum_i E_T^i \hat{n}_i, \tag{4.10}$$

where E_T^i is the transverse energy measured in the i-th tower of the calorimeter and $\hat{n}_i$ is the projection of the versor pointing from the event vertex to the i-th calorimeter tower onto the plane perpendicular to the beam axis. $\not{E}_T$ is also corrected for the number of vertices in the event, for the presence of muons (that only leave a tiny fraction of their energy in calorimeters). Finally corrections applied to jet cone are also used to correct $\not{E}_T$. In the following, unless indicated by the subscript "raw", we will mean a $\not{E}_T$ corrected for all effects.

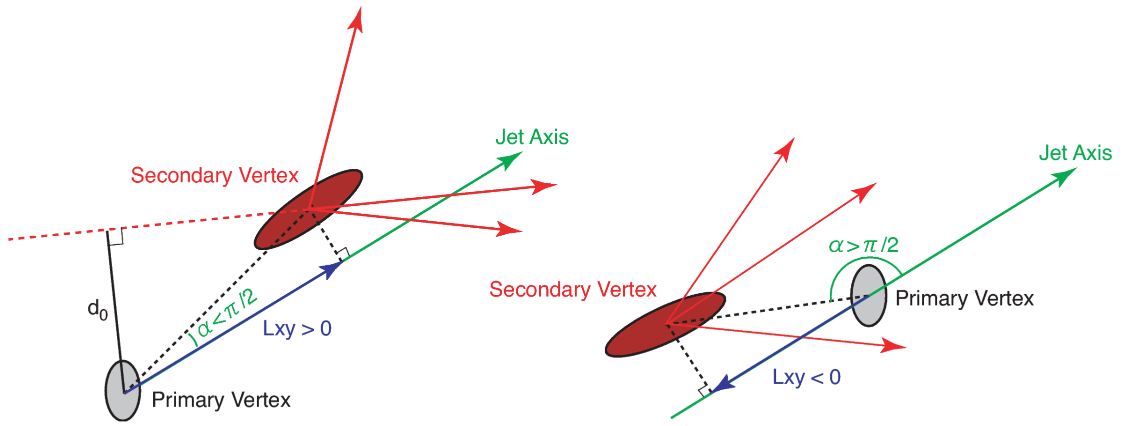


Figure 4.9: Right: true reconstructed secondary vertex (positive tag). Right: negative SecVtx tag, falsely reconstructed secondary vertex.

Chapter 5

Event Selection

In order to measure R we first need to isolate a sample enriched in events in which $t\bar{t}$ decay into $W^+qW^-\bar{q} \rightarrow \ell^+\ell^-\nu\bar{\nu}jj$. Our signal sample will be made of events with a final state containing 2 leptons, missing transverse energy and 2 jets. This chapter describes how we choose the selection requirements and the final data sample composition.

5.1 Monte Carlo Data Sets

Our Monte Carlo samples are produced by several generators, then are processed by the CDF simulation package, based on GEANT 4, that provides an output of the same format of real data. In this analysis we use the following samples:

- $t\bar{t}$: generated by PYTHIA [45] using a top mass of 172.5 GeV/c^2;
- Drell Yan / Z ($\rightarrow ee, \mu\mu, \tau\tau$) : simulated using ALPGEN [17] and showered by PYTHIA;
- WW, WZ and ZZ: generated and showered by PYTHIA;
- W+γ($W \rightarrow e\nu$ and $W \rightarrow \mu\nu$) : generated with BAUR.

Since part of the muon detector (CMX) was not fully operational before run 150145, our dataset starts from this run. As we use the silicon detector to measure secondary vertices originated by b-hadrons, we use the so called “silicon good run list”. Runs belonging to this list are the ones in which the silicon detector was fully operational. For comparison with other CDF analyses we will also use different GRL whenever needed.

5.2 Data Samples

The signal in this analysis is $t\bar{t} \rightarrow W^+qW^-\bar{q} \rightarrow \ell^+\ell^-\nu\bar{\nu}jj$. Therefore our final state of interest has two high p_T charged leptons (e or μ), two jets, and large missing transverse energy $\not\!\!E_T$ due to the presence of two escaping neutrinos. We consider as charged lepton in the final state just electrons and muons.

Events are collected using the 3-level trigger system with the following paths:

- ELECTRON_ CENTRAL18: selects events with an electron with $E_T \geq 18$GeV in the central calorimeter region (**bhel** offline dataset);
- MUON_ CMUP18: selects events with a muon with $p_T \geq 18$ GeV/c in the central muon region (**bhmu** offline dataset);
- MUON_ CMX18: requires a muon with hits in the CMX muon chambers and $p_T \geq 18$GeV/c (**bhmu** offline dataset);

Since in our sample both leptons could have been triggering, we pay special attention to dataset overlaps, considering every event just once. All samples are, therefore, orthogonal by construction.

5.3 Selected Sample Composition

In our analysis the sample is composed by the $t\bar{t}$ signal and the following backgrounds: ElectroWeak (EWK) processes and fakes.

Electroweak processes: this background is due to contribution coming from known electroweak processes with well predicted production cross sections and branching fractions:

- Diboson processes: WW, WZ, ZZ. WW can decay leptonically and produce jets from initial or final state radiation. In WZ the Z can decay in two charged lepton and the W hadronically. In ZZ, a Z boson can decay hadronically, while the other decays leponically
- Wγ: this contamination comes from events where the W decays leptonically(W$\rightarrow$ $e\nu_e$ or W$\rightarrow$ $\mu\nu_\mu$) and the photon converts, providing the second lepton.
- Drell-Yan/Z : DY/Z(DY $\rightarrow e^-e^+, DY \rightarrow \mu^-\mu^+, DY \rightarrow \tau^-\tau^+$) gives the largest background contribution. This process has naturally 2 leptons, jets from radiation and $\not\!\!E_T$ from energy mismeasurements.

In the second category, **Fakes** we collect events where one lepton is a fake. Therefore we include background contamination from detector model effects, like fake $\not\!\!E_T$ and lepton contamination from jets. Those effects, although tiny, have impact as they affect processes with large (with respect to signal) yield. For example W+jets processes, where the W provides a real lepton and $\not\!\!E_T$, and a jet can be reconstructed as the second lepton; or DY/Z+jets, with two real leptons and $\not\!\!E_T$ from energy mismeasurement.

The first step of our background estimate is to calculate the Standard Model processes contribution using the Monte Carlo simulation. Then we use data, defining appropriate control samples to estimate the fakes.

5.4 Data reduction

We purify the inclusive lepton samples collected by online system, by applying further requirements offline:

- Lepton: two and only two *tight* leptons: CEM or PHX electrons with $E_T \geq 20$ GeV, CMUP or CMX muons with $p_T \geq 20$ GeV/c. We apply the standard requirements for lepton categories and the *isolation cut*: in a cone of $R < 0.4$ around the lepton direction, the E_T not assigned to the lepton must be $< 10\%$ of the lepton E_T. The PHX electrons are accepted in events as second lepton.

- Missing Transverse energy: since there are two neutrinos from W decays in the final state, we expect to have large missing transverse energy. Thus we require $\not\!\!E_T \geq 25$ GeV, after correction for the presence of muons and for the jet energy scale. This requirement reduces the Drell Yan/Z background.

- Jets: reconstructed with JETCLU cone algorithm with a cone of $R = 0.4$, jets are required to have $E_T \geq 15$ GeV and $|\eta| \leq 2.4$. We use energy correction up to level 5, that includes effects of the η-depending response of the calorimeter and of multiple $\mathrm{p\bar{p}}$ interactions. To increase our statistics we require at least two tight jets, so we accept also events with 3 or more jets.

- B-tagging: the SECVTX algorithm is used to tag heavy flavor jets. Since tagging efficiency in data and MC is different, we apply a Scale Factor (SF) of 0.96 ± 0.05 to MC events.

- Z-cut: we require that the event primary vertex $|z_0| \leq 60$ cm with respect to the nominal center of the detector.

- Cosmic ray veto: applied only to data, it ensures that the detected muons are coming from the interaction point and not from cosmic rays coming from outside the detector.

- Conversion veto: we remove events with electrons flagged as coming from a photon conversion.

- Opposite charge veto: we require leptons to have opposite charge.

Other topological cuts are applied to improve the signal purity by rejecting the non $t\bar{t}$ events:

- H_T: we require $\mathrm{H}_T \geq 200$ GeV to remove events with initial state less energetic then $t\bar{t}$ (this partially reduce for example the W+jets contamination).

- Dilepton invariant mass: we require a dilepton invariant mass ≥ 5 GeV/c^2, as simulations does not properly model this region.

- L-cut: we require the event $\not\!\!E_T$ to be ≥ 50 GeV if there is any lepton or jet object ($\circ$) with $\triangle\phi(\circ, \not\!\!E_T) \leq 20°$. This rejects Z$\rightarrow \tau\tau$ and events with fake $\not\!\!E_T$, due to jets pointing to cracks in the calorimeter.

- Z-veto: we apply a requirement to ee or $\mu\mu$ events with dilepton invariant mass in the Z mass window [76 GeV/c^2;106 GeV/c^2]. This rejects Z events, which is the main background to our signal. Since Z events with two charged leptons have no neutrinos, we do not expect large $\not\!\!E_T$, but just $\not\!\!E_T$ from energy mismeasuring. We apply a cut on the significance of the missing trasverse energy $MetSig$, defined as:

$$MetSig = \frac{\not\!\!E_T}{\sqrt{E_T^{sum}}} \geq 4GeV^{1/2}. \tag{5.1}$$

  where E_T^{sum} is the sum of the raw transverse energy released in the calorimeters, corrected for the presence of muons and calculated considering level 4 corrected energy for tight jets. This requirement reduces the background by $\sim 25\%$ (mostly Drell Yan events), while the signal is reduced by $\sim$ %5.[18].

5.5 MC Based Estimate of Signal and Background

In this class falls every process whose normalization, cross section and kinematic can be precisely determined by Monte Carlo simulation. Those processes are: Wγ, Dibosons, $DY/Z \to \tau\tau$ and $t\bar{t}$, our signal.

