UNIVERSITA' DEGLI STUDI DI TRENTO

Facoltá di Scienze Matematiche, Fisiche e Naturali

DOTTORATO DI RICERCA IN FISICA - CICLO XX

Tesi di Dottorato

Measurement of the $t\bar{t}$ production cross section in the $\not\!\!E_T + jets$ channel at CDF

Coordinatore:
Prof. Renzo Vallauri

Supervisore:
Prof. Ignazio Lazzizzera

Dottorando:
Gabriele Compostella

---

6 Marzo 2008

To $\mathcal{M}.^{\mathbf{4}}$*,*
my irreplaceable source
of true happiness...

Contents

Introduction

This thesis is focused on an inclusive search of the $t\bar{t} \to \not\!\!E_T + jets$ decay channel by means of neural network tools in proton antiproton collisions at $\sqrt{s} = 1.96\ TeV$ recorded by the Collider Detector at Fermilab (CDF).

At the Tevatron $p\bar{p}$ collider top quarks are mainly produced in pairs through quark-antiquark annihilation and gluon-gluon fusion processes; in the Standard Model description, the top quark then decays to a W boson and a b quark almost 100% of the times, so that its decay signatures are classified according to the W decay modes. When only one W decays leptonically, the $t\bar{t}$ event typically contains a charged lepton, missing transverse energy due to the presence of a neutrino escaping from the detector, and four high transverse momentum jets, two of which originate from b quarks.

In this thesis we describe a $t\bar{t}$ production cross section measurement which uses data collected by a "multijet" trigger, and selects this kind of top decays by requiring a high-P_T neutrino signature and by using an optimized neural network to discriminate top quark pair production from backgrounds.

In Chapter 1, a brief review of the Standard Model of particle physics will be discussed, focusing on top quark properties and experimental signatures.

In Chapter 2 will be presented an overview of the Tevatron accelerator chain that provides $p\bar{p}$ collisions at the center-of-mass energy of $\sqrt{s} = 1.96\ TeV$, and proton and antiproton beams production procedure will be discussed.

The CDF detector and its components and subsystems used for the study of $p\bar{p}$ collisions provided by the Tevatron will be described in Chapter 3.

Chapter 4 will detail the reconstruction procedures used in CDF to detect physical objects exploiting the features of the different detector subsystems.

Chapter 5 will provide an overview of the main concepts regarding Artificial Neural Networks, one of the most important tools we will use in the analysis.

Chapter 6 will be devoted to the description of the main characteristics of the $t\bar{t} \to \not\!\!E_T + jets$ decay channel used to train our neural network to discriminate the top pair production from background processes. We will discuss the event selection method and the tecnique used for background prediction, that will rely on b-jets identification rate parameterization.

Finally, Chapter 7 will provide a description of the final data sample and a detailed discussion of the systematic uncertainties before determining the cross section measurement by means of a likelihood maximization.

Chapter 1

Theoretical Overview

Our present understanding of the fundamental constituents of matter and their interactions is expressed in a theory called the Standard Model. The Standard Model was developed during 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 Standard Model integrates two gauge theories: Quantum Chromodynamics (QCD), describing the strong interactions, and the electroweak (EW) theory of Glashow, Weinberg and Salam, 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 The Standard Model of particle physics

The Standard Model [1, 2, 3, 4] is a quantum field theory based on the gauge symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The first gauge group $SU(3)_C$ is related to the description of the strong interactions which affect quarks only and are mediated by gluons. $SU(3)_C$ defines the *Quantum Chromo Dynamics* (QCD) theory. On the other hand, $SU(2)_L \times U(1)_Y$ is the underlying symmetry which provides a theoretical description of electromagnetic and weak interactions.

According to the Standard Model there are two families of elementary particles (i.e. particles which do not have any internal structure): *fermions* (with spin 1/2) and *bosons* (with spin 1). Fermions are subject to interactions mediated through the exchange of *gauge bosons*. There are 12 elementary fermions: the 6 ones interacting by the electroweak force only are named leptons and the 6 ones interacting by both the electroweak and the strong force are named quarks. Leptons and quarks are further organized into three families, called *generations*: for each generation, particles have their corresponding anti-particles having the same properties as the partner particles but opposite charges (the *charge* of the particle is the quantum number that defines the coupling of the particle to the electroweak force carriers).

The first generation comprises the electron e^-, with electric charge $Q = -1$, its corresponding neutrino ν_e with $Q = 0$, and two types (conventionally named "flavours") of quarks, the *up* and *down*, and their corresponding antiparticles (e^+,

$\bar{\nu}_e$, $\bar{u}$ and $\bar{d}$). The up and down quarks, denoted by u and d, carry fractionary electric charges $Q_u = \frac{2}{3}$ and $Q_d = -\frac{1}{3}$ respectively. In addition to the electric charge, quarks also carry an additional quantum number related to strong interaction, the *color*, labeling the three degenerate indipendent state of the fundamental triplet (anti-triplet) of exact $SU(3)_C$ "color" simmetry in which quarks (anti-quarks) live. Since "colored" particles are not observed in nature, quarks must be confined into color-neutral composite particles, called *hadrons*, which are categorized as baryons and mesons depending on their quark composition: baryons are basically constituted by three "valence" quarks, like proton and neutron: $p \sim uud$ and $n \sim udd$. On the other hand, mesons are composed by a quark-antiquark pair, for instance pions $\pi^+ \sim u\bar{d}$ and $\pi^- \sim d\bar{u}$.

Second and third generation particles have identical properties to first generation ones, but different masses.

Interactions are mediated by gauge particles: the carrier of the electromagnetic force is the photon γ, which is massless and chargeless. The weak force is mediated by three massive vector bosons: $W^{\pm}$ and Z^0, with charge $Q = \pm 1$ and 0 respectively. The strong force among quarks is mediated by the eight gluons g_α, which are an octet of adjoint representation in color space; each gluon is massless and chargeless and has the possibility of interacting with other gluons as well as with quarks. Gravitational interactions are not part of the Standard Model framework. A spin 2 graviton boson is supposed to be the carrier of the gravitational force but has never been observed.

1.1.1 Quantum Electrodynamics

Elementary particles are spin-$\frac{1}{2}$ fermions: in absence of gauge fields their dynamics is described by the Dirac equation and the corresponding Lagrangian:

$$\mathcal{L}_{Dirac} = \bar{\Psi}(x)(i\partial_\mu \gamma^\mu - m)\Psi(x) \tag{1.1}$$

$\mathcal{L}_{Dirac}$ is invariant under the following global $U(1)$ transformation, acting on the fields and their derivatives:

$$\Psi \to e^{iQ\theta}\Psi \quad \bar{\Psi} \to e^{-iQ\theta}\bar{\Psi} \quad \partial_\mu \Psi \to e^{iQ\theta}\partial_\mu \Psi \tag{1.2}$$

It is possible to consider a local transformation of the same kind by allowing the parameter θ in Eq. 1.2 to have a dependence on the space-time point x; but by doing so the invariance of the Lagrangian in Eq. 1.1 is lost.

We can restore the invariance under local $U(1)$ transformations of the type $\Psi \to \Psi e^{iQ\theta(x)}$ if we introduce an additional boson field $A_\mu(x)$, a gauge vector associated to the photon, interacting with the field Ψ and whose transformations compensate the non-invariant terms in the Lagrangian. In this way, the $U(1)$ gauge invariant Lagrangian of Quantum Electrodynamic (QED) can be written as:

$$\mathcal{L}_{QED} = \bar{\Psi}(x)(iD_\mu \gamma^\mu - m)\Psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) \tag{1.3}$$

where we introduced the so called covariant derivative D_μ, defined as follows:

$$D_\mu \Psi = (\partial_\mu - ieQA_\mu)\Psi \tag{1.4}$$

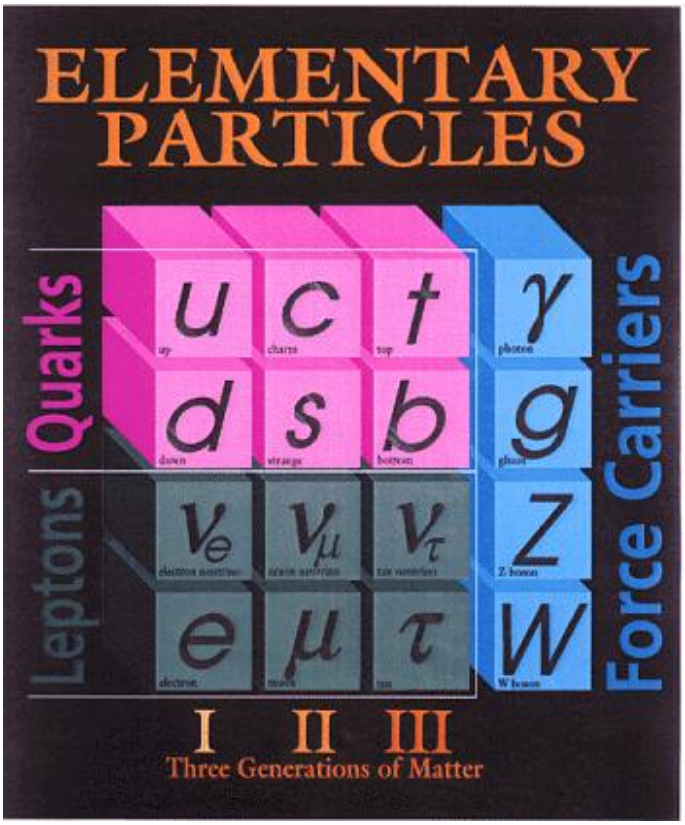


Figure 1.1: The three families of elementary particles in the Standard Model.

that contains the interaction term between the photon and fermions, and the field strength tensor $F_{\mu\nu}$, defined by:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.5}$$

in photon kinematical term of Eq. 1.3.

1.1.2 Electroweak Theory

The Electroweak theory unifies the weak isospin non-Abelian group $SU(2)_L$ acting on left-handed fermions and the weak hypercharge (Abelian) group $U(1)_Y$ in $SU(2)_L \times U(1)_Y$. Introducing the Pauli matrices σ_i with $i = 1, 2, 3$ we can write the four generators of $SU(2)_L \times U(1)_Y$ as $T_i = \frac{\sigma_i}{2}$ coming from $SU(2)_L$ and $\frac{Y}{2}$ from $U(1)_Y$. The commutation relations of the four group generators are the following:

$$[T_i, T_j] = i\epsilon_{ijk}T_k; \quad [T_i, Y] = 0; \quad i, j, k = 1, 2, 3. \tag{1.6}$$

Left-handed fermions are $SU(2)_L$ doublets:

$$f_L \to e^{i\vec{T}\vec{\theta}} f_L; \quad f_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \quad \begin{pmatrix} u_L \\ d_L \end{pmatrix}, ... \tag{1.7}$$

while right-handed fermions transform as singlets:

$$f_R \to f_R; \quad f_R = e_R, u_R, d_R, ... \tag{1.8}$$

Fermions	T	T^3	Q	Y
ν_L	$1/2$	$1/2$	0	-1
e_L	$1/2$	$-1/2$	-1	-1
e_R	0	0	-1	-2
u_L	$1/2$	$1/2$	$2/3$	$1/3$
d_L	$1/2$	$-1/2$	$-1/3$	$1/3$
u_R	0	0	$2/3$	$4/3$
d_R	0	0	$-1/3$	$-2/3$

Table 1.1: Fermion quantum numbers for the first generation in the Standard Model.

Fermions quantum numbers coming from the two groups are related to each other and to charge by the following equation:

$$Q = T_3 + \frac{Y}{2}. \tag{1.9}$$

The number of associated gauge bosons of the model is equal to the number of the symmetry group generators, so we have four bosons: W_μ^i $(i = 1, 2, 3)$ and B_μ, associated to $SU(2)_L$ and $U(1)_Y$ respectively.

In order to write the Lagrangian for the electroweak sector of the Standard Model we can follow the same procedure used previously for the Quantum Electrodynamics, building the model around the conservation of the weak isospin and weak hypercharge under local gauge transformations. We can thus change the $SU(2)_L \times U(1)_Y$ symmetry from global to local and replace the field derivatives with their corresponding covariant derivatives. For a generic fermion field f, we can define the covariant derivative as follows:

$$D_\mu f = \left(\partial_\mu - ig\vec{T}\cdot\vec{W}_\mu - g'\frac{Y}{2}B_\mu\right) f \tag{1.10}$$

where g and g' are the coupling constants associated to $SU(2)_L$ and $U(1)_Y$, respectively.

Similarly to QED the electroweak Lagrangian includes kinetic terms for the gauge fields:

$$\mathcal{L} = -\frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.11}$$

where the field strength tensors are defined as follows:

$$\begin{aligned} W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk}W_{\mu j}W_{\nu k} \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \tag{1.12}$$

where i, j, k are indeces of vector components in the adjoint representation of $SU(2)_L$.

The gauge invariant interactions and the fermion kinematics are generated by $\bar{f}iD_\mu\gamma^\mu f$ terms in the Lagrangian, while the physical gauge bosons fields $W_\mu^\pm$, Z_μ

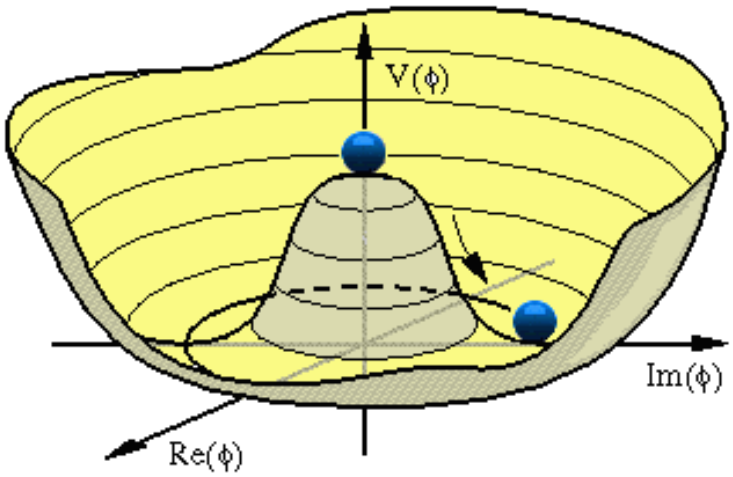


Figure 1.2: The Higgs potential, and a pictorial representation of the spontaneus symmetry breaking mechanism.

and A_μ can be obtained calculating the electroweak interaction eigenstates, that are found to be:

$$\begin{aligned} W_\mu^\pm &= \frac{W_{1\mu} \mp i W_{2\mu}}{\sqrt{2}} \\ Z_\mu &= W_{3\mu} \cos\theta - B_\mu \sin\theta \\ A_\mu &= W_{3\mu} \sin\theta + B_\mu \cos\theta \end{aligned} \tag{1.13}$$

where θ is the weak mixing angle.

The gauge invariance of the electroweak Lagrangian is complicated by the observed non-zero mass of the physical gauge bosons fields $W^\pm$ and Z_0, carriers of the weak force. In fact mass terms like $M_W^2 W_\mu W^\mu$, $M_Z^2 Z_\mu Z^\mu$ and $m_f^2 \bar{f} f$ cannot be added to the derived Lagrangian, since they explicitly violate $SU(2)_L \times U(1)_Y$ gauge invariance.

A method called *Higgs Mechanism*, based on spontaneous symmetry breaking, is then used to solve the mass generation problem and will be briefly described.