They are all present in our selected data sample as they provide a final state made of 2 leptons, 2 or more jets and $\not{E}_T$. The number of events is calculated using the producion cross sections, the integrated luminosity and the selection acceptance.

In detail for a given dilepton category:

$$N^i_{\mathrm{p\bar{p}}\to X} = \sigma_{\mathrm{p\bar{p}}\to X}\ \epsilon^i\ C_{\ell_1\ell_2}\ L^i \tag{5.2}$$

where:

- X is the given considered physics process;
- i: is the given dilepton category. Using just CEM, CMUP, CMX leptons as trigger leptons (PHX ones as possible second lepton), our categories are: CEM_ CEM, CEM_ PHX, CMUP_ CMUP, CMUP_ CMX, CMX_ CMX, CEM_ CMUP, CEM_ CMX, PHX_ CMUP, PHX_ CMX;
- $\sigma_{\mathrm{p\bar{p}}\to X}$ is the production cross section. In Tab. 5.1 we list the cross sections used in this analysis. $W\gamma$ is normalized to NLO calculations by applying a correction factor to the MC production LO cross section;
- L^i is the total luminosity for the given dilepton category;
- ϵ^i is total acceptance of our sample for the given physics process: it accounts for the event detection efficiency obtained from MC events in the pretag sample and Branching Ratio fractions in the i-th channel;
- $C_{\ell_1\ell_2}$ is a factor that accounts for the z_0 scale factor[1], the trigger efficiency (ϵ_{trig}) and the lepton identification scale factor (SF) for the two leptons $\ell_1\ell_2$ of the given i-th categories.

Those factors are calculated using the following factorization:

$$C_{\ell_1\ell_2} = \epsilon_{z_0} \cdot (\epsilon_{trig_1} + \epsilon_{trig_2} - \epsilon_{trig_1}\epsilon_{trig_2}) \cdot SF_1 SF_2 \tag{5.3}$$

The lepton identification SF is measured as the ratio of the data over the MC lepton identification and muon reconstruction efficiency. The value of trigger electron efficiency is a function of the event run number, the lepton E$_T$ and its η, to take into account trigger operations and turn on[13]. The efficiency and scale factor values, averaged over all data periods [19], are listed in Tab.5.2.

5.5.1 Jet Correction Factor

Since Monte Carlo does not model the jet multiplicity spectrum properly for jets from initial and final state radiation (ISR and FSR), we correct the fraction of Monte Carlo events with no prompt jets (WW,Wγ,DY$\to \tau\tau$) reconstructed with 2 or more jets. These factors (Njet) are computed for

[1]this factor accounts for the efficiency of the $|z_0 \leq 60$ cm cut.

Process	Cross section (σ pb)	NLO Factor
WW	11.34 $\pm$ 0.68	
WZ	3.47 $\pm$ 0.21	
ZZ	3.62 $\pm$ 0.22	
W($\rightarrow e\nu$)γ	32 $\pm$ 3.2	1.36
W($\rightarrow \mu\nu$)γ	32 $\pm$ 3.2	1.34
$t\bar{t}$	7.40 $\pm$ 0.49	
DY($\rightarrow \tau\tau$)	257 $\pm$ 16	

Table 5.1: Production cross section for Standard Model processes.

ϵ_{z_0}	0.9747 $\pm$ 0.0002
Trigger Efficiency	
CEM	0.982 $\pm$ 0.003
CMUP	0.866 $\pm$ 0.009
CMX	0.873 $\pm$ 0.001
Lepton ID SF	
CEM	0.973 $\pm$ 0.005
PHX	0.908 $\pm$ 0.008
CMUP	0.868 $\pm$ 0.008
CMX	0.940 $\pm$ 0.009
$C_{\ell_1 \ell_2}$	
CEM_ CEM	0.9225 $\pm$ 0.0095
CEM_ PHX	0.8456 $\pm$ 0.0090
CMUP_ CMUP	0.7212 $\pm$ 0.0134
CMUP_ CMX	0.7817 $\pm$ 0.0104
CMX_ CMX	0.8474 $\pm$ 0.0162
CEM_ CMUP	0.8212 $\pm$ 0.0087
CEM_ CMX	0.8894 $\pm$ 0.0097
PHX_ CMUP	0.6653 $\pm$ 0.0109
PHX_ CMX	0.7263 $\pm$ 0.0095

Table 5.2: Trigger efficiency, Lepton ID Scale factor and $C_{\ell_1 \ell_2}$ global factor for dilepton category.

WW and Wγ, which are PYTHIA Monte Carlo samples, by comparing data and MC prediction for the jet multiplicity of ee and $\mu\mu$ in the Z mass window [20].

For W$\rightarrow$ $\tau\tau$ we compare data and background events in the Z mass region using ALPGEN+PYTHIA Monte Carlo Sample. Those factors are reported with their statistical uncertainty in Tab.5.3. A systematic uncertainty of 5% is taken.

For the Wγ background, given that the second lepton comes from photon conversion, we apply a conversion inefficiency scale factor of 1.147 $\pm$0.11 for W$\gamma \rightarrow e\nu\gamma$ and 1.161 $\pm$0.12 for W$\gamma \rightarrow \mu\nu\gamma$.

Njets correction factor for WW, Wγ		
ee	$\mu\mu$	$e\mu$
1.032 ± 0.009	1.057 ± 0.011	1.042 ± 0.008
Njets correction factor for W$\rightarrow \tau\tau$		
ee	$\mu\mu$	$e\mu$
1.099 ± 0.009	1.047 ± 0.012	1.076 ± 0.007

Table 5.3: Njet correction factor for MC sample.

5.6 Drell-Yan background

Drell Yan is one of the largest background in our pretag sample, as it provides naturally two leptons and jets from radiation. Fake $\not\!\!E_T$ can be produced by an energy mismeasurement. Detector effects are difficult to model, therefore we use a data-driven approach. We select a data region enriched in DY and orthogonal to our signal region. We first remove non DY/Z events using a combination of MC calculations (for known processes) and a data driven fakes estimate. Using this event yield and Drell Yan simulations, we extrapolate the DY contribution into the signal region.

For clarity let us divide the total DY/Z contribution in two parts: one in the Z mass peak and one outside the Z mass region.

$$N_{DY} = N_{OUT} + N_{IN}$$

- N$_{OUT}$: outside the Z mass peak[76 GeV/c^2; 106 GeV/c^2], the requirements applied for the signal selection are the $\not\!\!E_T$ cut, the L-Cut and the H$_T$ cut. Therefore in order to obtain a region orthogonal to this one, we choose the Z mass region without any Z veto cut, so we ask for events that pass the L-Cut and the $\not\!\!E_T$ cut with dilepton invariant mass in the Z mass region ($N^{DT}_{\not E_T}$).

  To estimate the number of DY events that pass $\not\!\!E_T$ cut and L-Cut inside the Z-peak, we correct for the presence of non DY events ($N^{BKG}_{\not E_T}$) applying the same selection. As background (BKG) *here* we consider the background to DY processes, so $t\bar{t}$ ($\sigma = 7.4 \pm 0.49$ pb), dibosons, Wγ, DY$\rightarrow \tau\tau$ and fake leptons.

  Then, we use DY Monte Carlo simulation to get the fraction ($R^{OUT/IN}_{\not E_T}$) of DY events passing $\not\!\!E_T$ cut and L-Cut, outside the Z-peak over the ones inside the Z region. Now the DY events outside the Z mass region will be the fraction from MC normalized to number of DY found in the Z region in data ($N^{DT}_{\not E_T}$).

$$N_{OUT} = R^{OUT/IN}_{\not E_T}(N^{DT}_{\not E_T} - N^{BKG}_{\not E_T})$$

- N$_{IN}$: the $t\bar{t}$ signal selection in the Z mass region in addition to the other standard requirement, requires also the Z veto. To obtain a region orthogonal to this one we invert this cut.

  Our DY enriched sample is made of events with dilepton mass in the Z mass region that fail the MetSig cut (N^{DT}_{Low}), i.e have low MetSignificance. We remove, like before, events due to non DY processes that fail the Zveto N^{BCK}_{Low}. Then using DY Monte Carlo simulation, we find the fraction ($R^{High/Low}_{\not E_T}$) of DY events in the Z mass region, that pass all cuts (which have high MetSignificance) over the ones failing the MetSignificance requirement, but passing all other $\not\!\!E_T$ cut and L-Cut. To obtain our estimate in the $t\bar{t}$ signal region, in which this cut is applied, we normalize this fraction to the number of DY events found in the DY enriched region:

$$N_{IN} = R^{High/Low}(N^{DT}_{Low} - N^{BKG}_{Low}).$$

The values of those ratios are listed in Tab. 5.4.

Then we multiply both N_{IN} and N_{OUT} for the H_T requirement efficiency estimated in both cases from MC, depending on the dilepton flavor, calculated using Z Alpgen MC.

After $\not\!\!E_T$ and L-Cut		
	ee	$\mu\mu$
$R^{OUT/IN}$	0.18 $\pm$ 0.02	0.19 $\pm$ 0.02
ϵ_{H_T}	0.40 $\pm$ 0.03	0.55 $\pm$ 0.04
$R^{High/Low}$	0.0112 $\pm$ 0.0010	0.0102 $\pm$ 0.0015
ϵ_{H_T}	0.93 $\pm$ 0.02	0.98 $\pm$ 0.03

Table 5.4: Ratios and H_T efficiency from MC.

We also estimate the contribution of DY/Z$\rightarrow \mu\mu$ events to the $e\mu$ category. From Monte Carlo studies, we know that those are events in which a very energetic photon from final state radiation is almost collinear with a muon, thus it releases energy in the electromagnetic calorimeter. In this way the muon track is associated with the calorimeter cluster and reconstructed as an electron. Since determining a data control region for this effect is not possible, we use the MC prediction.

5.7 Fake Lepton Background

Jets with a large electromagnetic calorimeter cluster and a single high p_T track can fake electrons. The muon can be faked if in the jet an isolated pion or kaon is produced within the shower. Also semileptonic decay of b or c hadron can generate real leptons.

These jets faking a lepton can produce a final state with two leptons in several physics processes, the most important in our case is the associated production of W+ jets.

In order to estimate this source of background we follow the procedure outlined in [20]. In W+ jets events, we start by defining a fakeable jet, i.e. a jet that has the possibility to fake an electron. Then we measure the fake rate, i.e. the probability that a fakeable jet actually fakes a lepton. Then our fake background estimate will be found by weighting W+ jets events with fakeable jets using the fakeable probability.

A fakeable object is defined as a CDF_EM object with HAD/EM energy deposition <0.125 or a CDF Muon with $E/p < 1$. To remove real lepton contamination from this sample we ask the fakeable to fail at least one of the following lepton ID cuts (Tab.5.5).

The fake rate is the probability for a fakeable object to fake a lepton, i.e. to be reconstructed as a good lepton. To estimate this fake rate an indipendent multijet (QCD) sample is used. This sample is made of events that pass the JET50 trigger, so events with at least a jet with $E_T > 50$ GeV. The reconstructed leptons in that sample could be real lepton from Dibosons, DY, $t\bar{t}$, so we use MC estimates to remove this contribution from the QCD sample. The energy threshold for the most energetic jet is increased to 55 GeV to mimic the trigger turn-on of the Jet50 Trigger

List of ID Variable	
CEM	HAD/EM , Strip Chi2, LShr, CEM DeltaX, CES DeltaZ
PHX	HAD/EM , Pem3x3FitTow, Pem3x3Chi2, preProfileRatio5by9U, preProfileRatio5by9V
CMUP, CMX	Stubb, HAD, EM, stub deltaX

Table 5.5: List of variables for lepton ID

and ensure a 90 % trigger efficiency. As this trigger is prescaled an appropriate correction factor is introduced. The real leptons from MC are calculated as:

$$RealLeptons_{MC} = \frac{Acc_{MC}\ \sigma_{MC}\ L}{EffectivePreScale}$$

where Acc is the MC acceptance for events with a given lepton type requiring at least a jet of E_T >50 GeV, L is the luminosity of the QCD sample and the EffectivePreScale is the pre-scale factor averaged over the runs. In the same way we can also define the contribution to fakeable object from MC simulated processes.

Then the Fake Rate is defined:

$$FakeRate = \frac{GoodLepton - RealLepton_{MC}}{Fakeable - Fakeable_{MC}}$$

where the Good Lepton and Fakeable are the objects reconstructed in the QCD data sample, while RealLepton_{MC} and Fakeable_{MC} are the contribution from other known processes.

The fakeable are assigned to a lepton category depending on the detector region they point to and the ID requirements they fail.

We estimate the systematics on the fake rate, by varying the QCD sample used in the procedure, using JET20, JET70 and JET100 samples. At the end we assume a 30% systematics to account for the differences.

Values used for the different QCD samples are listed in Tab. 5.6.

	jet20	jet50	jet70	jet100
luminosity(pb^{-1})	9027.12	9028.68	9028.11	9028.76
Effective Pre-Scale	2341.02	81.02	8	1
Jet E_T threshold cut	35 GeV	55 GeV	75 GeV	105 GeV

Table 5.6: Parameter for the QCD jet Samples

The fake rates obtained are listed in Tab.5.7. To estimate the background contribution in our signal region due to fakes we search for events in our data that have a real lepton and fakeable jets, and pass all the selection cuts, except the ID cut on the second lepton. Then we weight every fakeable jet for its fake rate and use it as the second lepton in the event. Tight jets in a cone of R < 0.4 from the fakeable are removed.

The uncertainty of background has a statistical component, which is the sum of the fake rate uncertainty and the statistical uncertainty of the $lepton + fakeable$ sample. When summing together all the contribution for different dilepton categories, we consider 30% systematics uncorrelated for ee and $\mu\mu$ categories. They are considered fully correlated and then added linearly for the $e\mu$ categories .

Jet50 Fake rate period 0 - 38						
Fake lepton	[20;30]GeV/c	[30;40]GeV/c	[40;60]GeV/c	[60;100]GeV/c	[100;200]GeV/c	p_T >200 GeV/c
CEM	0574 ± 0.0004	0.0401 ± 0.0004	0.0265 ± 0.000	0.285 ± 0.0001	0.0336 ± ± 0.0005	0.0505 ± 0.0156
PHX	0.0072 ± 0.0004	0.1288 ± 0.0006	0.1331 ± 0.0001	0.2113 ± 0.0001	0.2733 ± 0.0010	0.3759 ± 0.3049
CMUP	0.0703 ± 0.0020	0.1101 0± 0.0042	0.1184 ± 0.0036	0.1723 ± 0.0049	0.4130 ± 0.0329	0.000 ± 0.000
CMX	0.1336 ± 0.0051	0.2809 ± 0.0119	0.3366 ± 0.0113	0.3581 ± 0.0163	0.41882 ± 0.0844	0.000 ± 0.000

Table 5.7: JET 50 Fake Rates as a function of the fakeable p_T for all data period.

5.8 Pretag Sample Estimate Summary

We have described the estimate of the yield of the processes composing our pretag sample. We divided the selected sample in 9 dilepton categories depending on the lepton in the events, CEM, PHX, CMUP or CMX, and then gathered in three dilepton flavor groups (ee, $\mu\mu$ and $e\mu$). In Tab.6.3, we give the final results of number of DIL candidates events versus the background and the Standard Model $t\bar{t}$ signal predictions for the full 8.7 fb^{-1} samples. The quoted uncertainties are the sum of the statistical and sistematic [2].

Number of pretag events passing the full selection				
Pretag Sample	ee	$e\mu$	$\mu\mu$	$\ell\ell$
WW	4.46 ± 0.53	6.37 ± 0.72	1.96 ± 0.27	12.79 ± 1.54
WZ	2.33 ± 0.26	0.95 ± 0.12	0.74 ± 0.10	4.02 ± 0.42
ZZ	1.41 ± 0.17	0.350163 ± 0.06	0.62 ± 0.08	2.38 ± 0.26
Wγ	0.70 ± 0.70	0 ± 0	0 ± 0	0.70 ± 0.70
DY→ $\tau\tau$	3.70 ± 0.56	5.11 ± 0.68	2.18 ± 0.36	10.99 ± 1.33
DY→ ee, $\mu\mu$	15.40 ± 3.18	2.52± 0.39	5.53 ± 1.69	23.45± 4.86
Fakes	7.16 ± 2.2 9	13.92 ± 4.49	6.05 ± 2.19	27.14± 8.43
Total background	35.17 ± 6.60	29.22± 5.99	17.08 ± 4.10	81.47 ± 16.16
tt(sigma=7.40 pb)	*79.69± 8.25*	*127.7 ± 12.47*	*40.56± 4.24*	*247.95± 25.65*
Total SM Prediction	114.86 ± 14.70	156.92 ± 18.28	57.64 ± 8.16	329.41 ± 40.87
DATA	106	161	51	318

Table 5.8: Event summary for the 8.7 fb^{-1} inclusive DIL sample. It is shown the number of background, SM expectation and data candidate events, divided by lepton flavor.

5.9 Tagged Sample Composition

In this section we describe the sample composition after b-tagging. We use Monte Carlo simulation as it is done also in other top-physics analysis (for example in the *lepton + jets*). After the pretag selection, we require at least one jet in the event to be b-tagged by SecVtx algorithm. For Standard Model background processes we rely on Monte Carlo estimates to compute their contribution to tagged sample. So for all processes (Dibosons,Wγ, DY/Z) we consider events passing the pretag selection with b-tagged jets in Monte Carlo simulation, then we apply the *SecVtx* tagging scale factor of 0.96 ± 0.05 to take into account differences between data and MC.

For Drell-Yan component the main difference is that we are not able to generate Drell-Yan $+c\bar{c}$ events below the Z mass peak region, so we use only DY/Z + $b\bar{b}$ events to describe DY/Z + Heavy flavor jets background, rescaling the $b\bar{b}$ events to account also for $c\bar{c}$ events. This factor is 1.82 ± 0.09 [20] [22].

[2]here we are considering as sistematics, the cross section uncertainties, the 6% of uncertainty on the luminosity estimates, the errors on the $C_{\ell_1 \ell_2}$ and on the number of jet scale factors.

Events contributing to the b-tagged signal without a real b-quark ("mistag") are due to cases in which the jet is tagged but is not coming from an heavy flavour quark.

We estimate this component in the following way using the WH Analysis Method Framework(WHAM). Every jet in an event is analized and assigned a weight. If the jet is b-tagged the SecVtx Scale factor is applied, otherwise we apply the mistag matrix (this assigns each taggable jet in the event the probability to be mis-tagged, as a function of the jet variables (η, E_T)). We divide our sample by the number of b-tags, so for every event WHAM computes the probability of having the requested number of tags, with all possible jets combination, and apply it as weight to the event.

To estimates events due to fake leptons we maintained the data driven approach, and apply the fake rate to events $lepton + fakeable$ passing the whole selection and having at least a b-tagged jet.

5.10 b-tagged Sample Estimate Summary

We now describe our estimates for the composition of our b-tagged sample. The sample is divided in orthogonal dilepton categories depending on number of btagged jets (1,2) and on the lepton in the events, CEM, PHX, CMUP or CMX, and then gathered in three dilepton flavor groups (ee, $\mu\mu$ and $e\mu$). In Tab.5.9, we provide the final results of number of dilepton candidates events, together with the background and the $t\bar{t}$ signal expectations for the full 8.7 fb^{-1} data set. The quoted uncertainties are the sum of the statistical and sistematics[3].