Spontaneous symmetry breaking

The spontaneous symmetry breaking happens when the Lagrangian describing the dynamics of a physical system has a symmetry that is not preserved by the system ground states.

Given a gauge theory based on a local invariance with respect to a symmetry group G, and being $H \subset G$ the symmetry group of the vacuum state, with $dim(G) = N$ and $dim(H) = M$, the general formulation of the Goldstone theorem states that $N - M$ massless bosons will be absorbed by $N - M$ massive vector bosons. Therefore, in the $SU(2)_L \times U(1)_Y$, where $dim(G) = 4$ and $H = U(1)_{em}$, three vector bosons will realize the desired mass spectrum. This mechanism requires the introduction of the *Higgs* field, a doublet of complex fields: three of its four degrees of freedom will be spent for the longitudinal polarization states of the

massive bosons. The remaining degree of freedom is associated to the presence of the undetected Higgs particle, H_0.

The result of this theoretical environment is that the spontaneous symmetry breaking mechanism is responsible for the reduction of the symmetry group of the theory from $SU(2)_L \times U(1)_Y$ to $U(1)_{em}$, the latter being related to the electric charge conservation only.

The simplest Lagrangian for the $SU(2)_L \times U(1)_Y$ group manifesting spontaneous symmetry breaking can be written as:

$$\begin{aligned} \mathcal{L}_{SSB} &= (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \\ V(\Phi) &= -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad \lambda > 0 \end{aligned} \tag{1.14}$$

where $\Phi = \binom{\phi^+}{\phi^0}$ is a complex doublet with hypercharge $Y(\Phi) = 1$, and $V(\Phi)$ is the simplest renormalizable potential we can choose. If we choose $(-\mu^2) < 0$, then the minimum of the potential is realized on a circle of radius $v = \sqrt{\mu^2/\lambda}$ (see Fig. 1.2), and

$$| < 0 | \Phi | 0 > | = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \tag{1.15}$$

As a consequence of this choice, the lowest energy state of the system has a vacuum expectation value which no longer reflects the symmetry of the potential $V(\Phi)$, and the physical spectrum is then realized by performing "small oscillations" around the vacuum state. By parameterizing $\Phi(x)$ as

$$\Phi(x) = \exp\left(i \frac{\vec{\xi}(x) \vec{\sigma}}{v} \right) \begin{pmatrix} 0 \\ (v + H(x))/\sqrt{2} \end{pmatrix} \tag{1.16}$$

and eliminating the unphysical fields $\vec{\xi}(x)$ by means of gauge transformations, the mass spectrum can be obtained from the following terms of $\mathcal{L}_{SM}$:

$$\begin{aligned} (D_\mu \Phi')^\dagger (D^\mu \Phi') &= \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} + \frac{1}{2} \frac{(g^2 + g'^2) v^2}{4} Z_\mu Z^\mu + \dots \\ V(\Phi') &= \frac{1}{2} 2 \mu^2 H^2 + \dots \\ \mathcal{L}_{YW} &= \lambda_e \frac{v}{\sqrt{2}} \bar{e'_L} e'_R + \lambda_u \frac{v}{\sqrt{2}} \bar{u'_L} u'_R + \lambda_d \frac{v}{\sqrt{2}} \bar{d'_L} d'_R + \dots \end{aligned} \tag{1.17}$$

The tree level mass predictions for gauge and Higgs bosons are then the following:

$$\begin{aligned} M_{W_\mu^\pm} &= \frac{gv}{2} \\ M_{Z_\mu} &= \frac{\sqrt{g^2 + g'^2} v}{2} \\ M_{A_\mu} &= 0 \\ M_{Higgs} &= \sqrt{2\lambda} v \end{aligned} \tag{1.18}$$

where

$$v = \sqrt{\frac{\mu^2}{\lambda}} \tag{1.19}$$

is determined from the muon decay: $v = (\sqrt{2}G_F)^{-1/2} \sim 246\ GeV$ and fixes the scale of the spontaneous symmetry breaking.

This mechanism is called the Higgs mechanism and gives rise to mass terms for $W^{\pm}$, Z, as well as for quarks and leptons preserving the gauge invariance of the theory, at the cost of introducing a new scalar particle, not yet experimentally observed, the Higgs boson, whose mass and self-interaction are not theoretically determined.

1.1.3 Quantum Chromo Dynamics

The gauge theory for strong interactions is based on $SU(3)_C$, which is a non Abelian Lie group generated by color transformations. The Quantum Chromo Dynamics invariant Lagrangian can be built similarly to the QED one, with the difference that the $SU(3)_C$ symmetry will require the *color* to be conserved. Since the gauge group is non Abelian, this will cause the bosons mediating the interaction, the *gluons*, to posses color charge and to interact among themselves as well as with quarks.

Moreover, the additional gluon-gluon interactions cause the strong coupling constant α_S to have a qualitatively different behaviour with the interaction momentum transfer scale with respect to the QED coupling constant α_{QED}.

We can introduce the QCD covariant derivative:

$$D_\mu q = \left(\partial_\mu - i g_s \frac{\lambda_\alpha}{2} A_\mu^\alpha\right) q \tag{1.20}$$

where

$$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \tag{1.21}$$

are the quark fields, g_s is the strong coupling constant, $\frac{\lambda_\alpha}{2}$ are $SU(3)$ generators given by 3×3 traceless hermitian matrices, and A_μ^α are gluon fields, $\alpha = 1, ..., 8$.

Then the QCD Lagrangian can be written as:

$$\mathcal{L}_{QCD} = \sum_q \bar{q}(x)(iD_\mu \gamma^\mu - m_q)q(x) - \frac{1}{4} F_{\mu\nu}^\alpha F_\alpha^{\mu\nu} \tag{1.22}$$

where the gluon field strength tensors are defined as follows:

$$F_{\mu\nu}^\alpha(x) = \partial_\mu A_\nu^\alpha(x) - \partial_\nu A_\mu^\alpha(x) + g_s f^{\alpha\beta\gamma} A_{\mu\beta} A_{\nu\gamma} \tag{1.23}$$

and the related term in Eq. 1.22 provides three and four gluon interaction vertices.

In expression 1.23, g_S, the strong coupling constant (which is usually denoted as $\alpha_S = \frac{g_S^2}{4\pi}$), is found to decrease as the interaction energy scale increases, due to vacuum polarization effects induced by gluon self-interactions:

$$\alpha_S(q^2) = \frac{4\pi}{\left(11 - \frac{2}{3}N_f(q^2)\right) \ln\left(\frac{-q^2}{\Lambda_{QCD}^2}\right)} \tag{1.24}$$

In eq. 1.24, Λ_{QCD} is the QCD energy scale, $N_f(q^2)$ is the number of quark flavours that can be pair-produced at a given energy (i.e. the number of quark flavours with $m_q < \sqrt{-q^2}/2$). The "running" of α_S with energy allows the strong coupling to be small enough at high energy, allowing a perturbative description of the strong force. However, at small momentum transfer comparable with the mass of the light hadrons, α_S becomes of order unity and the perturbation approximation breaks down. This large value of the coupling constant is the source of most mathematical complications and uncertainties in QCD calculations at low energy. On the other hand, it is of great importance that α_S tends to zero in the high energy limit. This property gives rise to the so-called "asymptotic freedom", and allows perturbation theory to be used in theoretical calculations to produce experimentally verifiable predictions for hard scattering processes. At the same time the behaviour of the strong coupling constant at low energy is responsible for quark confinement into hadrons.

Trying to separate colored particles requires increasing energy density in the binding color string, since the interaction potential grows linearly with the distance between the outgoing partons, until the creation of new color-singlet hadronic states becomes energetically favorable and energy is materialized into colored quark pairs. The fact that quarks are forced into color singlets yields final state color-neutral hadrons rather than free quarks and gluons. Thus a hard scattered parton evolves into a shower of partons and finally into hadrons. This process is called parton shower evolution or hadronization.

1.2 Physics beyond the Standard Model

Recent developments show that the Standard Model of particle physics is incomplete and many issues still remain open, for example the recently proved non zero masses of neutrinos, that would require an extension of the model.

Another problem, the so-called *hierarchy problem* concerns the corrections to the Higgs mass: in fact the Higgs boson mass receives divergent quadratic radiative corrections which need to be controlled by means of fine-tuning cancellations in order to keep the mass at the electroweak energy scale fixed. Several ways of solving this issue have been explored, for example the hypothesis of new strong dynamics that could appear at the TeV scale (Technicolor theories). Another possible explanation allows the divergent corrections to m_H to be cancelled by a new spectrum of particles at the electroweak scale: supersymmetric (SUSY) theories propose a supersymmetric partner for each SM particle with different spin, solving the hierarchy problem by considering radiative corrections from supersymmetric partners. SUSY requires additional Higgs fields in order to provide mass to fermions and their superpartners, for example in the minimal supersymmetric extension of SM, the MSSM, there are five Higgs bosons: h, H, A and $H^{\pm}$ which are associated to two complex doublets.

Furthermore, Grand Unification would require an extension of the Standard Model to include gravitational interactions in the theory.

1.3 The Top quark

The top quark was discovered during Run 1 of the Tevatron operation by CDF and DØ collaborations at Fermilab in 1995 [5, 6]. It was another success of the Standard Model, which had strongly predicted its existence.

Several experimental results and theoretical arguments already prior to the top quark discovery had provided evidence for its existence. These hints are mainly based on theoretical self-consistency (namely the absence of anomalies), the absence of flavour changing neutral currents (FCNC), and the measurement of weak isospin of the b-quark $T_3 = -1/2$, thus demanding a $T_3 = 1/2$ partner in its isospin multiplet.

In 1964 Christenson and collaborators observed violation of CP symmetry in rare decays of neutral kaons at the Brookhaven National Laboratory [7].

To accomodate this result in the theory, in 1973, Kobayashi and Maskawa added a phase factor $e^{i\delta}$ into their quark mixing matrix [8]. At that time, only three quarks (u, d, s) were known. In their work they concluded that the only way to have a renormalizable theory of weak interactions with CP violation was to introduce additional fields, thus proposing the existence of three complete generations of quarks, since the smallest unitary matrix which can exhibit a non removable complex phase is 3×3 in size.

Afterwads, in 1974, at Brookhaven [9] and SLAC [10], two experiments independently observed a new resonance at 3.1 GeV/c^2, the particle J/ψ, which was immediately interpreted as a $c\bar{c}$ bound state: this discovery of the charm-quark completed the second generation of quarks.

Furthermore, one year later, in 1975, M. L. Perl and collaborators at SLAC made the first observation of the τ lepton [11], evidence for the existence of a third lepton and quark generation.

In 1977 the FNAL-E-0288 experiment collaboration at Fermilab discovered the b-quark ($\Upsilon = b\bar{b}$) [12]. The searches for a companion, the top quark, initiated immediately, based on the existence of the b and the empirically observed generation grouping of the quarks and leptons previously discovered.

The quark model suggested that within any family fermions must appear in left-handed doublets and right-handed singlets of weak isospin [13]. So, in accordance with the structure of the first generation, the left-handed b-quark was expected to be part of a doublet of weak isospin ($T^3_{b_L} = -1/2$), while the right-handed b was associated to a isospin singlet: $T^3_{b_R} = 0$. In the hypothesis that the t-quark did not exist, a b-quark would have appeared only as a singlet state: $T^3_{b_L} = T^3_{b_R} = 0$. However the weak isospin of b-quarks was determined on the basis of the measurement of the forward-backward asymmetry and of the total width of the $b\bar{b}$ production, by the JADE collaboration at DESY [14] and more recently from LEP experiments [15], determining the b-quark to be part of a doublet of weak isospin.

Additionally, the experimentally determined absence of *flavour changing neutral currents*, an important feature of the Standard Model that excludes processes like $b \to \mu^+\mu^- X$ or $b \to sX$, where X is a state with no net flavour quantum numbers, implies that the b quark is a member of a $SU(2)$ doublet.

Another compelling argument for the existence of the top quark follows from a theoretical consistency requirement. The renormalizability of the Standard Model

demands the absence of triangle anomalies. Triangular fermion loops built-up by an axial-vector charge combined with two electric vector charges Q would break the renormalizability of the Standard Model. In order to avoid this from happening it is sufficient to impose a constraint on the sum of the electric charges of all the left-handed fermions:

$$\sum_L Q = 0 \tag{1.25}$$

This condition is met in a complete standard generation in which the electric charge of the leptons and those of the quarks of all color components add up to zero:

$$\sum_L Q = -1 + 3 \times \left[\left(+\frac{2}{3} \right) + \left(-\frac{1}{3} \right) \right] = 0 \tag{1.26}$$

The absence of the top quark in the third generation would violate the condition of Eq. 1.25.

1.3.1 Top quark production

At hadron colliders top quarks are produced mainly in pairs through strong interactions. Even if protons and antiprotons are not elementary particles, but composed of quarks and gluons, thanks to the asymptotic freedom property of QCD if the momenta of the initial particles are high enough ($\gg \Lambda_{QCD} \sim 200\ MeV$), we can consider the interaction to take place between just two elementary particles (quarks or gluons), one in each incoming hadron, neglecting interactions among the other constituents of proton and antiproton.

The initial momentum of the interacting partons is however unknown, since a given parton carries a fraction x of the proton (or antiproton) momentum according to a statistical distribution named "parton distribution function" (PDF). For each parton type these functions describe the probability to find it with a momentum xP inside the proton [16], where P is the momentum of the proton (Fig. 1.3). The valence quarks (u and d) are most likely to carry a large fraction of the proton momentum, while gluons and sea quarks tend to carry smaller fractions. All allowed parton-parton interaction channels contribute to the experimental $t\bar{t}$ production cross section $\sigma_{t\bar{t}}$ to an amount depending on their distribution functions in the primary hadrons, so in order to calculate it we must sum over all the possible interactions, weighted by their probability specified by the PDF's. For proton-antiproton collisions:

$$\sigma(p\bar{p} \to t\bar{t}) = \sum_{i,j} \int dz_i dz_j f_{i/p}(z_i, \mu^2) f_{j/\bar{p}}(z_j, \mu^2) \hat{\sigma}(ij \to t\bar{t}; \hat{s}, \mu^2, M_{top}) \tag{1.27}$$

where the sum is over light quarks and gluons contained in the initial proton and antiproton, carrying momentum z_i and z_j of the initial hadron respectively; $f_{i/p}$ and $f_{j/\bar{p}}$ are the parton distribution functions for proton and antiproton respectively; $\hat{\sigma}$ is the parton-parton cross section. The center-of-mass energy of the $i-j$ parton system is denoted by $\hat{s}$ and the parameter μ is a factorization scale which is introduced to include resultant contributions from higher order Feynman diagrams.