In order to proceed with our measurement we need the background and the signal evaluated separately for 0, 1 and 2 tags. Tab.5.10 shows the background for 2 jets -1 tag sample, Tab. 5.11 shows the backgrounds for 2 jets-2 tags sample. Background for the 0 tag is obtained by subtracting from the pretag background (Tab.6.3) the 1 and 2 tag estimates. In the same way the 0 b-tag bin in data is the number of pretagged events minus tagged in each category.

Number of Tagged events passing the full selection				
	ee	$e\mu$	$\mu\mu$	$\ell\ell$
WW	0.18 ± 0.08	0.06 ± 0.051	0.13± 0.05	0.70 ± 0.13
WZ	0.04 ± 0.02	0.02 ± 0.003	0.02 ± 0.01	0.08 ± 0.02
ZZ	0.11 ± 0.03	0.015 ± 0.002	0.07 ± 0.019	0.19 ± 0.03
Wγ	0.0012 ± 0.0002	0 ± 0	0 ± 0	0.0012 ± 0.0002
DY+LF	0.81 ± 0.13	0.34± 0.06	0.62 ± 0.10	1.76 ± 0.28
DY+HF	0.45 ± 0.12	0.18 ± 0.03	0.24 ± 0.22	0.87 ± 0.19
Fakes	1.26 ± 0.51	3.16 ± 1.22	2.25± 1.06	6.67 ± 2.33
Total background	2.84 ± 0.73	3.78± 1.33	3.31 ± 1.30	10.26 ± 2.83
$t\bar{t}$ (σ=7.4 pb)	*46.77 ± 5.45*	*74.37 ± 8.28*	*23.50 ± 2.76*	*144.64 ± 16.43*
Total SM Expectation	49.61 ± 6.09	78.15 ± 9.41	26.81 ± 3.85	154.90 ± 19.02
DATA	42	73	22	137

Table 5.9: Events summary for the 8.7 fb^{-1} tagged DIL sample. It shows the number of background, SM expectation and data candidate events, divided by lepton flavor with at least one b-tag jets.

[3]here we are considering as sistematics, the cross section uncertainties, the 6% of uncertainty on the luminosity estimates, the errors on the $C_{\ell_1\ell_2}$ and on the number of jet scale factors and the error due to b-tagging.

Number of Single Tagged events passing the full selection				
	ee	$e\mu$	$\mu\mu$	$\ell\ell$
WW	0.18 ± 0.08	0.05 ± 0.05	0.12 ± 0.052	0.68± 0.13
WZ	0.03 ± 0.02	0.02 ± 0.0034	0.02 ± 0.01	0.07 ± 0.02
ZZ	0.10 ± 0.02	0.01 ± 0.002	0.06 ± 0.02	0.18 ± 0.04
Wγ	0.0012 ± 0.0002	0 ± 0	0 ± 0	0.0012 ± 0.0002
DY+LF	0.79 ± 0.13	0.34 ± 0.05	0.61 ± 0.10	1.74 ± 0.27
DY+HF	0.33 ± 0.09	0.15 ± 0.03	0.21 ± 0.08	0.69 ± 0.16
Fakes	1.06± 0.46	2.78± 1.12	1.77 ± 0.90	5.61 ± 2.06
Total background	2.50 ± 0.65	3.33 ± 1.22	2.79 ± 1.00	8.96 ± 2.49
$t\bar{t}$(σ=7.4 pb)	*35.30 ± 3.74*	*55.67 ± 5.77*	*17.91 ± 1.93*	*108.88 ± 11.33*
Total SM Expectation	37.80 ± 4.30	59.00 ± 6.72	20.70 ± 2.71	117.84 ± 13.56
DATA	30	55	17	102

Table 5.10: Events summary for the 8.7 fb^{-1} for the single tagged DIL sample. It shows the number of background, SM expectation and data candidates, divided by lepton flavor in events with at least 2 jets, but just a b-tagged one.

Number of Double Tagged events passing the full selection				
	ee	$e\mu$	$\mu\mu$	$\ell\ell$
WW	0.00076 ± 0.00023	0.02081 ± 0.00501	0.00174 ± 0.00044	0.02332 ± 0.00570
WZ	0.00587 ± 0.00572	0.00006± 0.00002	0.00009 ± 0.00002	0.00602 ± 0.00574
ZZ	0.00619 ± 0.00605	0.00032 ± 0.00007	0.01087 ± 0.00030	0.00693 ± 0.00607
Wγ	0 ± 0	0 ± 0	0 ± 0	0 ± 0
DY+LF	0.01172 ± 0.00336	0.00440 ± 0.00126	0.01088± 0.00217	0.02698 ± 0.00746
DY+HF	0.11798 ± 0.05914	0.03675 ± 0.00749	0.02163 ± 0.00754	0.17658 ± 0.06522
Fakes	0.19900± 0.11561	0.38300 ± 0.20791	0.47600 ± 0.40398	1.05800 ± 0.36619
Total background	0.34153± 0.49450	0.44534 ± 0.21587	0.52123 ± 0.49796	1.2978 ± 0.41502
$t\bar{t}$ (σ=7.4 pb)	*11.46950 ± 1.73439*	*18.70000 ± 2.56793*	*5.58676 ± 0.85638*	*35.75626 ± 5.11939*
Total SM Expectation	11.81 ± 1.82	19.15 ± 2.70	6.11 ± 1.23	37.05 ± 5.49
DATA	12	18	5	35

Table 5.11: Summary of events for the 8.7 fb^{-1} for the double tagged DIL sample. It shows the number of background, SM expectation and data candidates, divided by lepton flavor in events with at least 2 jets and 2 tags.

We report some of the kinematical distribution in the tagged data sample compared to the same distributions for the $t\bar{t}$ MC events passing our selection criteria (Fig.5.1-5.7). Jets are ordered in E$_T$ (highest E$_T$ is first).

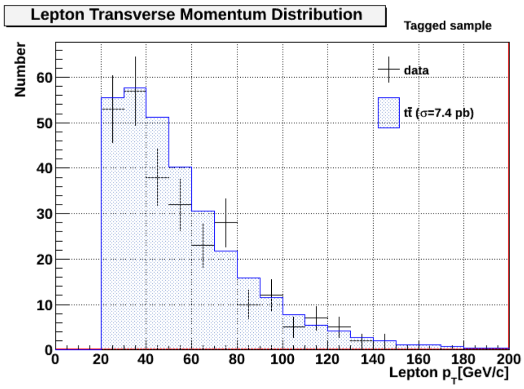


Figure 5.1: Lepton transverse momentum distribution for data (black) and $t\bar{t}$ events passing our selection requirements.

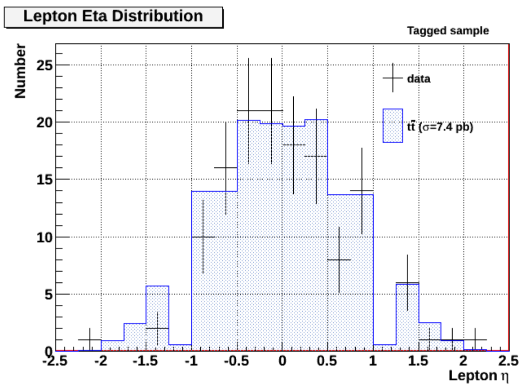


Figure 5.2: Lepton η distribution for data (black) and $t\bar{t}$ events passing our selection requirements.

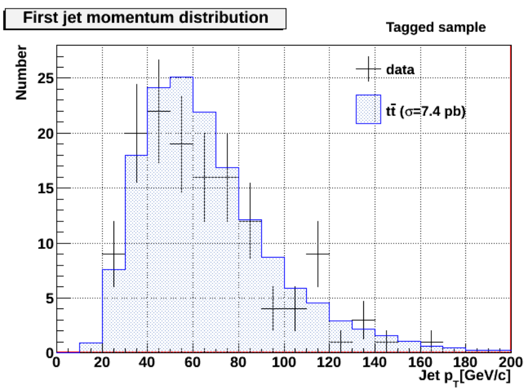


Figure 5.3: First jet transverse momentum distribution for data (black) and $t\bar{t}$ events passing our selection requirements.

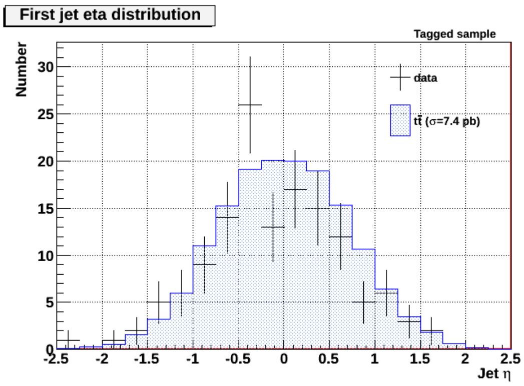


Figure 5.4: First jet η distribution for data (black) and $t\bar{t}$ events passing our selection requirements.

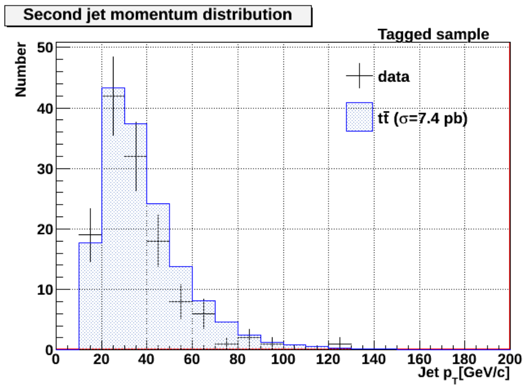


Figure 5.5: Second jet p_T distribution for data (black) and $t\bar{t}$ events passing our selection requirements.

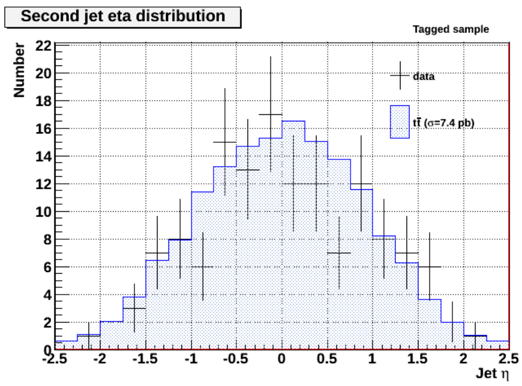


Figure 5.6: Second jet η distribution for data (black) and $t\bar{t}$ events passing our selection requirements.

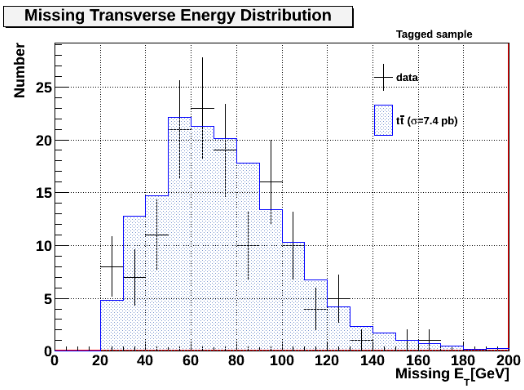


Figure 5.7: Missing Transverse Energy $\not\!\!E_T$distribution for data (black) and $t\bar{t}$ events passing our selection requirements.

Chapter 6

Analysis Strategy and Cross Section Measurement

Since our goal is the measurement of the top quark branching ratio to bottom quark, with respect to the total nuber of top decays

$$R = \frac{B(t \to Wb)}{B(t \to Wq)} = \frac{\mid V_{tb} \mid^2}{\mid V_{td} \mid^2 + \mid V_{ts} \mid^2 + \mid V_{tb} \mid^2} \tag{6.1}$$

for this analysis is crucial the number of b-jets actually found in events with respect to the number of b-jets predicted under the assumption we made for the background and $t\bar{t}$ signals. We use SecVtx algorithm, based on the reconstruction of a secondary vertex in the event, for the b-tagging, i.e. in order to identify a jet coming from b-quark fragmentation.

Due to the high purity $t\bar{t}$ signal in dilepton events, it is possible to perform a kinematic measurement of $t\bar{t}$ cross section. Our strategy is to use this result to predict the $t\bar{t}$ events in the tagged sample as a function of R . The comparison between observed events and the prediction, given by the sum of the expected $t\bar{t}$ estimate and the background, is made using a Likelihood function. Our measured value for R is the one which maximizes the Likelihood, i.e. gives the largest probability to model our data.

In order to better exploit the S/B rates for each subsample, we divide our data in 3 dileptons bins $(ee, \mu\mu, e\mu)$ and three tagging states (0,1,2). In all samples we ask for at least 2 jets.

6.1 Cross section

We first measure the $t\bar{t} \to W^+W^-\mathrm{b}\bar{\mathrm{b}} \to \ell^+\ell^- jets$ cross section applying the kinematical requirements, described in Chap.5 without looking at the b-tag content of the events.

The measured cross section is calculated as:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{A \cdot L} \tag{6.2}$$

with

$$A \cdot L = \sum_i A_i \cdot L_i \tag{6.3}$$

where

$$A_i = A_{\ell_1\ell_2} \cdot C_{\ell_1\ell_2}$$

and N_{obs} is the number of dilepton candidate events, N_{bkg} is the total number of expected background events, and the denominator is the weighted sum of the corrected acceptance for each class of dilepton events grouped by lepton reconstruction type. The corrected acceptance A_i for the given i-th dilepton category, containing the leptons $\ell_1\ell_2$, is the product of the acceptance for the $t\bar{t}$ ($A_{\ell_1\ell_2}$) simulated by PYTHIA MC and the $C_{\ell_1\ell_2}$, correction factor that takes into account the lepton trigger efficiencies and scale factors (described in Chap.5). The i-th acceptance must be multiplied also for the luminosity L appropriate for the dilepton category taken into account as there are several differences due to the different trigger paths. To calculate the luminosity for each category we require CDF subdetectors responsible of the lepton identifications and the silicon detector to be fully functional. The values for the acceptances, the correction factors and the luminosities for each category are reported in Tab.6.1. The sample composition and the backgrounds were described in Chap.5.

	Acceptance ($t\bar{t}$)		$C_{\ell_1\ell_2}$		Luminosity (pb^{-1})
Dilepton Category	Value	Uncertainty	Value	Uncertainty	Value
CEM_CEM	0.0954	0.0012	0.9225	0.0095	8763.13
CEM_PHX	0.0410	0.0008	0.8456	0.0090	8763.13
CMUP_CMUP	0.0359	0.0008	0.7212	0.0134	8755.53
CMUP_CMX	0.0381	0.0008	0.7817	0.0104	8711.31
CMX_CMX	0.0092	0.0004	0.8474	0.0162	8711.31
CEM_CMUP	0.1323	0.0014	0.8212	0.0087	8755.53
CEM_CMX	0.0669	0.0010	0.8894	0.0097	8711.31
PHX_CMUP	0.0282	0.0007	0.6653	0.0109	8755.53
PHX_CMX	0.0147	0.0005	0.7263	0.0095	8711.31

Table 6.1: List, by dilepton category, of raw acceptances $A_{\ell_1\ell_2}$, correction factor $C_{\ell_1\ell_2}$ and luminosities used to calculate the denominator for the 8.7 fb^{-1}.

In Tab. 6.2, we list the yield for data and background processes after our complete selection. In the yield uncertainty estimate, we considered the statistical contribution and the systematics due to theoretical cross section uncertainties, luminosity and the different correction factors.

PRETAG	N Event	Uncertainty
WW	12.79	1.54
WZ	4.02	0.43
ZZ	2.38	0.26
Wγ	0.70	0.71
DY$\rightarrow \tau\tau$	10.99	1.33
DY	23.45	4.86
FAKE	27.14	8.43
Data	318	

Table 6.2: List, for all dilepton categories combined, of the yields for the data and backgrounds for the 8.7 fb^{-1}. In the uncertainty estimate we considered the statistical contribution and the systematics due to theoretical cross section uncertainties, luminosity and correction factors.

Using these numbers, we measure $\sigma_{t\bar{t}}$ in the pretag sample:

$$\sigma_{p\bar{p}\rightarrow t\bar{t}} = 7.05 \pm 0.53_{stat} \pm 0.42_{lumi}\ pb. \tag{6.4}$$

As a check, we compare the cross section with the value of the official dilepton analysis of the $t\bar{t}$ production cross section in the dilepton channel at CDF [20]. In that analysis the production cross section is evaluated (both in the pretag and in the tagged signal region) with a larger lepton acceptance with respect to ours, as more lepton categories are used. That analysis uses different Good Run List with respect to ours for the pretag sample. To be consistent we do the same and in this way, we estimate the cross section for the pretag signal in the dataset corresponding to a luminosity of 9.1 fb^{-1} and the cross section for the tagged signal in a subsample of data corresponding to 8.7 fb^{-1}. For both analyses the selection criteria are the same. We report in Tab.6.3 and Tab.6.4 the yield for data and background processes, after the complete selection consistent with the official analysis.

PRETAG	N Event	Uncertainty
WW	13.10	1.60
WZ	4.12	0.45
ZZ	2.42	0.26
DY$\rightarrow \tau\tau$	11	1.5
Wγ	0.73	0.73
DY	23.6	4.90
FAKE	28.7	10.11
Data	329	

Table 6.3: List, for all dilepton categories combined, of the pretag yields for the data and backgrounds for the 9.1 fb^{-1}. In the uncertainty estimate we considered the statistical contribution and the systematics due to theoretical cross section uncertainties, luminosity and correction factors.

TAG	N Event	Uncertainty
WW	0.70	0.13
WZ	0.08	0.02
ZZ	0.19	0.03
DY+LF	1.76	0.28
Wγ	0.00	0.00
DY+HF	0.87	0.19
FAKE	6.67	2.33
Data	137	

Table 6.4: List, for all dilepton categories combined, of the tagged yields for the data and backgrounds for the 8.7 fb^{-1}. In the uncertainty estimate we considered the statistical contribution and the systematics due to theoretical cross section uncertainties, luminosity, correction factors and b-tagging.

For the pretag signal, we measure the $t\bar{t}$ production cross section:

$$\sigma_{p\bar{p}\rightarrow t\bar{t}} = 7.22 \pm 0.53_{stat} \pm 0.43_{lumi}\ pb.$$

For comparison the CDF official result is, for the pretag analysis:

$$\sigma_{p\bar{p}\rightarrow t\bar{t}} = 7.61 \pm 0.44_{stat} \pm 0.52_{syst} \pm 0.46_{lumi}\ pb.$$

For the b-tagged signal, we measure a $t\bar{t}$ production cross section:

$$\sigma_{p\bar{p}\to t\bar{t}} = 7.18 \pm 0.58_{stat} \pm 0.43_{lumi}\ pb.$$

The CDF result is, for the same signal [21]:

$$\sigma_{p\bar{p}\to t\bar{t}} = 7.09 \pm 0.49_{stat} \pm 0.52_{syst} \pm 0.42_{lumi}\ pb.$$

At the moment, the values we found for the cross section are just preliminary. Indeed we are not quoting any systematic uncertainty to the cross section, because they are still under study. The official dilepton analysis cross section values have systematic uncertainties comparable with the statistical one. Therefore as a preliminary estimate, we decided to take as total error on the cross section a conservative value of 1 pb, to account also for systematic uncertainties. We quote separately the $\sim 6\%$ luminosity contribution to the cross section uncertainty coming from luminosity measurements.

All measurements are compatibile one another within uncertainties.

However, in the following R measurement, we use the cross section value we measure (Eq.(6.4)), as already mentioned, in order to avoid any effect of the b-tagging algorithm to our $t\bar{t}$ yield and to reduce any systematics due to the use of different samples (for example due to different Good Run List requirements).

Chapter 7

Measurement of R

Since our aim is to measure R, we have to choose a physical observable in our data which is related to R. The variable we choose is the number of tagged jets in the event. In fact, for greater values of R we expect to have a larger number of b-quark in the final states, so a larger number of b-tagged jets in our data. Vice versa for smaller values of R.