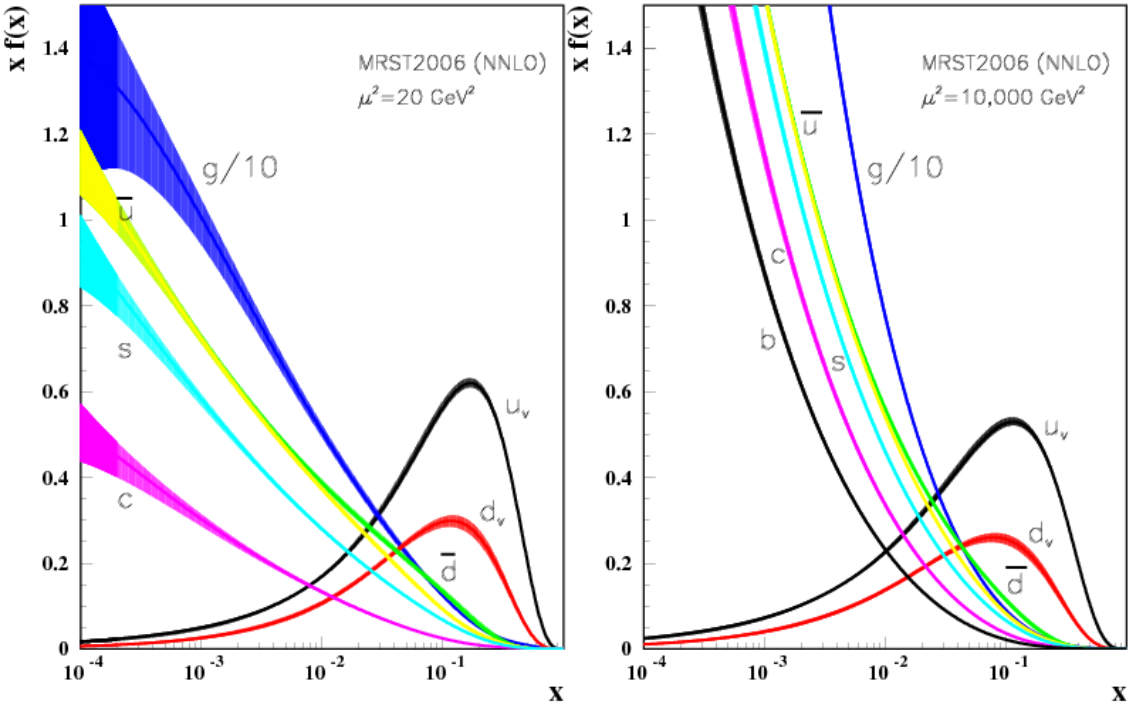


Figure 1.3: Parton distribution functions of quarks and gluons in the proton at two different momentum transfers μ^2 [16].

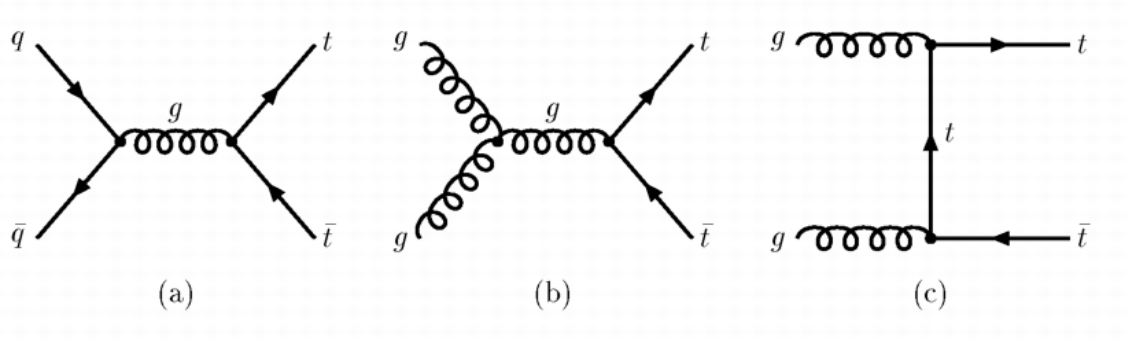


Figure 1.4: Leading order Feynman diagrams for $t\bar{t}$ production via strong interaction: (a) $q\bar{q}$ annihilation, (b) and (c) gg fusion.

At the Tevatron center-of-mass energy of $\sqrt{s} = 1.96\ TeV$ top quark pair production occurs 85% of the times via quark-antiquark annihilation ($q\bar{q}$) and for the remaining 15% via gluon fusion (gg). The leading order Feynman diagrams are shown in Fig. 1.4.

The theoretical Standard Model prediction for $t\bar{t}$ production at $\sqrt{s} = 1.96\ TeV$, depends on the top mass value M_{top} as shown in Fig. 1.5, and is $\sigma_{t\bar{t}} = 6.7^{+0.7}_{-0.9}\ pb$ for a top mass of 175 GeV/c^2 [17, 18], meaning that, since the total $p\bar{p}$ inelastic cross section is about 80 $mbarn$, we expect roughly one in 10^{10} collisions ($\approx 7 \cdot 10^{-4} Hz$

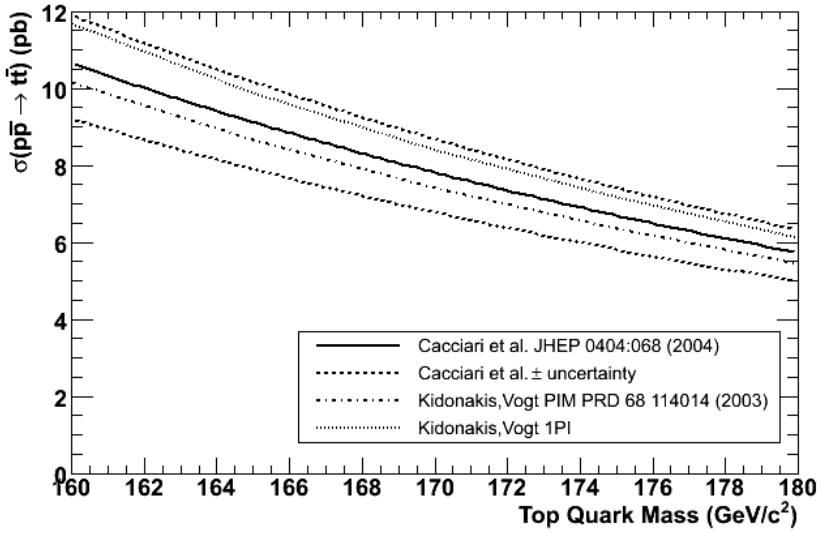


Figure 1.5: Dependence of $t\bar{t}$ top pair production cross section on top quark mass.

rate) at Tevatron to produce a $t\bar{t}$ top quark pair, thus providing a real challenge in the discrimination of the top events among a huge background.

The measurement of the production cross section for $t\bar{t}$ pairs can be a test for QCD, since a significant deviation from the predicted value could indicate some kind of non Standard Model production mechanism. Fig. 1.6 shows some recent cross section measurements by CDF.

Single-top Production

Within the Standard Model, a single top quark can also be produced via electroweak interaction trough the following processes (see Fig. 1.7):

- t-channel: a space-like W boson ($q^2 \leq 0$) strikes a b quark in the proton sea, promoting it to a top quark; this channel is often referred to as $W-gluon$ fusion, since the b quark arises from a gluon splitting to $b\bar{b}$.

- s-channel: rotating the t-channel diagram so that the W boson becomes time-like ($q^2 \geq (m_t + m_b^2)$), single top production can happen trough $q\bar{q}$ annihilation.

- associated production: single top may be also produced via weak interaction in association with a real W boson ($q^2 = M_W^2$); one of the initial partons is a b quark in the proton sea, as in the t-channel.

The cross sections for all these processes are proportional to the matrix element $|V_{tb}|^2$ of the CKM matrix (see next section), therefore measuring the single top production cross section provides a direct probe of this SM parameter.

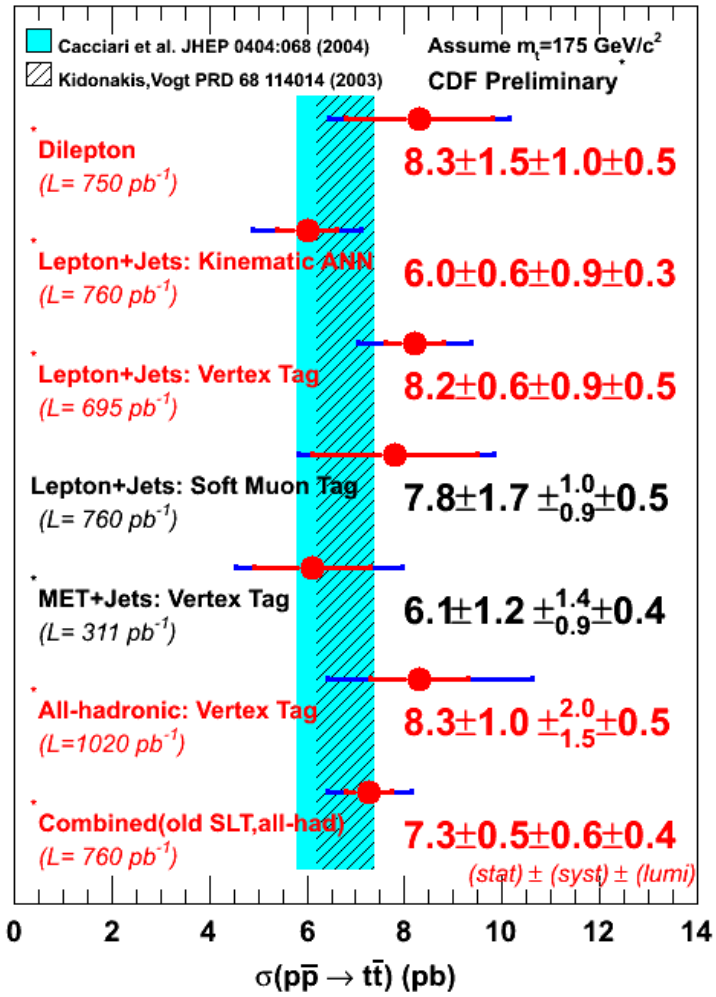


Figure 1.6: Cross section measurements by CDF compared with theoretical predictions (shaded). This plot is updated to March 2007.

1.3.2 Top quark decays

In the Standard Model the top quark decay is mediated by the weak force, and its dominant decay signature is $t \to W^+ b$ or $\bar{t} \to W^- \bar{b}$ with branching ratio $BR(t \to Wb) \sim 1$. The additional decay channels $t \to Wd$ and $t \to Ws$ are allowed by the Standard Model but highly suppressed, thus giving minimal contribution, due to the very small values of the off-diagonal elements in the quark flavour mixing matrix of weak eigenstates, the Cabibbo-Kobayashi-Maskawa (CKM) matrix. CKM matrix arises because of the difference of mass and weak eigenstates for quarks, and can be expressed as a 3×3 unitary matrix operating on the charge $-1/3$ quark mass eigenstates (d, s and b) [16]:

$$V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \approx \begin{pmatrix} 0.973 & 0.227 & 0.0039 \\ 0.021 & 0.972 & 0.042 \\ 0.0081 & 0.041 & 0.999 \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

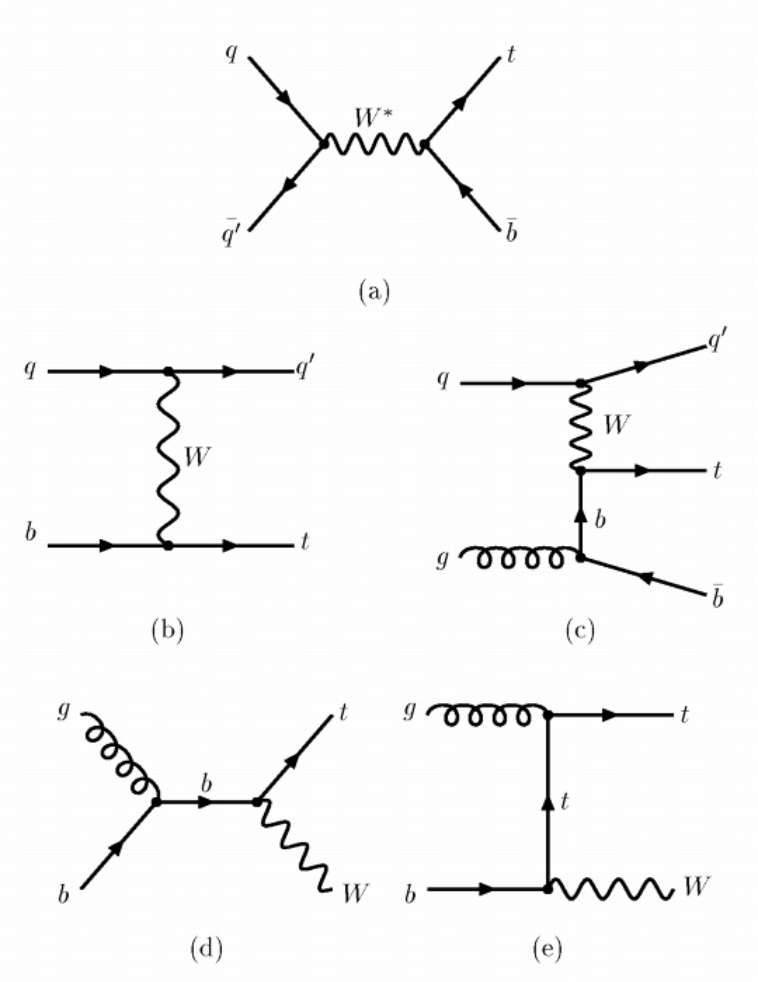


Figure 1.7: Leading order Feynman diagrams for electroweak production of single top quarks: (a) s-channel, (b,c) t-channel and (d,e) associated production with a W.

The Standard Model predicts the top quark decay width to be [16]:

$$\Gamma(t \to Wb) = \frac{G_F M_{top}^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{M_{top}^2}\right) \left(1 + 2\frac{M_W^2}{M_{top}^2}\right) \left[1 - \frac{2\alpha_S}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \quad (1.28)$$

For $M_{top} = 175\ GeV/c^2$ we have:

$$\Gamma(t \to Wb) \approx 1,55\ GeV \quad \longrightarrow \quad \tau_{top} = \frac{1}{\Gamma_{top}} \approx 4 \cdot 10^{-25}\ s \quad (1.29)$$

This large width ($\Gamma_{top} \gg \Lambda_{QCD}$) causes the top quark to decay before hadronizing (its width is smaller then the characteristic hadronization time of QCD $\tau_{had} \approx 28 \cdot 10^{-25}\ s$), allowing its observation as a free particle. In particular, this feature enables precision mass measurements, otherwise impossible for the other quarks due to non-perturbative effects in the hadronic bound state.

Thus to detect the top quark we just need to identify and reconstruct its decay products; consequently, the top pair decay signatures are classified according to the

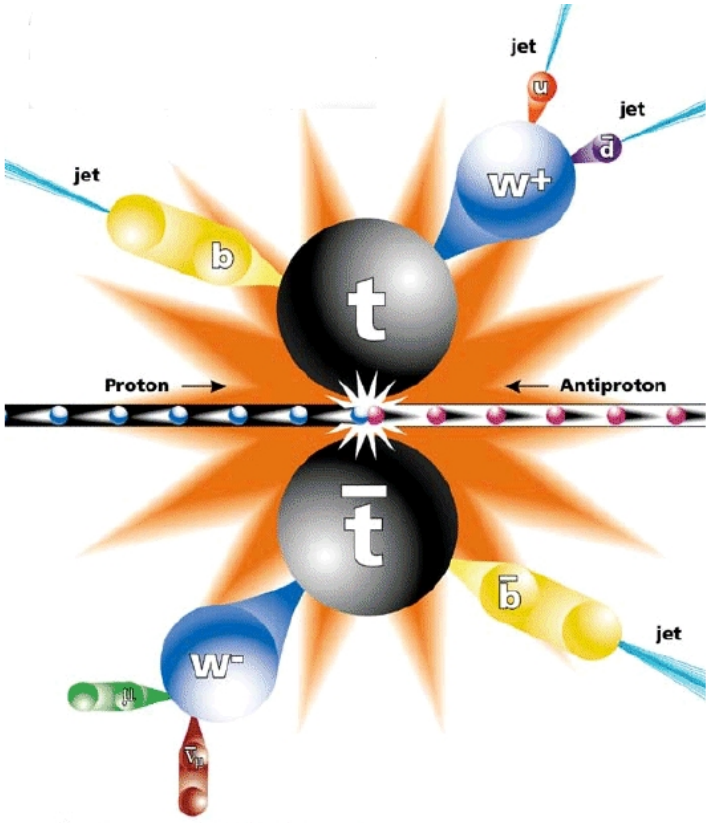


Figure 1.8: Pictorial view of $t\bar{t}$ top pair production at tree level by $q\bar{q}$ annihilation followed by the pair decay into the $\mu + jets$ channel.