We divide our data in bin of dilepton flavor $(ee, e\mu, \mu\mu)$ and in bins of tags (0, 1 and 2 tags) so we have a total of 9 independent subsamples where we can compare data to prediction. Our prediction is the total number of events we expect to have in each bin, i.e. the sum of the events due to the background and the ones due to the $t\bar{t}$ signal.

The number of $t\bar{t}$ signal expected for each bin of tags is a function of the probability for a jet to be tagged, so it is a function of R. In Fig.7.1 it is shown the dependence of the presence of b quark in the final state of $t\bar{t}$ decay on R.

For a jet coming from a b quark, generated in the quark decay, the probability to appear and be tagged is:

$$P_{TAG,b} = SF\ \epsilon_{TAG}\ R, \tag{7.1}$$

while if the top decays into a light quark, it can be tagged with probability:

$$P_{TAG,q} = \epsilon_{MISTAG}\ (1 - R), \tag{7.2}$$

where SF is the SecVtx Scale factor, $\epsilon_{MISTAG} = 0.007 \pm 0.001$ is the average probability of mis-tagging a jet [8] and $\epsilon_{TAG} = 0.413 \pm 0.023$ is the average b-tagging efficiency[1].

The probability for a b-jet of not being tagged is:

$$P_{NOTAG,b} = (1 - SF\epsilon_{TAG})R, \tag{7.3}$$

while for a light quark jet this probability is:

$$P_{NOTAG,q} = (1 - \epsilon_{MISTAG})\ (1 - R). \tag{7.4}$$

Combining the various components, the overall probability of having a given number of tags is:

$$P(0tag) = ([\epsilon_{MISTAG} - \epsilon_{TAG}SF]R + [1 - \epsilon_{MISTAG}])^2, \tag{7.5}$$

[1] We measured the value of the tagging efficiency, as the ratio of the number of tagged jets over the number of taggable jets using a $t\bar{t}$ MC sample, removing the contribution due to mis-tags.

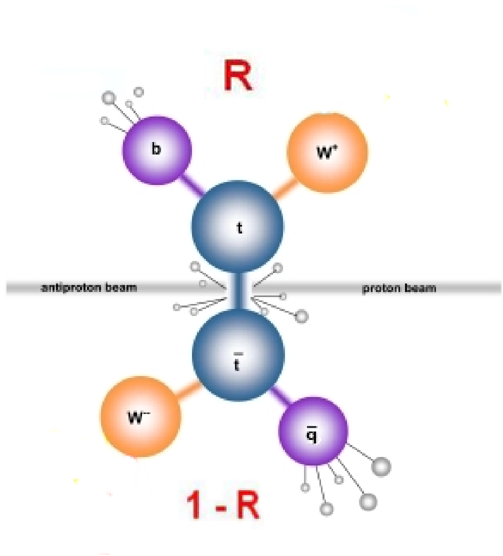


Figure 7.1: From the top quark decay the probability of a b-quark in the final state is R, while the probability of a light quark in the final state is 1-R.

$$P(1tag) = 2\cdot(R[\epsilon_{MISTAG}+\epsilon_{TAG}SF]+(\epsilon_{MISTAG}))\cdot([1-\epsilon_{MISTAG}]R+[1-\epsilon_{MISTAG}](1-R)), \quad (7.6)$$

$$P(2tags) = ([\epsilon_{TAG}SF - \epsilon_{MISTAG}]R + \epsilon_{MISTAG})^2, \quad (7.7)$$

In each i-th dilepton flavor bin, the number of $t\bar{t}$ events with j tags we expect is:

$$\lambda^{ij} = P(j)N^i_{PRETAG} = P(j)\ A^i\ L\ \sigma_{t\bar{t}}, \quad (7.8)$$

where N^i_{PRETAG} is the number of events expected for the pretag signal in the i-th dilepton flavor bin, so the sum of the expected pretag events over the dilepton categories corresponding to that particular dilepton flavor. To be explicit, A^i is the total acceptance of the dilepton flavor, calculated as the sum over the single dilepton categories, and L is the luminosity of 8.7 fb^{-1}.

For background processes, since there is no dependency upon R in our yield estimates, we just use the number of predicted events in each of the 9 categories (Tab. 5.8, Tab. 5.9). The number of events in the 0 tag bin is taken as the pretag number minus the number of events with 1 or 2 tags.

To measure R we compare our predictions to the observed data, using a likelihood function.

7.1 Likelihood function

The likelihood function L used in this analysis is:

$$L = \prod_i \wp(\mu^i_{exp}\ (R, x_j)|N_{obs}) \prod_j G(x_j|\bar{x}_j, \sigma_j). \quad (7.9)$$

In this equation $\wp(\mu^i_{exp}\ (R, x_j)|N_{obs})$ is the Poissonian probability to observe in the i-th bin N$_{obs}$ events, given the expected mean value μ_{exp}, calculated as described in the previous paragraph, as the expected yield of the $t\bar{t}$ signal and the other physical processes.

The index i runs over the 9 combinations of the three dilepton flavors and three tag bins. The Gaussian functions $G(x_i|\bar{x}_j, \sigma_j)$ are functions of the nuisance parameters x_j, with mean $\bar{x}_j$ and variance σ_j and are used to constraint the sources of systematic uncertainties. For example they describe luminosity, background estimates, selection acceptances and relevant efficiencies. In this way we can correlate uncertainties among different channels, using the same parameter for the common source of uncertainty and allowing them to vary with respect to their central values. We list our nuisance parameters in Tab.7.1.

R is left as a free parameter and we obtain it by performing a maximum likelihood fit to our model predictions given observed data. In fact it is chosen the value of R that best matches our prediction with observed data. The equivalent minimization of the -log(L) is done using MINUIT analisis package [46].

The fit from the negative logarithmic likelihood returns two results for the fitted parameter: its value and its uncertainty. Minuit gives the opportunity to choose the type of uncertainty (symmetrical or asymmetrical) to use:

- symmetric uncertainty: is obtained calculating the Hessian matrix at the minimum and inverting it. The error matrix obtained is checked to be positive-definite and takes into account the correlations between the parameters.

- asymmetric uncertainty: in order to take into account possible non- linearities introduced by the formulation of the problem. This method in general provides different positive and negative errors and the difference between the symmetric and asymmetric uncertainties is a measure of the non-linearity of the model. The errors obtained in this way are usually larger than uncertainties derived from the error matrix.

We tried both methods and found not significant variations. Therefore we are quoting just the symmetrical uncertainty.

In Tab.7.1, we list the nuisance parameter values returned by the fit.

Nuisance parameters	
ϵ_{TAG}	0.413 ± 0.023
Luminosity	8.690 ± 0.247
$\sigma_{t\bar{t}}$	7.039 ± 0.241
A_{ee}	1.207 ± 0.105
$A_{e\mu}$	2.005 ± 0.155
$A_{\mu\mu}$	0.626 ± 0.070
$B_{0,ee}$	35.215± 7.432
$B_{0,e\mu}$	32.32 ± 7.450
$B_{0,\mu\mu}$	16.539 ± 4.898
$B_{1,ee}$	2.430 ± 0.912
$B_{1,e\mu}$	3.229± 1.694
$B_{1,\mu\mu}$	2.638 ± 1.378
$B_{2,ee}$	0.346 ± 0.202
$B_{2,e\mu}$	0.453 ± 0.306
$B_{2,\mu\mu}$	0.514 ± 0.650
SF	0.95998 ± 0.024
ϵ_{MISTAG}	0.007 ± 0.002

Table 7.1: List of the nuisance parameters of the likelihood function, A are the acceptances in each dilepton flavor and B are the backgrouds estimates for the given bin (numer of tags, dilepon flavor). Values shown are the ones returned by the fit.

In Fig. 7.2, we plot the likelihood logarithm shape about its minimum.

The overall uncertainty on R is obtained as the value of which $\triangle log(L) = 0.5$. The value of R, found with this procedure is :

$$R = 0.86^{+0.064}_{-0.062} = 0.86 \pm 0.063.$$

The uncertainty on R is the combination of statistics and systematrics uncertainty.

In Tab.7.2 we list the number of observed events in data, the number of expected event with our maximum likelihood estimate of R. We quote also the expected uncertainty calculated by varying R of $\pm 1\sigma$ about its mean value.

We show (Fig.7.3-Fig.7.5) the number of events expected ($t\bar{t}$ (R=1) and background), data and predicted number of events after the fit with R=0.86 for $t\bar{t}$ signal, for the different dilepton flavor for 0, 1 and 2 tags.

For the evaluation of the systematics on R, we performed the fit, varying one by one each nuisance parameter value by a standard deviation from its mean. The most important contributions to R variation, during this procedure are reported in Tab.7.3. We report also the bidimensional

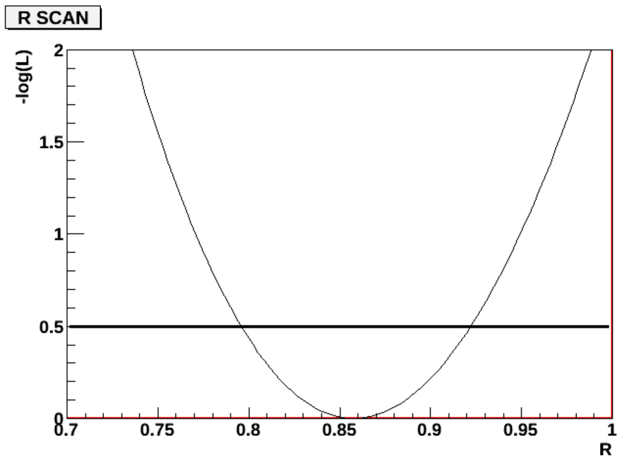


Figure 7.2: View of the logarithm likelihood distribution about its minimum.