Channel	Decay Mode	Branching Ratio
All-hadronic	$t\bar{t} \to q\bar{q}'b\ q\bar{q}'\bar{b}$	36/81
Lepton+jets	$t\bar{t} \to q\bar{q}'b\ e\nu\bar{b}$	12/81
	$t\bar{t} \to q\bar{q}'b\ \mu\nu\bar{b}$	12/81
	$t\bar{t} \to q\bar{q}'b\ e\tau\bar{b}$	12/81
Di-lepton	$t\bar{t} \to e\nu b\ e\nu b$	1/81
	$t\bar{t} \to \mu\nu b\ \mu\nu\bar{b}$	1/81
	$t\bar{t} \to e\nu b\ \mu\nu\bar{b}$	2/81
	$t\bar{t} \to e\nu b\ \tau\nu\bar{b}$	2/81
	$t\bar{t} \to \mu\nu b\ \tau\nu\bar{b}$	2/81
	$t\bar{t} \to \tau\nu b\ \tau\nu\bar{b}$	1/81

Table 1.2: Standard Model $t\bar{t}$ decay modes and their associated relative branching ratios.

W decay modes. The W bosons decay to either one of the three generation leptons, $W^+ \to e^+\nu_e$, $W^+ \to \mu^+\nu_\mu$, $W^+ \to \tau^+\nu_\tau$, or into the lightest two generations of quarks: $W^+ \to u\bar{d}$, $W^+ \to c\bar{s}$.

This gives rise to different decay channels that produce different experimental

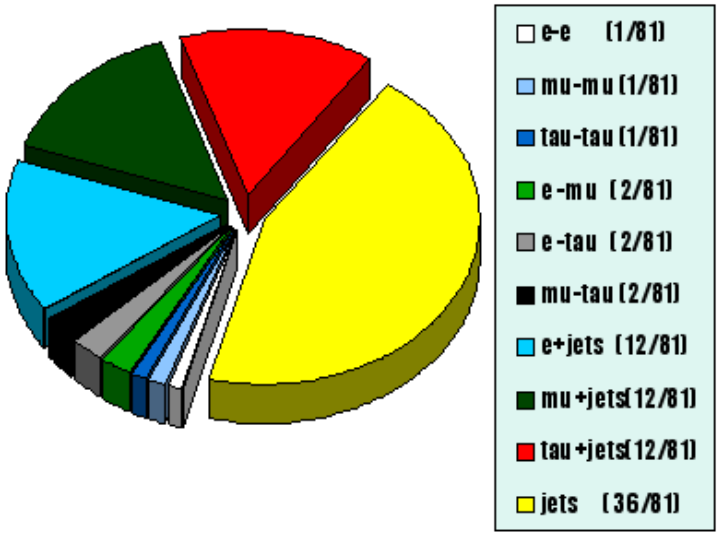


Figure 1.9: Standard Model $t\bar{t}$ decay signatures.

signatures in the detector. The *di-lepton* category represents the case in which both W bosons decay leptonically; this channel is somehow complicated by the two non-observable neutrinos in the final state, has the lowest branching ratio, but is the cleanest among the three, due to the clear signature of its two leptons. Main background sources are di-boson and Drell-Yan events.

The *lepton+jets* signature on the other hand arises when one of the W decays hadronically and the other into $l\nu$; those involving τ's are difficult to isolate because of the poor tau signature.

Finally, the *all-hadronic* channel corresponds to the case in which both W bosons decay into quarks; this channel has the largest branching ratio but suffers from a large QCD background of multijet states.

The possible $t\bar{t}$ decay modes and their corresponding branching ratios are summarized in Tab. 1.2 and Fig. 1.9.

1.3.3 Top quark mass

The top quark mass M_{top}, is an important parameter in different areas of Particle Physics. Its precise measurement is important to set basic parameters in the calculation of the electroweak processes, and provides a constraint on the mass of the Higgs boson.

In fact, W mass theoretical calculation is subject to radiative corrections that arise from creation and absorbtion of virtual quarks and bosons. Quark corrections depend on top mass while boson corrections depend on $log(M_H)$, where M_H is the Higgs boson mass. Measuring with high precision W and top mass we can thus obtain a constraint on the mass of the Higgs. Fig. 1.10 shows the limits on Higgs mass that can be derived from direct and indirect measurement of top quark and W masses.

The current value of the top mass is set at $170.9 \pm 1.1\ (stat)\ \pm 1.5\ (syst)\ GeV/c^2$ (which corresponds to a total uncertainty of $1.8\ GeV/c^2$) as a result of a combi-

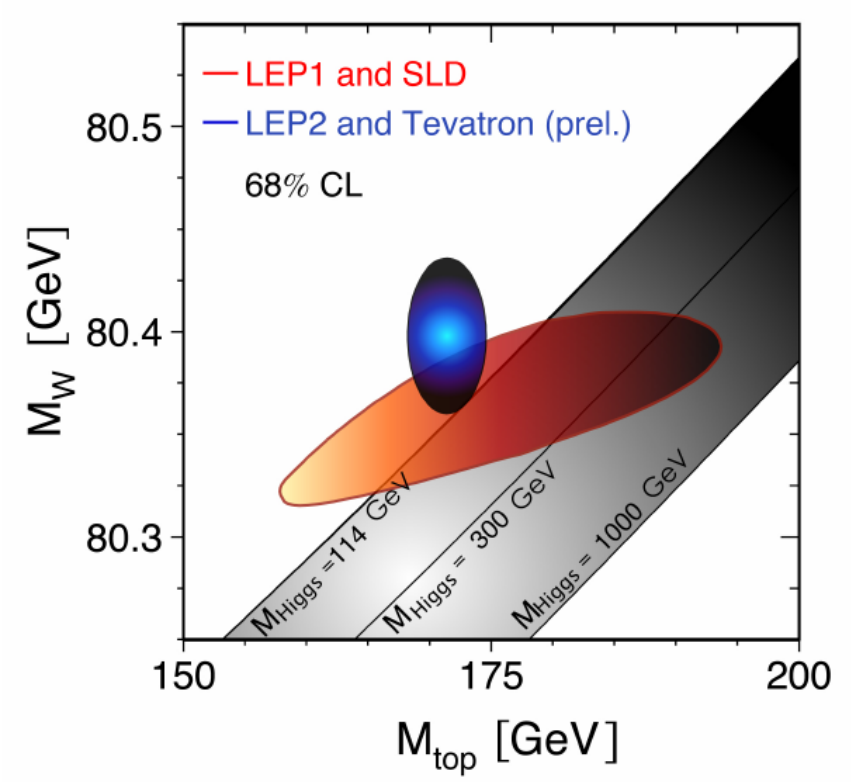


Figure 1.10: Relationship between M_W and M_{top} as a function of the Higgs mass. Expectations for a number of H masses are shown within the shaded band. Available EW data and Run 1 Tevatron measurements of M_W and M_{top} favour low M_H values. The small ellipse (1σ radius in the two observables) indicates the expected constraint by higher precision measurements of M_W, M_{top} at the end of Run 2. Results are from CDF, DØ, LEP and SLD.

nation of Tevatron Run I and Run II measurements [19], making it the heaviest known elementary particle (see Fig. 1.11 for CDF results).

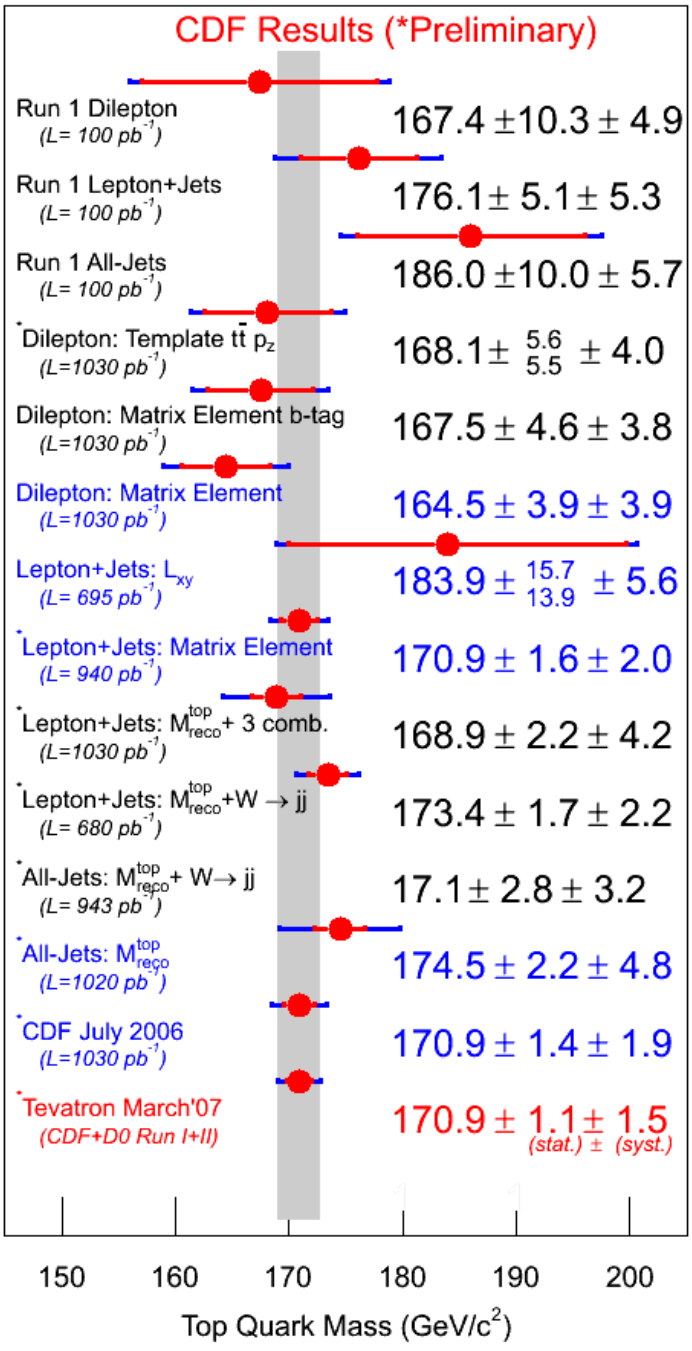


Figure 1.11: Most recent CDF results using different techniques and channels compared to the Tevatron average. Measurements in blue were included in the CDF combination of March 2007.

Bibliography

[1] F. Mandl and G. Shaw, **Quantum Field Theory**, *Wiley and Sons (1984).*

[2] M. Peskin and D. Schroeder, **An Introduction to Quantum Field Theory**, *HarperCollins (1995).*

[3] C. Quigg, **Gauge Theories of the Strong, Weak, and Electromagnetic Interactions**, *Addison-Wesley (1997).*

[4] S.F. Novaes, *Standard Model: An Introduction*, hep-ph/0001283.

[5] F. Abe *et al.* [CDF Collaboration], *Evidence for top quark production in anti-p p collisions at* $\sqrt{s}$ *= 1.8 TeV*, Phys. Rev. Lett. **73**, (1994) 225.

[6] S. Abachi *et al.* [DØ Collaboration], *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632.

[7] J. H. Christenson *et al.*, *Evidence For The 2 Pi Decay Of The K(2)0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[8] M. Kobayashi and T. Maskawa, *CP Violation In The Renormalizable Theory Of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[9] J. J. Aubert *et al.*, *Experimental Observation Of A Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[10] J. E. Augustin *et al.*, *Discovery Of A Narrow Resonance In E+ E- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406.

[11] M. L. Perl *et al.*, *Evidence For Anomalous Lepton Production In E+ E- Annihilation*, Phys. Rev. Lett. **35** (1975) 1489.

[12] S. W. Herb *et al.*, *Observation Of A Dimuon Resonance At 9.5 GeV In 400 GeV Proton - Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.

[13] S. Weinberg, *A Model Of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[14] W. Bartel *et al.*, *A Measurement Of The Electroweak Induced Charge Asymmetry In E+ E-* $\rightarrow$ *B Anti-B*, Phys. Lett. B **146** (1984) 437.

[15] LEP collaborations, *A combination of Preliminary LEP Electroweak Measurement and contraints on the Standard Model*, $CERN - PPE$ (1995) 95.

[16] A. B. Balantekin *et al.* [Particle Data Group], *Review of particle physics*, J. Phys. G: Nucl. Part. Phys. **33** (2006).

[17] M. Cacciari *et al.*, JHEP 0404:068 (2004).

[18] N. Kidonakis and R. Vogt, Phys. Rev. D **68** 114014 (2003).

[19] Tevatron Electroweak Working Group [for the CDF and D0 Collaborations], *A Combination of CDF and D0 Results on the Mass of the Top Quark*, hep-ex/0703034, CDF Internal Note 8735, DØInternal Note 5378.

Chapter 2

Tevatron Accelerator complex

The Tevatron [1] is a proton-antiproton accelerator hosted at the Fermi National Accelerator Laboratory. With its center-of-mass energy of $\sqrt{s} = 1.96\ TeV$ it is the source of the highest energy $p\bar{p}$ collisions to date, up to now the only machine capable of letting us examine dimensions up to $10^{-15}\ m$, looking at the hadron constituents, the quarks.

The Tevatron is the final and largest element of the Fermilab accelerator complex, illustrated in Fig. 2.1, and works primarily as a $p\bar{p}$ collider; however, it can also accelerate a single proton beam and operate in fixed target mode to provide a number of neutral and charged particle beams. The Tevatron collider obtained the first collisions in 1985, and during the course of its lifetime provided several physics runs, as listed in Tab. 2.1.

In the following, after spending a few words on one of the fundamental accelerator parameters, the luminosity, a description of the acceleration apparatus will be given.

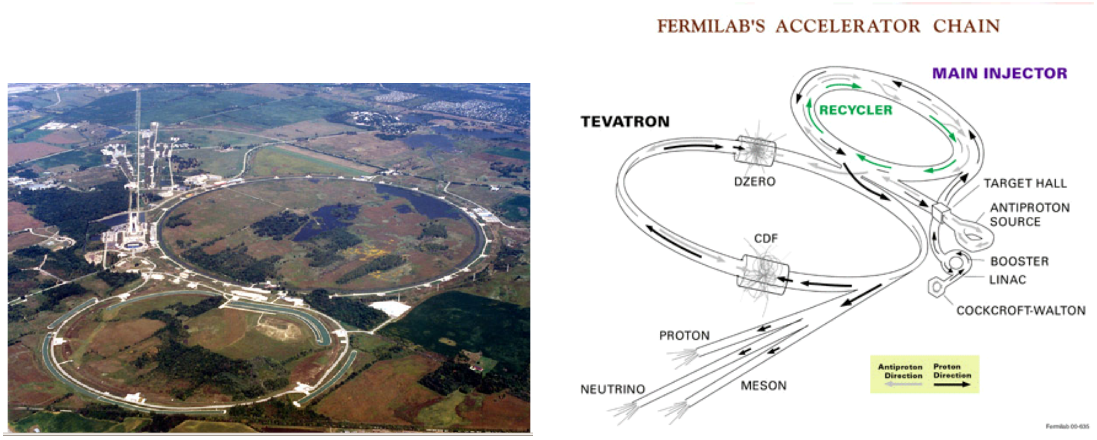
2.1 Instantaneous and integrated Luminosity

While building an accelerator, a fundamental construction parameter is the design luminosity that needs to be achieved; in fact luminosity is a resource directly related to the computation of the probability $W_{i\to f}$ for a generic process $i \to f$, where i and f are the initial and final states, respectively. In the case of Tevatron, the initial state is made up by two particles, a proton and an antiproton, while the final state is composed by a generic number N of particles. Taking into account the overall four-momentum conservation, the probability amplitude for the $p, \bar{p} \to f$ process has the following general structure:

$$\langle f|T|p;\bar{p}\rangle = (2\pi)^2\delta^{(4)}(P_f - p_p - p_{\bar{p}})\langle P_f|M|p_p;p_{\bar{p}}\rangle \tag{2.1}$$

where we made the following assumptions regarding each particle a in the initial and final states: a is described by a narrow wave packet that obeys, as obvious, the on-shell mass condition, the Klein-Gordon equation, and moreover, that is peaked around a four-momentum p_a, giving the following equation (were we hided all remaining quantum numbers):

$$\mathcal{F}_p^a(x) \equiv \langle x|a\rangle = \frac{1}{(2\pi)^{3/2}}\int d^4q\ \theta(q_0)\delta(q^2 - m_a^2)\hat{f}_p(q)e^{-i\,qx} \tag{2.2}$$



(a) (b)

Figure 2.1: (a): An airplane view of the Fermilab laboratory. The ring at the bottom of the figure is the Main Injector, the above ring is the Tevatron. On the left are clearly visible the paths of the external beamlines: the central beamline is for neutral beams and the side beamlines are for charged beams (protons on the right, mesons on the left). (b): A sketch of the Fermilab accelerator chain.

Run	Period	Int. Lum. (pb^{-1})
First Test	1997	0.025
Run 0	1988-1989	4.5
Run 1A	1992-1993	19
Run 1B	1994-1995	90
Run 1C	1995-1996	1.9
Run 2A	2001-2004	400
Run 2B	2004-	>2000

Table 2.1: Integrated luminosity delivered by the Tevatron in its physics runs. Run2B is still in progress.

Integrating the square modulus of Eq. 2.1 over its space dependencies and after other manipulations that use approximations allowed by the narrowness of the wave packets, assuming that protons and antiprotons are grouped in bunches, we end up with the following transition probability $W_{i\to f}$:

$$W_{i\to f} \approx (2\pi)^4 \delta^4(P_f - p_p - p_{\bar{p}}) |\langle P_f | M | p_p; p_{\bar{p}} \rangle|^2 \frac{1}{4\omega_p \omega_{\bar{p}}} \int d^4x \; \rho_p(x) \rho_{\bar{p}}(x) \qquad (2.3)$$

where ω's denote energies and ρ's, that have the meaning of probability density of particle location, are the time component of conserved four-currents given by:

$$i(\mathcal{F}^* \partial_\mu \mathcal{F} - \mathcal{F} \partial_\mu \mathcal{F}^*) \qquad (2.4)$$

The square amplitude in Eq. 2.3 is what can be computed by means of the Standard Model theory. What appears in the integral depends on the experimental setup; the integral itself has dimension of an inverse cross section and is a measure of the

probability that incoming protons and antiprotons have to come in interaction. We can assume that the densities ρ are Gaussian near the collision points and that, for simplicity, the collisions themselves are head-on; then, parameterizing the bunches path by s and calling $x(s)y(s)$ a plane orthogonal to the path in s, we can write approximately

$$\rho^{\pm}(x(s), y(s), s \pm vt) = \frac{N^{\pm}}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_s}\, exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} - \frac{(s \pm vt)^2}{2\sigma_s^2}\right) \tag{2.5}$$

where $\pm$ refers to proton/antiproton, N is the number of particles in a bunch, v is the speed of the bunches and σ's denote the radii of the portion of the crossing bunches that effectively overlap. In Eq. 2.3 we have consequently:

$$\begin{aligned} \int d^4x\ \rho_p(x)\rho_{\bar{p}}(x) &\equiv \nu\Delta \int dx\ dy\ ds\ dt \rho_p(x, y, s + vt)\rho_{\bar{p}}(x, y, s - vt) \\ &= \nu \frac{N_p N_{\bar{p}}}{4\pi\sigma_x\sigma_y}\frac{\Delta}{2v} \\ &\equiv \mathcal{L}\frac{\Delta}{2v} \end{aligned} \tag{2.6}$$

where Δ is the whole lasting of the data taking, long with respect to the duration of each effective crossing of the colliding bunches, ν is the frequency of the crossing of the proton and anti-proton bunches, and (the lab reference frame is also the center of mass frame in our case)

$$v = \frac{|\vec{p}|}{\omega}, \quad |\vec{p}| = |\vec{p_p}| = |\vec{p_{\bar{p}}}|, \quad \sqrt{m_p^2 + \vec{p}^2} = \omega = \omega_p = \omega_{\bar{p}} \tag{2.7}$$

Thus we have

$$\begin{aligned} \frac{dW_{i\to f}}{dt} &\approx \delta^4(P_f - p_p - p_{\bar{p}})\frac{(2\pi)^4}{2\omega|\vec{p}|}|\langle f|M|p_p; p_{\bar{p}}\rangle|^2\mathcal{L} \\ &= \frac{(2\pi)^4\delta^4(P_f - p_p - p_{\bar{p}})}{\sqrt{(p_p \cdot p_{\bar{p}})^2 - m_p^2 m_{\bar{p}}^2}}|\langle f|M|p_p; p_{\bar{p}}\rangle|^2\mathcal{L} \\ &\equiv \sigma_{int}\mathcal{L} \end{aligned} \tag{2.8}$$

$\mathcal{L}$ is usually called (instantaneous) luminosity, while its integral over time L is called integrated luminosity. The bigger the luminosity, the bigger the probability to observe an interaction. For this reason the Tevatron has undergone a series of improvements during its lifetime in order to increase this fundamental parameter. Tab. 2.1 shows the luminosity served by the accelerator during its different physics runs.

2.2 The proton source

The process leading to $p\bar{p}$ collisions begins in a *Cockroft-Walton* generator (see Fig. 2.2) in which H^- gas is produced by hydrogen ionization. H^- ions are immediately accelerated by means of a multi step voltage divider up to an energy of

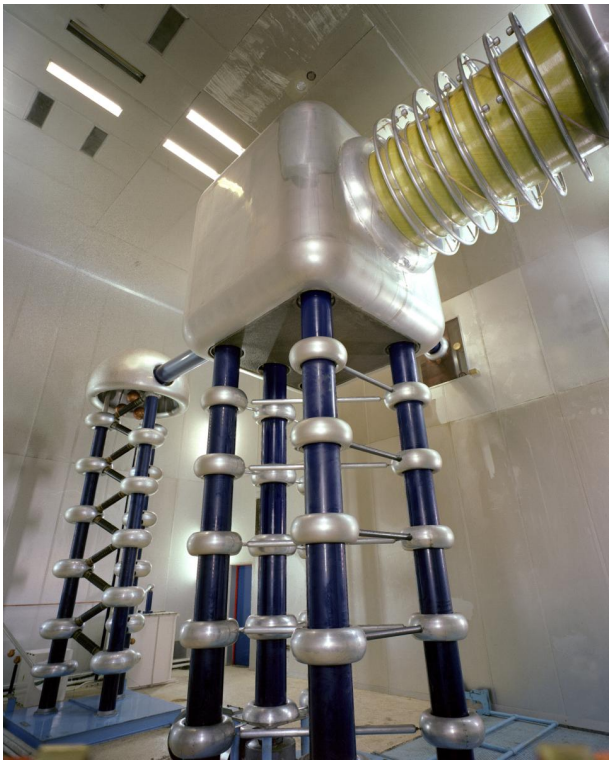


Figure 2.2: The Cockroft-Walton generator, the starting point of the proton acceleration chain.

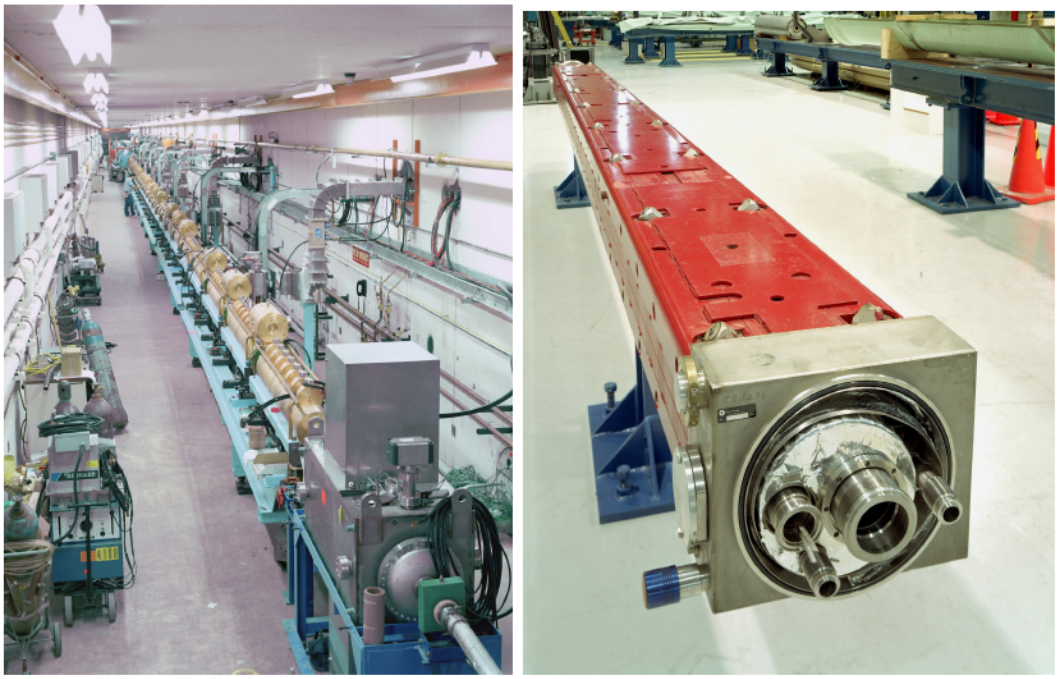


Figure 2.3: Left: upstream view of the 400 MeV section of the Linac. Right: Tevatron Superconducting Dipole Magnet.

750 KeV and then transported through a transfer line to the linear accelerator, the Linac.

The second stage of the accelerator chain is a 150 meters long linear accelerator (see Fig. 2.3): the *Linac* [2, 3] picks up the H^- ions at energy of 750 KeV, and accelerates them up to the energy of 400 MeV inducing an oscillating electric field between a series of electrodes.

The *Booster* [4] takes the 400 MeV negative hydrogen ions from Linac, strips the electrons off, which leaves only protons, and accelerates them up to 8 GeV. The Booster is the first circular accelerator in the Tevatron chain, and consists of a series of magnets arranged around a 75-meter radius circle with 18 radio frequency cavities. The Booster loading scheme overlays the injected beam of negative H^- ions from the Linac with the one of H^+ already circulating in the machine in order to increase beam intensity; then the mixed beams go through a carbon foil, which strips off the electrons turning the negative hydrogens into protons. When the bare protons are collected in the Booster, they are accelerated to the energy of 8 GeV by the conventional method of varying the phase of RF fields in the accelerator cavities [1], and subsequently injected into the Main Injector. The final "batch" will contain a maximum of 5×10^{12} protons divided among 84 bunches spaced by 18.9 ns, each consisting of 6×10^{10} protons.

2.3 The Main Injector

The *Main Injector* [5] is a circular synchrotron with a 3 km circumference (seven times the circumference of the Booster) and plays a crucial role in linking the Fermilab acceleration facilities: the Main Injector can accelerate or decelerate particles between the energies of 8 GeV and 150 GeV. The sources of these particles and their final destination are variable, depending on the Main Injector operation mode: it can accept 8 GeV protons from the Booster, or 8 GeV antiprotons from the Recycler; it can accelerate protons up to 120 GeV for antiproton production or deliver a proton beam to fixed target experiments. The beam energy, for both proton and antiproton, can reach 150 GeV during the collider mode when particles are injected to the Tevatron for the last stage of the acceleration. Furthermore, once Tevatron collisions end, the Main Injector can accept back the 150 GeV antiprotons in order to decelerate them down to 8 GeV before injecting them in the Recycler.

The Linac accelerates protons to 400 MeV, and the Booster guides them up to 8 GeV. Afterwards the proton beam, through a transfer line, reaches the Main Injector where by means of radio frequency systems it is accelerated and bunched.

We can summarize the functions of the Main Injector as:

- *Antiproton production*: Providing beam to the antiproton production target is one of the simplest tasks of the Main Injector. In this mode, a single batch of protons is accepted from the Booster, accelerated up to 120 GeV and extracted towards the target, which yields 8 GeV antiprotons.

- *Fixed target modes*: During fixed target operation, protons are accelerated to the desired energy and then extracted to a stationary target, external to

the ring. Extraction takes place from the Main Injector at 120 GeV. The target can be anything from a sliver of metal to a flask of liquid hydrogen depending on the experiment needs.

- *Collider operations*: in Collider Mode, in addition to supplying 120 GeV protons for antiproton production, the Main Injector must also feed the Tevatron protons and antiprotons at 150 GeV. The Main Injector maximum stored beams are $\sim 3 \cdot 10^{13}$ protons and $\sim 2 \cdot 10^{12}$ antiprotons, beams are stored in 36 bunches in the Tevatron. When the collision operation ends, another task of the Main Injector is to recover antiprotons from the Tevatron, decelerate and then send them to the Recycler.

2.4 The antiproton source

The number of antiprotons available is an important limiting factor in producing the high luminosity desired for Tevatron physics.

The Fermilab antiproton source [6] is comprised of a target station, two rings called the Debuncher and the Accumulator, and the transfer lines between these rings and the Main Injector.

Antiprotons $\bar{p}$ are produced from the 120 GeV proton beam extracted from the Main Injector and focused on a nikel target. Antiprotons are collected at 8 GeV with wide acceptance around the forward direction, injected into the Debuncher Ring, debunched into a continuous beam and stochastically cooled. The beam is then transferred between cicles to the Accumulator were antiprotons are stored at a rate of about $25 \cdot 10^{10}$ $\bar{p}/hour$ (improvements in the storage rate are still being made). Stacking within the accumulator acceptance is limited to a stored beam of about 10^{12} antiprotons.

When enough antiprotons have been accumulated in the Accumulator, their transfer starts. Antiproton beam destination can be either the Main Injector or the Recycler ring (see Sec. 2.4.1).

Overall it can take from 10 to 20 hours to build up a stack of antiprotons, which is then used in the Tevatron collisions. Antiproton availability is the most limiting factor attaining high luminosities, so in this context it is important to mention the Recycler, the part of the acceleration chain designed to collect antiprotons left at the end of a collider *store*, the period of time in which the colliding beams are retained in the Tevatron (roughly 20 hours).

2.4.1 The Recycler ring

The Recycler [7, 8] is a 3.3 Km long storage ring of fixed 8 GeV kinetic energy, and is located directly above the Main Injector. It is composed primarily of permanent gradient magnets and quadrupoles. Three main tasks are designed for the Recycler operations:

1. The most important feature of the Recycler is that it allows antiprotons left over at the end of Tevatron Collider stores to be re-cooled and re-used, allowing to recycle almost 75% of the antiprotons left after a store.