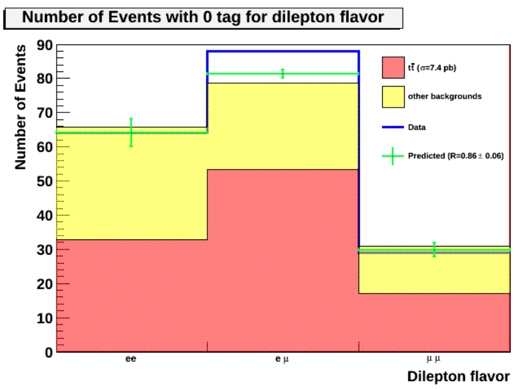


Figure 7.3: Number of events with 0 tag expected for R=1 ($t\bar{t}$ and background superimposed) and for R=0.86 (green). Events found in data are reported (blue). The bin are the dilepton flavors $ee, e\mu, \mu\mu$.

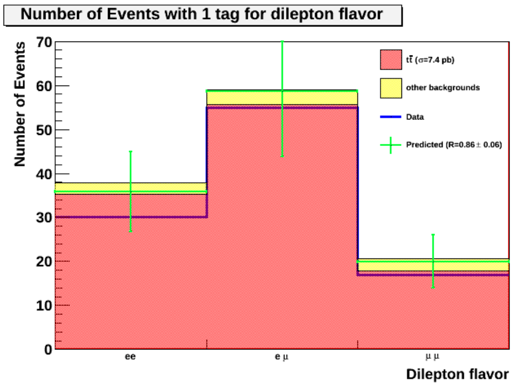


Figure 7.4: Number of events with 1 tag expected for R=1 ($t\bar{t}$ and background superimposed) and for R=0.86 (green). Events found in data are reported (blue). The bin are the dilepton flavors $ee, e\mu, \mu\mu$.

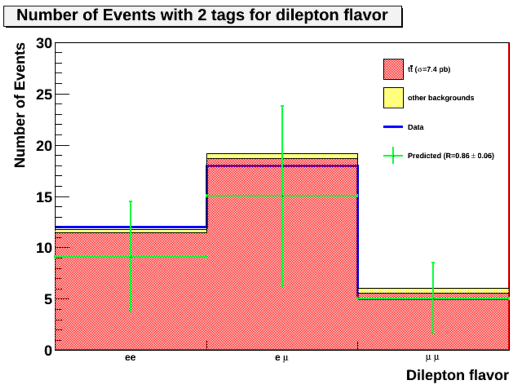


Figure 7.5: Number of events with 2 tags expected for R=1 ($t\bar{t}$ and background superimposed) and for R=0.86 (green). Events found in data are reported (blue). The bin are the dilepton flavors $ee, e\mu, \mu\mu$.

Number of Events		
Bin	Observed	Predicted
ee 0 tag	64	64.15 ± 4.05
$e\mu$ 0 tag	88	81.35 ± 1.15
$\mu\mu$ 0 tag	29	29.83 ± 2.09
ee 1 tag	30	35.79 ± 9.13
$e\mu$ 1 tag	55	58.63 ± 14.78
$\mu\mu$ 1 tag	17	19.94 ± 6.06
ee 2 tags	12	9.12 ± 5.40
$e\mu$ 2 tags	18	15.02 ± 8.77
$\mu\mu$ 2 tags	5	5.06 ± 3.47

Table 7.2: Number of observed events in data and the number of expected event with our maximum likelihood estimate of R. The uncertainty is calculated by varying R of $\pm 1\sigma$ about its mean value.

plots in the space of the parameters to see the dependence of R, within a 2σ variation, on the different parameters (Fig. 7.6 - Fig. 7.9)

Nuisance parameter	Systematic variation
ϵ_{TAG}	± 0.01
SF	+0.048, -0.044
$\sigma_{t\bar{t}}$	0.02, -0.003
Lumi	0.016, -0.014
ϵ_{MISTAG}	0.005, -0.006

Table 7.3: List of the nuisance parameters of the likelihood function that give the largest systematic contribution to our measurement.

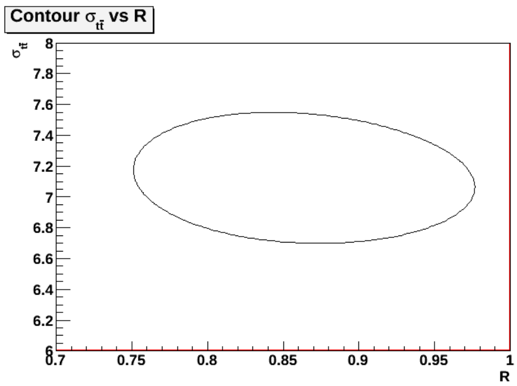


Figure 7.6: Contour plot of R as a function of the $t\bar{t}$ cross section

Fig.7.7 shows the contour plot of R as a function of the tagging efficiency. This is the most significantly correlated variable to the R parameter, with a correlation coefficient $\rho = -0.436$. Instead R is not largely correlated with the other parameters, because distributions of R are rather flat as a function of the other parameters.

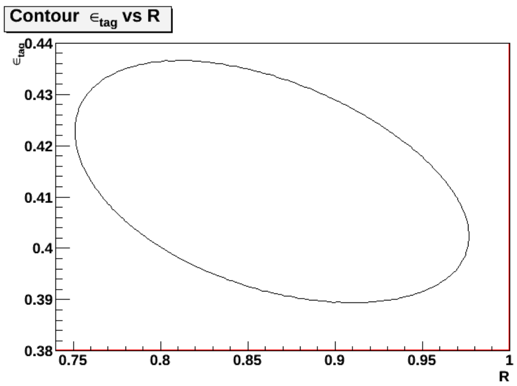


Figure 7.7: Contour plot of R as a function of the tagging efficiency.

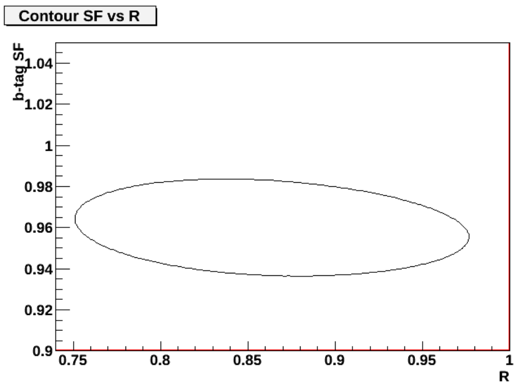


Figure 7.8: Contour plot of R as a function of the tagging Scale Factor(SF).

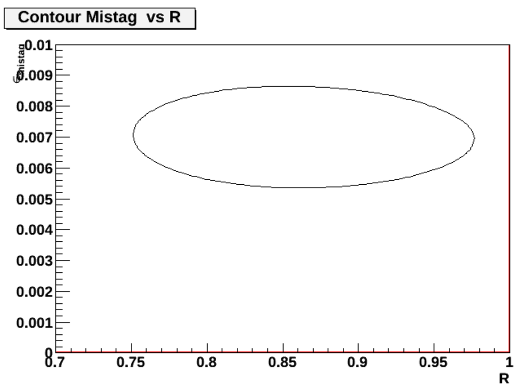


Figure 7.9: Contour plot of R as a function of the mis-tagging efficiency.

In the procedure to calculate R, we considered events with two jets in the final state in our model. A third jet in the $t\bar{t}$ sample could be produced by initial or final state radiation and could, in principle, be wrongly tagged. We estimate the third jet contribution and perform the fit again: no significant deviation of the mean value of R is found.

7.2 Conclusions

In this thesis we performed an update to the CDF measurement of the top branch ratio

$$R = \frac{\mathcal{B}(t \to Wb)}{\mathcal{B}(t \to Wq)} = \frac{\mid V_{tb} \mid^2}{\mid V_{td} \mid^2 + \mid V_{ts} \mid^2 + \mid V_{tb} \mid^2}$$

with a measurement of the top pair production cross section $\sigma_{p\bar{p}\to t\bar{t}}$ using the dilepton decay channel for the $t\bar{t}$ couple.

The work exploits several analysis tools, such as the WHAM framework for event selection and the maximum likelihood technique to compare observed data to the expectation.

As a conclusion we provide an estimate of the CKM matrix element assuming three generation of quarks and the unitarity of the CKM matrix. The results for R and $|V_{tb}|$ are summarized in Table 7.4 and are in agreement with the Standard Model, with the previous CDF measurements and with the up to date measurement of R performed by DØ .

Parameter	Our Values	CDF (1 + jets)	DØ	DØ (dilepton only)
$\sigma_{t\bar{t}}(pb)$	$7.05 \pm 0.53_{stat} \pm 0.43_{lumi}$	7.5 ± 1.0	$7.74^{+0.67}_{-0.57}$	$8.19^{+1.06}_{-0.92}$
R	0.86 ± 0.06	0.94 ± 0.09	0.90 ± 0.04	0.86 ± 0.05
$\|V_{tb}\|$	0.93 ± 0.03	0.97 ± 0.05	0.95 ± 0.02	0.94 ± 0.04

Table 7.4: Our results compared to the DØ measurements. We reported also the DØ measurement performed only in the dilepton channel. All the errors reported are combination of statistical and systematic uncertainty, if not otherwise indicated.

Bibliography

[1] M. Kobayashi and T. Maskawa *CP-Violation in the Renormalizable Theory of Weak Interaction* Progress of Theoretical Physics, 49: 652 (1973).

[2] F.Abe et al., *Evidence for Top Production in p anti-p Collisions at* $s^{1/2}$ *= 1.8 TeV* Phys. Rev. Lett. 73, 225 (1994).

[3] The Tevatron Electroweak Working Group, CDF Note 10976 (2013).

[4] D. Acosta et al., *Measurement of B(t→Wb)/B(t→Wq) at the Collider Detector at Fermilab* Phys. Rev. Lett. 95, 102002 (1995).

[5] V.M. Abazov et al., *Precision Measurement of the Ratio B(t→Wb)/B(t→Wq) and Extraction of Vtb*, Phys. Rev. Lett. 107, 121802 (2011).