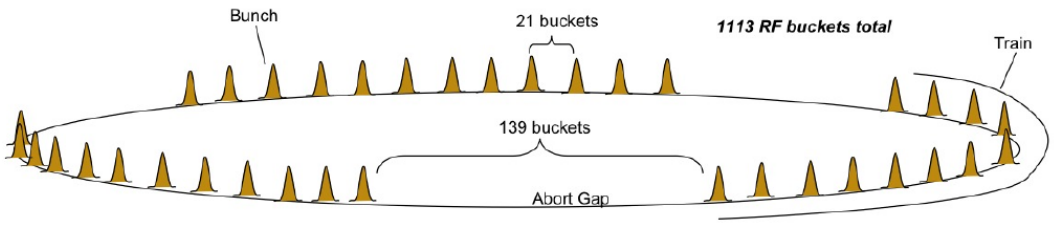


Figure 2.4: Bunch structure of the Tevatron beams in Run 2.

2. It allows the Accumulator to operate optimally: since the antiproton production rate decreases as the beam current in the Accumulator ring increases, the Recycler is designed to act as a post-Accumulator cooler ring, being the final storage for 8 GeV antiprotons.

3. The usage of permanent magnets in the construction of the Recycler allows to dramatically reduce the probability of unexpected losses of antiprotons. In fact, the ring has been designed so that Fermilab-wide power could be lost for an hour with the antiproton beam surviving.

After a store has been circulating in the Tevatron for several hours, as particles are gradually lost, the beam size slowly grows, and the luminosity degrades: a decision is then made to terminate the store and load a fresh one. To do so, since proton and antiproton orbits follow different paths in the Tevatron, large chunks of metal are slowly moved into the proton beam until only the antiprotons are left. At this point, antiprotons can be decelerated from 1 TeV to 150 GeV and then transferred to the Main Injector. While the antiprotons are still circulating at 150 GeV, they are decomposed back into fewer bunches (usually seven). The antiprotons are then decelerated to 8 GeV and transferred to the Recycler ring. This procedure is repeated until no antiprotons from the store are left into the Tevatron ring.

2.5 The Tevatron ring

The Tevatron [9] is the last stage of the Fermilab accelerator chain. The Tevatron is a 1 km radius synchrotron able to accelerate the incoming 150 GeV beams from Main Injector to 980 GeV, providing a center of mass energy of 1.96 TeV. The accelerator employs superconducting magnets (see Fig. 2.3) requiring cryogenic cooling and consequently a large scale production and distribution of liquid helium. During Run II the Tevatron operates at the 36×36 bunches mode.

The antiprotons are injected after the protons have already been loaded. When the Tevatron loading is complete, the beams are accelerated to the maximum energy and collisions begin. The beam revolution time is 21 μs. The beams are split in 36 bunches organized in 3 trains each containing 12 bunches (see Fig. 2.4). Within a train the time spacing between bunches is 396 ns. An empty sector 139

buckets long (2.6 μs) is provided in order to allow the kickers to raise to full power and abort the full beam into a dump in a single turn. This is done at the end of a run or in case of an emergency. In the 36×36 mode, there are 72 regions along the ring where the bunch crossing occurs. While 70 of these are parasitic, in the vicinity of CDF and DØ detectors additional focusing and beam steering is performed, to maximize the chance of a proton striking an antiproton. The focusing reduces the beam spot size and thus increases the luminosity, as seen in Eq. 2.7 that shows how smaller values of σ_x, σ_y imply larger luminosity values. During collisions the instantaneous luminosity decreases in time as particles are lost and the beams begin to heat up. When the luminosity becomes too low to allow a significant datataking (approximately after 15-20 hours) the current store is dumped and a new cycle starts. A number of reasons can cause unwanted early termination of runs. Typical failures are a vacuum leak, a power supply failure or a magnet quench, a loss of magnet superconductivity in the ring.

Table 2.2 summarizes the accelerator parameters for Run II.

Parameter	Value
Particles collided	$p\bar{p}$
Maximum beam energy	0.980 TeV
Time between collisions	0.396 μs
Crossing angle	0 μrad
Energy spread	0.14×10^{-3}
Bunch length	57 cm
Beam radius	$39\mu m$ for p, $31\mu m$ for $\bar{p}$
Filling time	30 min
Injection energy	0.15 TeV
Particles per bunch	24^{10} for p; 3×10^{10} for $\bar{p}$
Bunches per ring per species	36
Average beam current	$66\mu A$ for p, $8.2\mu A$ for $\bar{p}$
Circumference	6.12 Km
$\bar{p}$ source accumulation rate	$13.5 \times 10^{10}/hr$
Max number of $\bar{p}$ in accumulation ring	2.4×10^{12}

Table 2.2: Accelerator parameters for Run II configuration [10].

Fig. 2.5 shows the Tevatron peak luminosity as a function of the time from the beginning of Run II. The blue squares show the peak luminosity at the beginning of each store. The red triangle displays a point representing the last 20 peak values averaged together.

Fig. 2.6 on the other hand shows the weekly and total integrated luminosity to date; while Fig. 2.7 shows the total luminosity delivered by the Tevatron compared to the total luminosity recorded by the CDF experiment as a function of the store number.

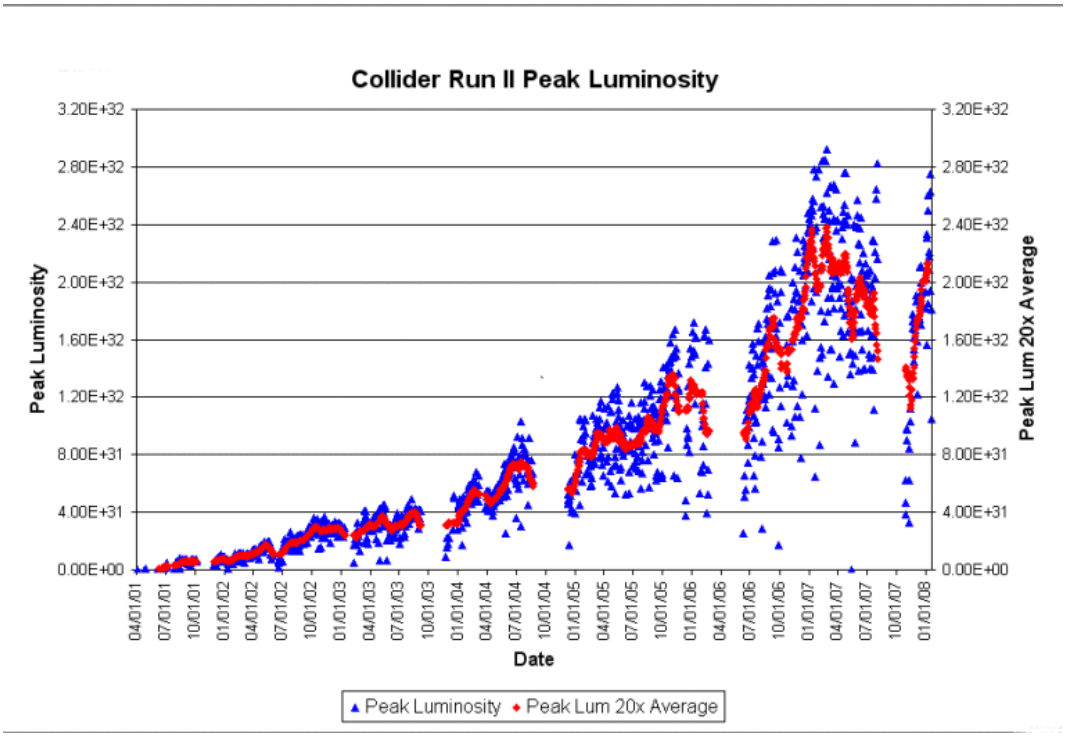


Figure 2.5: Run II peak luminosity.

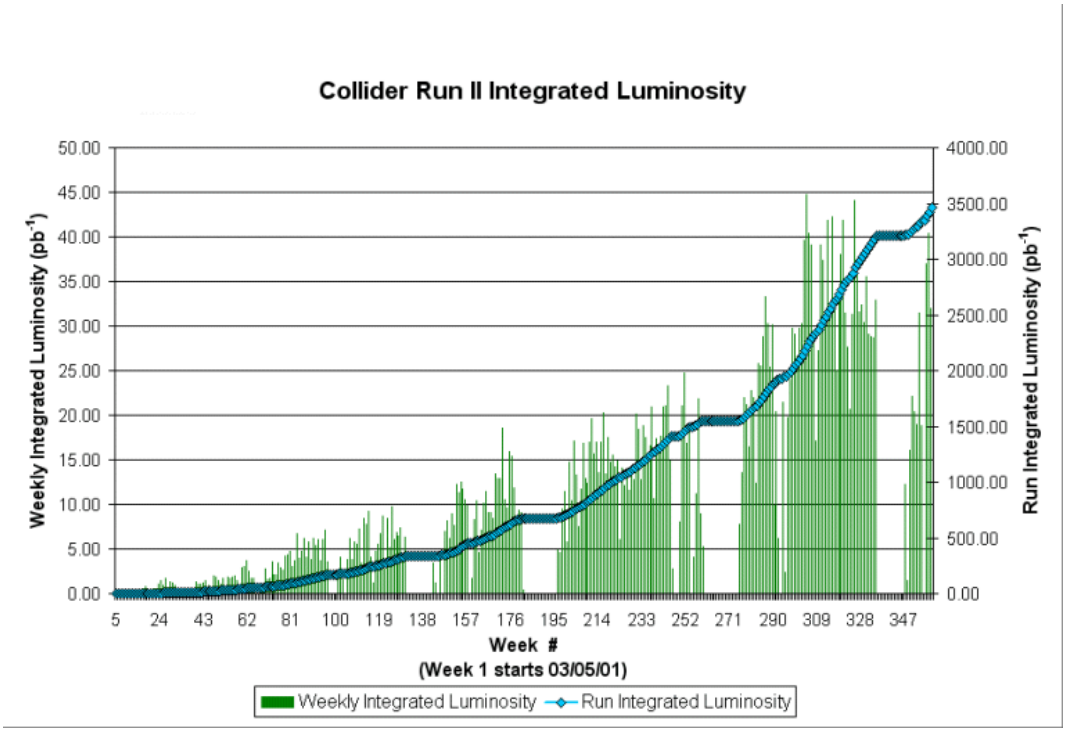


Figure 2.6: Weekly and total integrated luminosity.

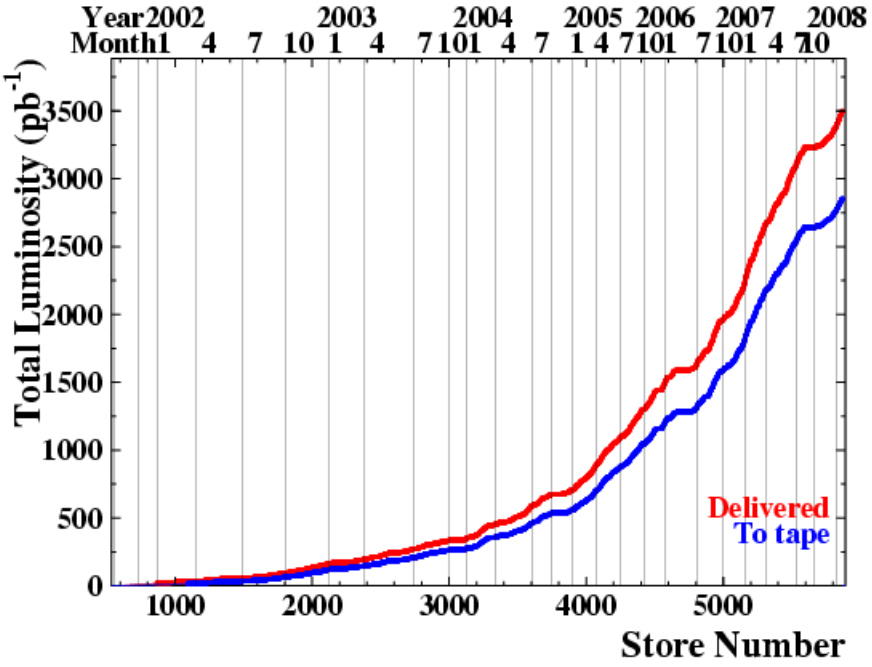


Figure 2.7: Delivered and CDF acquired integrated luminosity as a function of the Store number.

Bibliography

[1] Fermilab Beams Division, *Run II Handbook*, http://www-ad.fnal.gov/runII/index.html.

[2] Fermilab Beams Division, *Fermilab Linac Upgrade. Conceptual Design*, FERMILAB-LU-Conceptual Design, http://lss.fnal.gov/archive/linac/fermilab-lu-999.shtml.

[3] Fermilab Beams Division, *The Linac rookie book*, http://www-bdnew.fnal.gov/operations/rookie_books/LINAC_v2.pdf.

[4] Fermilab Beams Division, *The Booster rookie book*, http://www-bdnew.fnal.gov/operations/rookie_books/Booster_V3_1.pdf.

[5] Fermilab Beams Division, *The Main Injector rookie book*, http://www-bdnew.fnal.gov/operations/rookie_books/Main_Injector_v1.pdf.

[6] Fermilab Beams Division, *The Antiproton Source rookie book*, http://www-bdnew.fnal.gov/operations/rookie_books/Pbar_V1_1.pdf.

[7] Fermilab Beams Division, *The Recycler rookie book*, http://www-bdnew.fnal.gov/operations/rookie_books/Recycler_RB_v1.pdf.

[8] G. Jackson, *The Fermilab recycler ring technical design report. Rev. 1.2*, FERMILAB-TM-1991 (1991).

[9] Fermilab Beams Division, *The Tevatron rookie book*, http://www-bdnew.fnal.gov/operations/rookie_books/Tevatron_v1.pdf.

[10] A. B. Balantekin *et al.* [Particle Data Group], *Review of particle physics*, J. Phys. G: Nucl. Part. Phys. **33** (2006).

Chapter 3

The CDF Detector in Run II

This Chapter is dedicated to a detailed description of the CDF detector used to study $p\bar{p}$ interactions provided by the Tevatron, and to a specific explanation of all the detector sub-systems, whose role is crucial for the reconstruction of the physical objects needed for our analysis.

At the center of mass energy available at the Tevatron, proton-antiproton interactions are interpreted in terms of collisions between their constituents. At this level, the phenomenology is usually described in the framework of *Quantum Chromo Dynamics* (QCD), as already highlighted in Chapter 1. At the end of the interaction process and after *hadronization*, collimated *jets* of particles emerge from the scattering, whose energies and directions carry a reminiscence of initial partons ones.

In the collisions, apart from QCD processes, electroweak production of W and Z bosons takes place as well. For that reason, aside the detection of collimated spray of particles, the capability of detecting charged leptons and neutrinos as missing energy is of great importance in the design of a particle detector.

The *Collider Detector at Fermilab* (CDF) is described below as configured for Run II; additional technical details covering all parts of the detector can be found in CDF *Technical Design Report* [1] and in a series of guides for experimenters [2] and official lectures [3].

A detector elevation view is presented in Fig. 3.1. The CDF architecture is quite common for this type of detectors: radially from the inside to the outside it features a tracking system contained in a superconducting solenoid, calorimeters (electromagnetic and hadronic) and muon detectors. The whole CDF detector weighs about 6000 tons.

CDF is located around one of the the two interaction points along the Tevatron ring and has been designed in order to perform precise measurements of energy and momentum of the jets and charged leptons produced by the $p\bar{p}$ collisions, as well as the missing energy due to the neutrinos created in W and Z decays. Besides, it has been studied to provide a first identification of the produced particles, particularly of the ones with relatively long lifetime coming from heavy quarks hadronization.

The reconstruction of an event begins with the identification of jets performed by the calorimetry system. In order to determine the direction of the jet momenta, a precise measurement of the event interaction center is needed; moreover, the identification of the jets originated by heavy quarks requires an accurate reconstruction

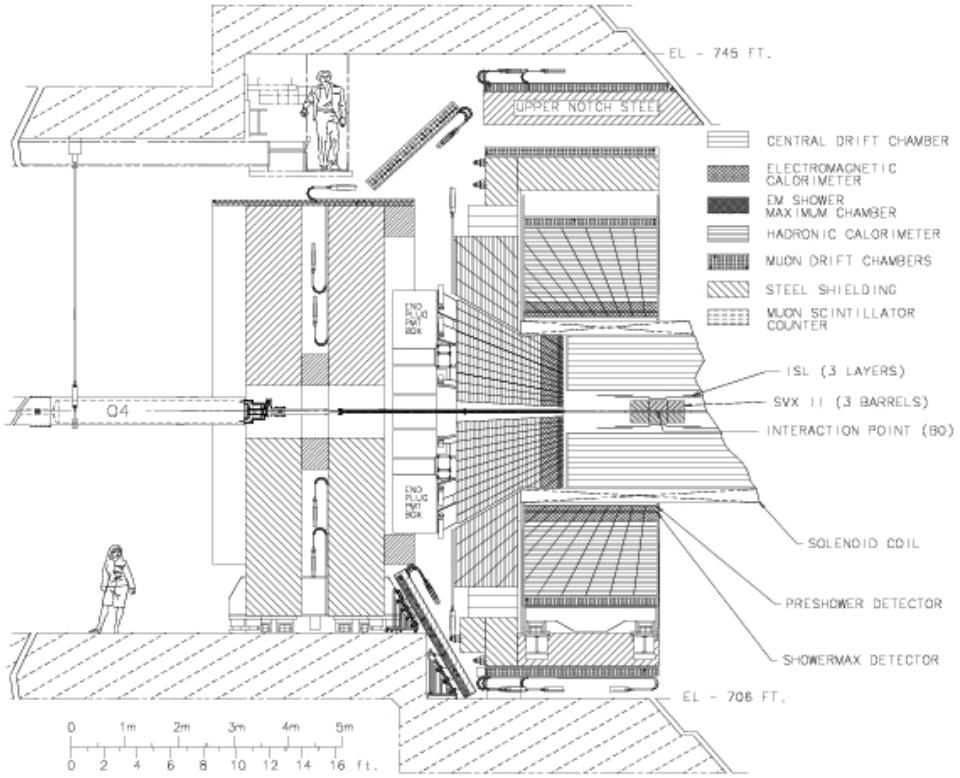


Figure 3.1: Elevation view of the CDF II Detector.

of the secondary vertices produced in heavy flavour decays. These measurements take advantage of the presence in the jets of charged particles, whose transverse momentum and trajectories can be reconstructed by a performant tracking system situated between the beam pipe and the calorimeter. Calorimetric and tracking informations are also used to identify electrons produced in the event. Outside the calorimeter, a complex of drift chambers for muon identification is arranged. Muons are very penetrating and leave a modest quantity of energy in the calorimeter: in order to identify them, tracks with high transverse momentum are extrapolated and matched to low energy calorimetric deposits and to stubs reconstructed in the external muon chambers.

In the following the structure of the detector CDF II will be examined in detail.

3.1 CDF Coordinate systems

CDF uses a Cartesian coordinate system centered in the nominal point of interaction, with the z axis coincident with the beamline and oriented parallel to the motion of the proton beam. The x axis is in the horizontal plane of the accelerator ring, pointing radially outward, while the y axis points vertically up (see Fig. 3.2).

For the simmetry of the detector, it is often convenient to work with cylindrical (z, r, ϕ) or polar (r, θ, ϕ) coordinates. The azimuthal angle ϕ is measured in the $x-y$ plane starting from the x axis, and it is defined positive in the anti-clockwise

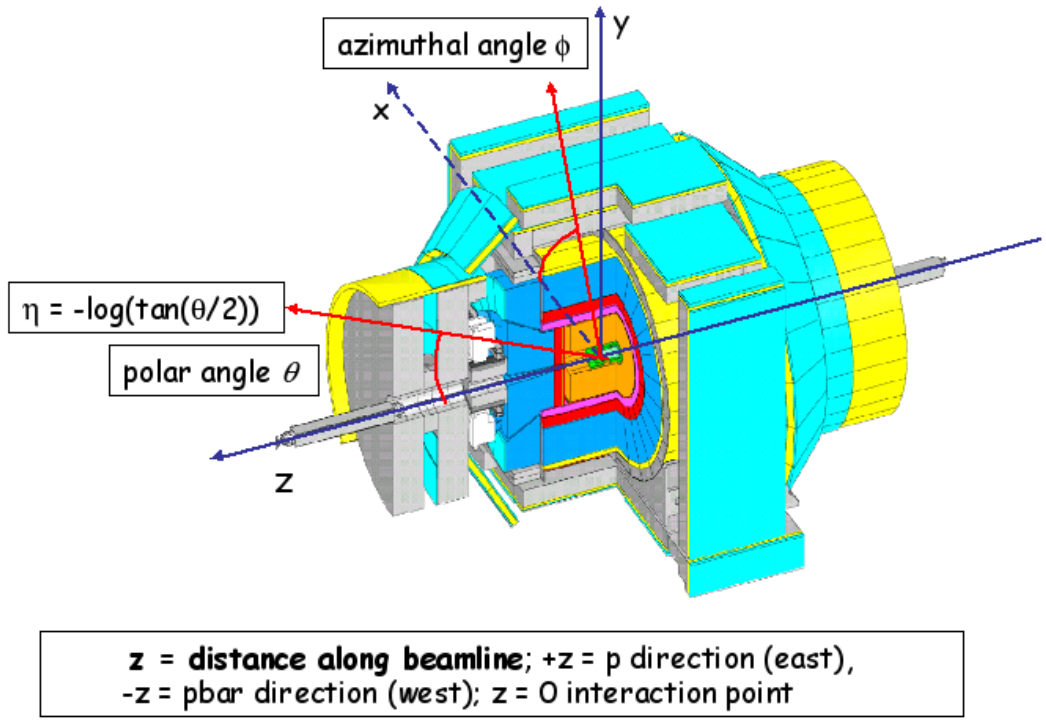


Figure 3.2: Isometric view of the CDF II Detector and its coordinate system.

direction; on the other side, the polar angle θ is measured from the positive direction of the z axis. The coordinate r defines the transverse distance from the z axis. Another important coordinate that can be used instead of the polar angle θ, is called *pseudorapidity* and it is defined as:

$$\eta = -\log \tan \frac{\theta}{2} \tag{3.1}$$

The pseudorapidity is usually preferred to θ at hadron colliders, where events are boosted along the beamline, since it transforms linearly under Lorentz boosts, i.e. η intervals are invariant with respect to boosts. For these reasons, the detector components are chosen to be as uniformly segmented as possible along η and ϕ coordinates.

3.2 Tracking system

Charged particles passing through matter cause ionization typically localized near the trajectory followed by the particle through the medium. Detecting ionization products gives geometrical information that can be used to reconstruct the particle's path in the detector by means of the *tracking* procedure.

The inner part of CDF II is devoted to the tracking system, whose volume is immersed in an uniform magnetic field of magnitude $B = 1.4\,T$, oriented along

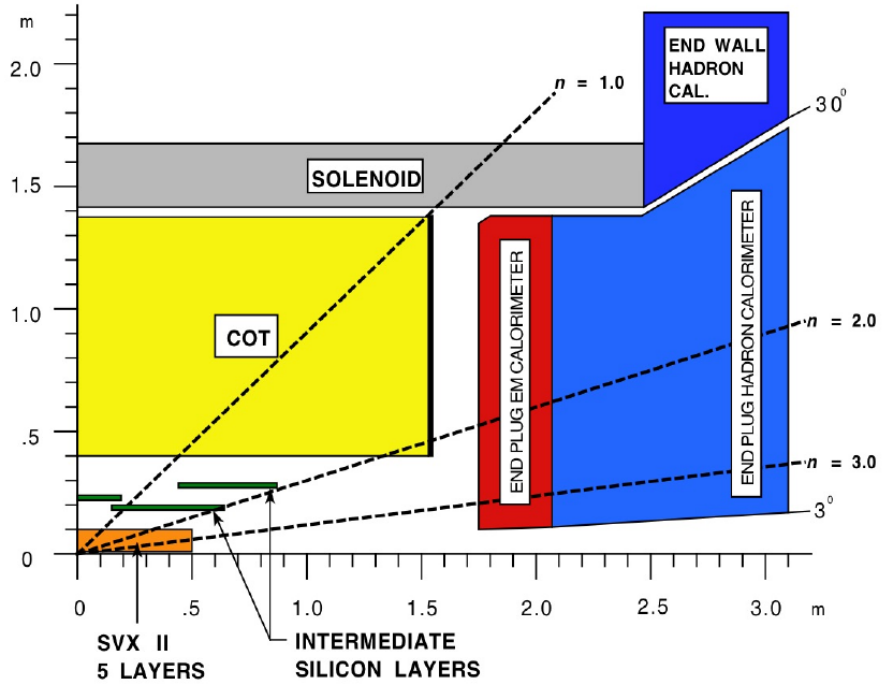


Figure 3.3: The CDF II detector subsystems projected on the z/y plane.

the z-axis. The Lorentz force induced on charged particles constrains them to an helicoidal trajectory, whose radius measured in the transverse plane $x-y$ is directly related to the particles transverse momentum P_T (see 4.1 for details).

The CDF II tracking system is basically divided into an inner silicon strip detector aimed to provide a precise vertex determination, and an outer drift chamber for momentum measurements. Fig. 3.3 shows the overall CDF II tracking volume, covering a pseudorapidity range up to $|\eta| = 2$.

3.2.1 Silicon vertex detector

The silicon vertex detector is crucial for precise determination of particle positions, and in particular its information can be used to infer the presence in the event of secondary decay vertices produced by heavy flavour decays.

The basic principle on which silicon strip detectors are based relies on the fact that a charged particle traveling across a silicon crystal produces electron-hole pairs. In fact the fundamental characteristic of semiconductor materials, such as silicon, is the presence of a full valence band that is separated from the conduction band by an energy gap of only few eV.

When an electron is excited from the valence band to the conduction band, a positive "hole" is left in the valence band while the excited electron becomes a negative charge carrier in the conduction band.

By introducing impurities (*doping*) with a different number of valence electrons, the number of available charge carriers in the semiconductor can be increased.

Doped semiconductors can be divided in two categories:

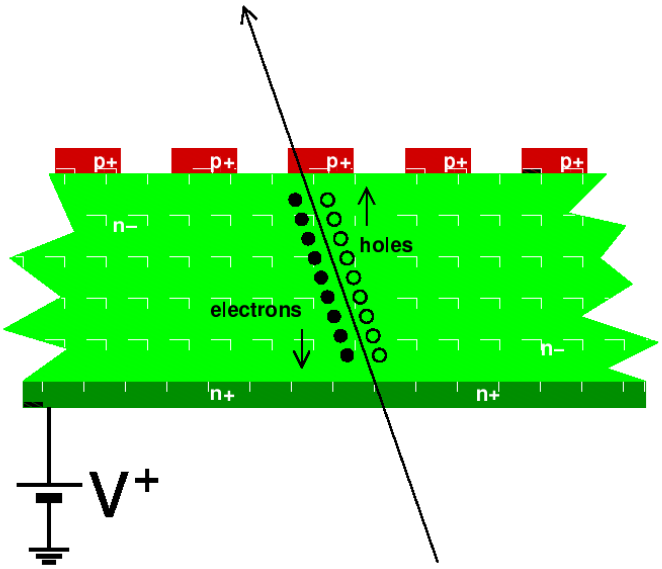


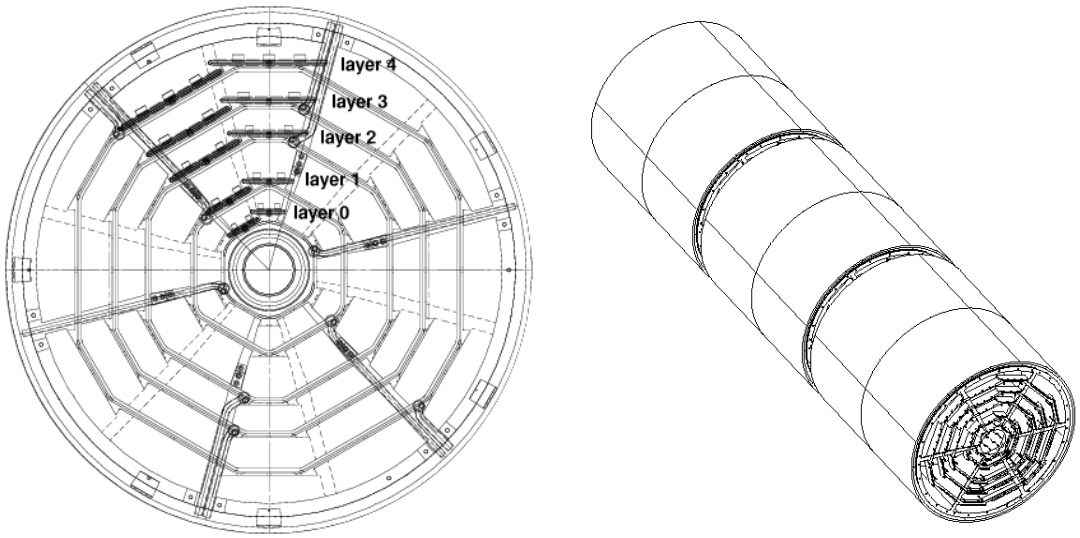
Figure 3.4: A generic silicon micro-strip detector.

1. n-type semiconductors, when the introduced impurity has one more valence electron than the silicon: the semiconductor will have additional electrons for excitation into the conduction band.

2. p-type semiconductors, when the introduced impurity has one less valence electron than the silicon: the semiconductor has an excess of "holes" as charge carriers in the valence band.

When one n-type semiconductor and one p-type semiconductor are placed together, the resulting device, called $n-p$ junction, has some very special properties. Due to the fact that each semiconductor contains charge carriers of differing polarity, the negative electrons in the n-type semiconductor will be drawn towards the positive holes in the p-type semiconductor and viceversa. After the equilibrium is reached, the n-type side possesses a net positive charge and the p-type side possesses a net negative charge: an electrical potential barrier is created and a depletion region arises between the n-type and p-type regions. For the silicon, the size of this depletion region is typically 10 μm and the potential through the junction is 0.6 eV. For each μm of depletion region traversed by an ionizing particle typically 100 electron-hole pairs are produced, whose identification becomes easier as the size of the depletion region increases; that's why the size of the depletion region is thus increased applying external electric voltages to the $p-n$ junction (reverse-biasing). This results in a larger sensitivity in detecting the ionization signals produced by incoming charged particles.