[6] The CMS collaboration, CMS-PAS-12-035 (2012)

[7] Fermilab Beams Division, *The Main Injector rookie book*, 2010, http://www-bdnew.fnal.gov/operations/rookie_ books/Main_ Injector_ v1.1.pdf.

[8] The CDF Collaboration, Measurement of R = B(t→Wb)/B(t→Wq) in the SecVtx tagged lepton plus jets sample with 8.7 fb^{-1}, CDF note 10887 (2012).

[9] P.W. Higgs, *Broken Symmetries and the masses of gauge bosons*, Phys.Rev.Lett. 13 (1964)508-509.

[10] F.Englert and R.Brout, *Broken Symmetries and the masses of gauge vector mesons*, Phys.Rev.Lett. 13 (1964)321-323

[11] G.S. Guralnik et al., *Global coservation laws and masless particles*, Phys. Rev.Lett 13(1964) 585-587.

[12] H.Bachachou, CDF Note 7632 (2004).

[13] F. Sforza, *Evidence for Diboson Production in the Lepton plus Heavy Flavour Jets Final States*, PhD Thesis, Pisa University (2012).

[14] *p0-38 B-tagging Prescriptions (p18-28, 616p+ reprocessed)* http://www-cdf.fnal.gov/internal/physics/top/RunIIBtag/p38_results/ .

[15] V.Giakoumopoulou et al. *Summer 2010 SecVtx Scale Factor Calculated electron Method through Perioriod 28*, CDF Note 10178 (2010).

[16] T.Sjostrand et al., *Comput.Phys.Commun* 135, p.238 (2001).

[17] M.L.Mangano et al.,*ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP 0307:001, hep-ph/0206293 (2003).

[18] http://www-cdf.fnal.gov/ kimura/internal/DIL/ValMETsig/valmetsig.html

[19] E. Rolli et al. (CDF) (2011), http://ncdf70.fnal.gov:8001/PerfIDia/PerfIDia.html.

[20] C.S.Moon at al. *Measurement of Top Dilepton Cross Section with CDF Full data using the DIL Selection* , CDF NOTE 10877,(2013).

[21] CDF collaboration *Measurement of the* $t\bar{t}$ *production cross-section in events with two leptons and bottom-quark jets using the full CDF data set* , CDF NOTE 10972,(2013).

[22] A.Ivanov et al., CDF NOTE 9417,(2009).

[23] I. J. R. Aitchison and A. J. G. Hey. Gauge Theories in Particle Physics 2 Volume Set. Taylor & Francis, 3 edition, January 2004.
[24] R. Ellis, W.J. Stirling, and B.R. Webber QCD and Collider Physics Cambridge University Press 2003.
[25] S. Willenbrock The Standard Model and the Top quark, arXiv : hep - ph/0211067 (2002).
[26] A. Abulencia et al, Phys. Rev. D, 73, 112006 (2006).
[27] D. Acosta et al., Phys. Rev. Lett., 95, 102002 (2005).
[28] T.Affolder et al., Phis. Rev. Lett., 86, 15 (2001).
[29] M. Beneke et al. arXiv : 1112.4606v1 (2010).
[30] D.Acosta et al., Nucl.Instrum.Meth, A461, 540 (2001).
[31] C. M. Ginsburg, "CDF Run 2 muon system," *The European Physical Journal C - Particles and Fields*, vol. 33, pp. s1002–s1004, 2004. 10.1140/epjcd/s2004-03-1795-6.
[32] G. Ascoli *et al.*, "Central Muon Level-1 Trigger Electronics," FERMILAB-PUB-87-188-E.
[33] D. Mohl, "Physics and technique of stochastic cooling," *Phys. Reports*, vol. 58, 1980.
[34] I. Vila, "Performance and First Physics Results of the SVT Trigger at CDF II," *ArXiv High Energy Physics - Phenomenology e-prints*, jul 2003.
[35] B. Ashmanskas, *et al.*,"The CDF Silicon Vertex Trigger," *Nuclear Instruments and Methods in Physics Research A*, vol. 518, pp. 532–536, Feb. 2004.
[36] T. Affolder *et al.*, "CDF central outer tracker," *Nucl. Instrum. Meth.*, vol. A526, pp. 249–299, 2004.
[37] http://www-cdf.fnal.gov/cdfsim/validation/cot/cotVal.html.
[38] R. Blair *et al.*, "The CDF-II detector: Technical design report," FERMILAB-DESIGN-1996-01.
[39] P. Gatto, "Performance of the new tracking system at CDF II," Master's thesis, University of Padova, 2001.
[40] S. Menzemer, "A tracking and alignment software package for the CDF II silicon detector," Master's thesis, University of Karlsruhe, 2002.
[41] F. D. Snider, "Tracking at CDF: Algorithms and experience from Run I and Run II," *Nuclear Instruments and Methods in Physics Research A*, vol. 566, pp. 133–141, Oct. 2006.
[42] G.C. Blazey et al. *Run II Jet Physics* CDF Note 5293.
[43] F. Abe et al. (CDF Collaboration) *The Topology of Three Jet Events in* $p\overline{p}$ *Collisions at* $\sqrt{s}$ *= 1.8 TeV* Phys. Rev. D 45 (1992) 1448.
[44] A. Bhatti et al. (CDF Collaboration) *Determination of the Jet Energy Scale at the Collider Detector at Fermilab* Nucl. Instrum. Meth.A566 (2006) 375.
[45] T. Sjostrand et al. *High-Energy-Physics Event Generation with PYTHIA 6.1* Comput. Phys. Commun. 135 (2001) 238.
[46] Cern Computing and Networks Division, *MINUIT Reference Manual* http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html.

Appendix A

Instantaneous Luminosity Measurement

The Cherenkov light is detected using photomultiplier tubes. The momentum threshold for light emission in CLC is 9.3 MeV/c for electrons and 2.6 GeV/c for pions. The number of $p\bar{p}$ interactions in a bunch crossing follows a Poisson distribution with mean μ, where the probability of empty crossing is given by:

$$P_0(\mu) = e^{-\mu}$$

which is correct if the acceptance of the detector and its efficiency were 100%. In practice, there are some selection criteria, α, to define an "interaction." An interaction is defined as a $p\bar{p}$ crossing with hits above a fixed threshold on both sides of the CLC detector. Therefore, an empty crossing is a $p\bar{p}$ crossing with no interactions. Given these selection criteria, the experimental quantity P_0, called $P_0^{exp}\{\alpha\}$, is related to μ as:

$$P_0^{exp}\{\mu, \alpha\} = (e^{\epsilon_w \cdot \mu} + e^{\epsilon_e \cdot \mu} - 1) \cdot e^{-(1-\epsilon_0)\cdot\mu}$$

where the acceptances ϵ_0 and $\epsilon_{w/e}$ are, respectively, the probability to have no hits in the combined east and west CLC modules and the probability to have at least one hit exclusively in west/east CLC module. The evaluation of these parameters is based on Monte Carlo simulations, and typical values are respectively 0.07 and 0.12. From the measurement of μ we can extract the luminosity. Since the CLC is not sensitive at all to the elastic component of the $p\bar{p}$ scattering, the rate of inelastic $p\bar{p}$ interactions is given by:

$$\mu \cdot f_{bc} = \sigma_{in} \cdot \mathcal{L}$$

where f_{bc} is the bunch-crossing frequency at the Tevatron and σ_{in} is the inelastic $p\bar{p}$ cross section. $\sigma_{in} = 60.7 \pm 2.0$ mb is obtained by extrapolating the combined results for the inelastic $p\bar{p}$ cross section of CDF at $\sqrt{s} = 1.8$ TeV and E811 measurements to $\sqrt{s} = 1.96$ TeV. Different sources of uncertainties are taken into account to evaluate the systematic uncertainties on the luminosity measurement. The dominant contributions are related to the detector simulation and the event generator used, and have been evaluated to be about 3%. The total uncertainty in the CLC luminosity measurements is 5.8%, which includes uncertainties on the measurement (4.2%) and on the inelastic cross section value (4%).

Appendix B

Tracking Algorithm

Using the hit positions in the tracking system, pattern recognition algorithms reconstruct the particle original trajectory measuring the five parameters of the helix that best match to the observed hits.

CDF employs several algorithms for track reconstruction, depending on which component of the detector a particle travels through. The principal one is the Outside-In (OI) reconstruction. This algorithm, which exploits the information from both the central drift chamber and the silicon detectors, is used to track the particles in the central region ($|\eta| < 1$). It first reconstructs tracks in the `COT` and then extrapolates them inwards toward the beam.

The first step of pattern recognition in the `COT` looks for circular paths in the axial superlayers. Cells in the axial superlayers are searched for sets of 4 or more hits that can be fit to a straight line. These sets are called "segments". Once segments are found, there are two approaches to track finding [41] ("segment linking" and "histogram linking" algorithms). One approach is to link together the segments which are consistent with lying tangent to a common circle. The other approach is to constrain its circular fit to the "beam-line" (see Sec. 4.1.1). Once a circular path is found in the r-ϕ plane, segments and hits in the stereo superlayers are added depending on their proximity to the circular fit, producing a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not. The outcone is a high track reconstruction efficiency in the `COT` for tracks passing through all 8 superlayers (97% for tracks with $p_T > 10\,GeV/c$) [1].

Once a track is reconstructed in the `COT`, it is extrapolated inward to the silicon system. Based on the estimated errors on the track parameters, a three dimensional "road" is formed around the extrapolated track. Starting from the outermost layer, and working inwards, silicon hits found inside the road are added to the track. As hits are added, the road gets narrowed according to the knowledge of the updated track parameters and their covariance matrix. A reduction of the road width decreases the chance of adding wrong hits to the track, and also reduces the computation time. In the first pass of this algorithm, axial hits are added. In the second pass, hits with stereo information are added to the track. At the end, the track combination with the highest number of hits and lowest χ^2/ndf for the five parameters helix fit is kept.

[1] The track reconstruction efficiency mostly depends on how many tracks are reconstructed in the event. If there are many tracks close to each other, hits from one track can shadow hits from the other track, resulting in efficiency losses.