In a typical silicon micro-strip detector (Fig. 3.4), finely spaced strips of strongly doped p-type silicon (p^+) are implanted on a lightly doped n-type silicon substrate (n^-) $\sim$ 300 μm thick. On the opposite side, with respect to the p-type silicon implantation, a thin layer of strongly doped n-type silicon (n^+) is deposited. A positive voltage applied to the n^+ side depletes the n^- volume of free electrons



(a) $r-\phi$ view of SVX II. (b) Perspective view of SVX II.

Figure 3.5: CDF II Silicon Vertex Detector.

Layer	Radius [cm]		# of strips		Strip pitch [μm]		Stereo	Ladder [mm]	
	stereo	r - phi	stereo	r - ϕ	stereo	r - ϕ	angle	width	length
0	2.55	3.00	256	256	60	141	90°	15.30	4×72.43
1	4.12	4.57	576	384	62	125.5	90°	23.75	4×72.43
2	6.52	7.02	640	640	60	60	+1.2°	38.34	4×72.43
3	8.22	8.72	512	768	60	141	90°	46.02	4×72.43
4	10.10	10.65	896	896	65	65	-1.2°	58.18	4×72.43

Table 3.1: SVX summary.

and creates an electric field. When a charged particle crosses the active volume it creates a trail of electron-hole pairs from ionization and the presence of the electric field drifts the holes to the p^+ implanted strips producing a well localized signal. Usually the signal is collected by a cluster of strips, rather than being concentrated in just one strip. This allows to calculate the crossing point of the particle with a precision greater than the strip spacing, by weighting the strip positions by the amount of charge collected by each strip. With this method the Silicon Vertex Detector installed by CDF collaboration can achieve individual hit position accuracy of 12 μm.

The CDF II Silicon VerteX Detector is shown in Fig. 3.5(a) and 3.5(b), and it is known as SVX II [1]. It is composed of three different barrels each 29 cm long, each barrel supporting five layers of double-sided silicon micro-strip detectors between 2.5 and 10.7 cm from the beamline. The layers are numbered from 0 (innermost) to 4 (outermost); layers 0, 1 and 3 have wires parallel to the beam axis on one side (*axial* strips for r - ϕ measurement) and tilted by 90°on the other side (*stereo* strips for r - z measurement); layers 2 and 4 have axial strips on one side and stereo strips tilted by a small angle (1.2°) on the other (Tab. 3.1).

To reach better performances in terms of resolution and tracking coverage, two

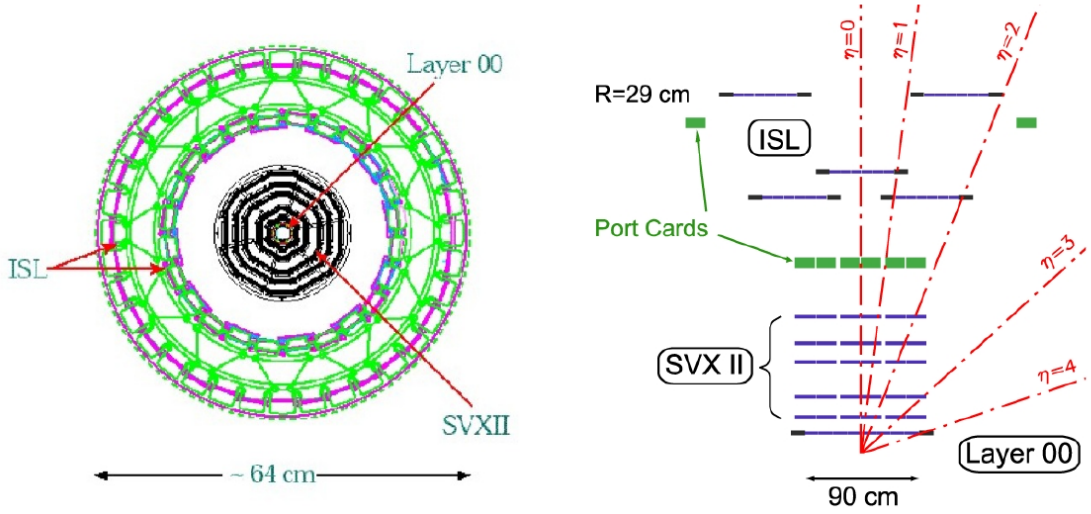
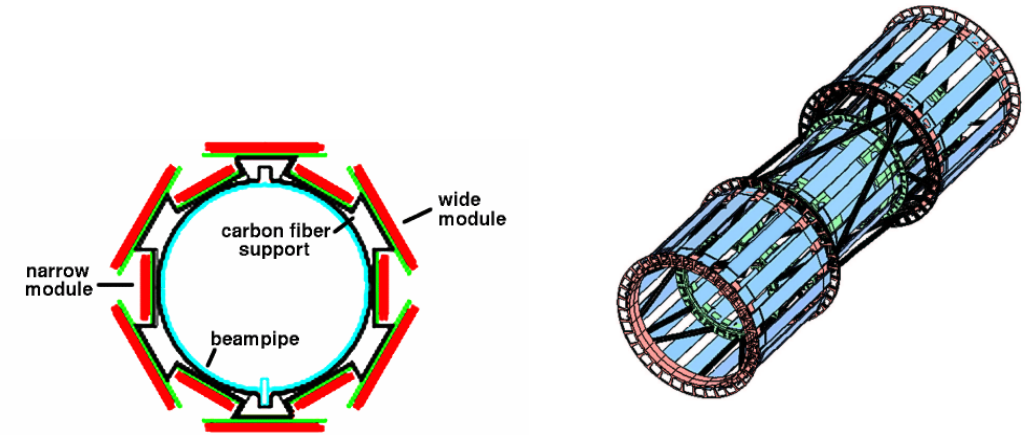


Figure 3.6: Left: cutaway transverse to the beam of the three subsystems of the silicon vertex system. Right: sketch of the silicon detector in a z/y projection showing the η coverage of each layer.

special sub-detectors are added to the silicon tracker: the Layer 00 (L00) [4, 5] and the Intermediate Silicon Layer (ISL), as illustrated in Fig. 3.6.

- **L00:** L00 is composed of a set of silicon strips assembled directly onto the beam pipe (Fig. 3.7(a)). This device has six narrow and six wide groups of ladder in ϕ at radii 1.35 and 1.62 cm respectively, providing 128 read out channels for the narrow groups and 256 channels for the wide groups. The silicon wafers are mounted on a carbon-fiber support which also provides cooling. L00 sensors are made of light-weight radiation-hard single-sided silicon (different from the ones used within SVX). Being so close to the beam, L00 allows to reach a resolution of $\sim 25/30$ μm on the impact parameter of tracks of moderate p_T, providing a powerful help to signal long-lived hadrons containing a b quark. L00 allows to overcome the effects of multiple scattering for tracks passing through high density regions of SVX thus making it possible to improve vertexing resolution.

- **ISL:** The Intermediate Silicon Layers (ISL) consist of double-sided silicon crystals: one side has axial microstrips to provide measurements in the r-ϕ plane, while the other one supplies z information by means of stereo strips. The arrangement of this device, shown in Fig. 3.7(b), varies according to the η range: in the central region ($|\eta| < 1$) it consists of a single layer placed at ~ 22 cm from the beam line, while for $1 < |\eta| < 2$ ISL is made of two layers placed at $r = 20$ and 29 cm respectively (see Fig. 3.3). The two layers at $1 < |\eta| < 2$ are important to help tracking in a region where the COT coverage is incomplete. In both regions, the stereo sampling enables a full three-dimensional stand-alone silicon tracking. The ISL is intended to improve the tracking resolution in the central region, while in the $1.0 <$



(a) Transverse view of Layer 00, the innermost silicon layer.

(b) Perspective view of ISL.

Figure 3.7: Layer 00 and ISL.

$|\eta| < 2.0$ region it provides a useful tool for silicon stand alone tracking in conjunction with SVX layers.

3.2.2 The COT

In addition to the silicon detector, the Central Outer Tracker (COT) [1] is located at larger radii, and is used both to improve the momentum resolution and to provide useful informations to the trigger system.

This system is installed in the region $|z| < 155\ cm$ and between the radii of 43 and 133 cm.

The COT is a cylindrical multi-wire open-cell drift chamber with a mixture of 50:35:15 Ar-Ethane-CF4 gas used as active medium. The COT contains 96 sense wire layers, which are radially grouped into eight "superlayers" (see Fig. 3.8). Each superlayer is divided in ϕ "supercells", and each supercell has 12 sense wires and it is designed so that the maximum drift distance is approximately the same for all supercells. Therefore, the number of supercells in a given superlayer scales approximately with the radius of the superlayer. Half of the 30,240 sense wires within the COT run along the z direction ("axial"), while the others are installed at a small angle (2°) with respect to the z direction ("stereo").

A charged particle passing through the gas mixture leaves a trail of ionization electrons. These electrons are carried towards sense wires of the corresponding cell. The electron drift direction is not aligned with the electric field, being affected by the 1.4 T magnetic field provided by the solenoid. Thus electrons originally at rest move in the plane perpendicular to the magnetic field forming an angle α with respect to the electric field lines. The value of α, the so-called Lorentz angle, depends on the magnitude of both fields and on the properties of the gas mixture. In the COT $\alpha \simeq 35°$(see Fig. 3.9).

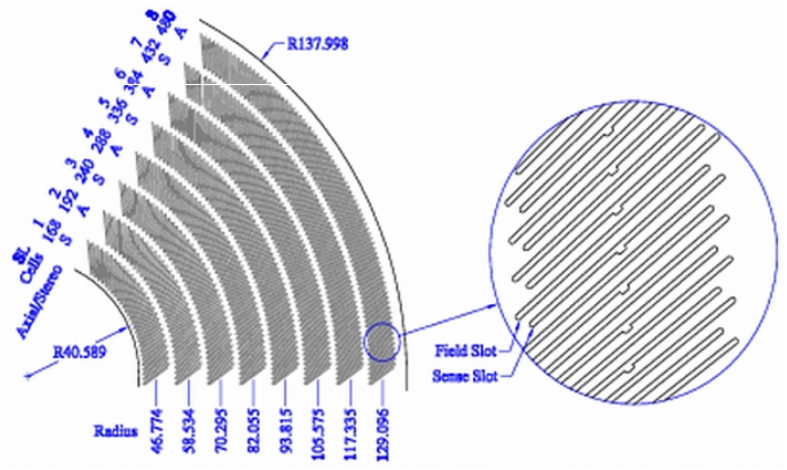


Figure 3.8: COT section: the eight superlayers (left) and the alternation of field plates and wire plates (right).

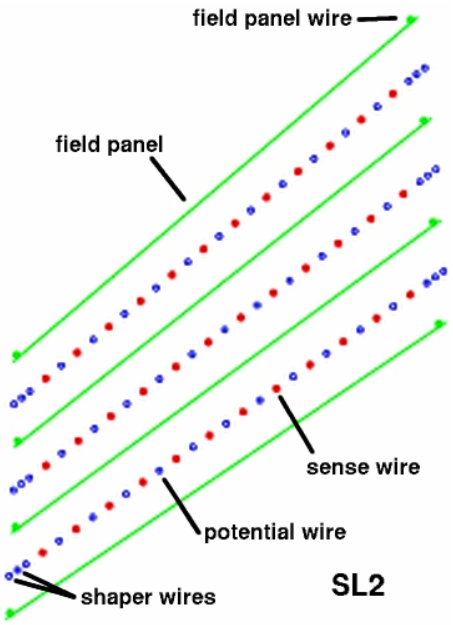


Figure 3.9: Cross-sectional view of some COT cells. The radial direction in the picture is horizontal and the angle between wire plane of the central cell and the radial direction is $\alpha \simeq 35°$.

The optimal situation in terms of resolution power is realized when the drift direction is perpendicular to that of the track. Usually the optimization is done for high P_T tracks, which are almost radial. As a result, all COT cells are tilted 35°away from the radial direction, so that the ionization electrons drift in the ϕ direction. When the electrons get near the sense wire, the local $\frac{1}{r}$ electric field accelerates them causing further ionization. The r - ϕ position of the track with

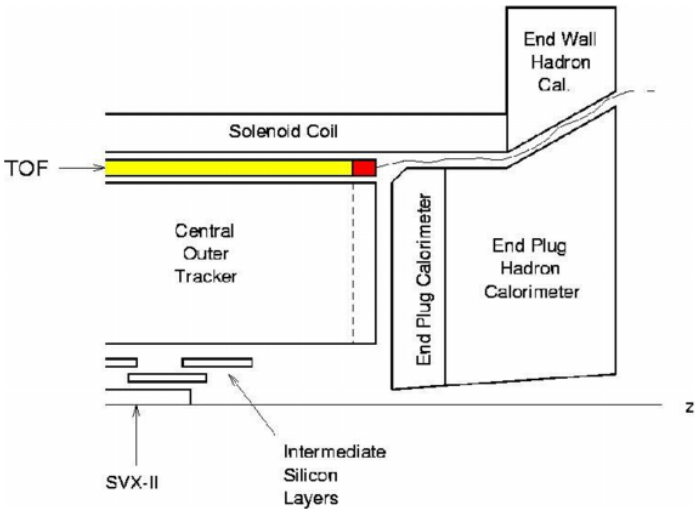


Figure 3.10: Time of Flight detector position.

respect to the sense wire is inferred by the signal arrival time.

A measurement of COT performance is given by the single hit position resolution and has been measured to be about 140 μm, which translates into the transverse momentum resolution $\frac{\delta p_T}{p_T} \sim 0.15\% \frac{p_T}{GeV/c}$.

3.2.3 Time of flight detector

The Time of Flight system (TOF) [4, 6] is a Run II upgrade to the CDF detector and it expands the particle identification capability of CDF II in the low P_T region. The TOF consists of 216 scintillator bars installed at a radial distance of about 138 cm from the z axis in the 4.7 cm space between the outer shell of the COT and the superconducting solenoid (see Figure 3.10). Bars are approximately 279 cm long and $4 \times 4\, cm^2$ in cross-section. With its cylindrical geometry TOF provides 2π coverage in ϕ, and covers the pseudorapidity range $|\eta| < 1.0$. Scintillator bars are read out at both ends by photomultiplier tubes, capable of providing adequate gain even if used inside the 1.4 T magnetic field. The TOF detector measures the arrival time t of a particle with respect to the collision time t_0. The mass m of a particle traversing the device is determined using the path length L and momentum P measured by the tracking system via the relationship

$$m = \frac{P}{c}\sqrt{\frac{(ct)^2}{L^2} - 1} \tag{3.2}$$

A resolution of $\sim$ 110 ps has been achieved which allows a 2σ separation of kaons from pions up to $\sim$ 1.6 GeV at $|\eta| < 1$.

	CEM	CHA	WHA	PEM	PHA
η coverage	< 1.1	< 0.9	$0.7 < \|\eta\| < 1.3$	$1.3 < \|\eta\| < 3.6$	$1.3 < \|\eta\| < 3.6$
n. of modules	48	48	48	24	24
η towers/mod	10	8	6	12	10
n. of channels	956	768	676	960	864
Absorber (mm)	Pb (3.0)	Fe (25.4)	Fe (50.8)	Pb(4.6)	Fe (50.8)
Thickness	$19X_0, 1\Lambda_0$	$4.5\Lambda_0$	$4.5\Lambda_0$	$21X_0, 1\Lambda_0$	$7\Lambda_0$
Position res.	0.2×0.2	10×5	10×5		
Energy res.	$\frac{13.5\%}{\sqrt{E_T}} \oplus 1.7\%$	$\frac{75\%}{\sqrt{E_T}} \oplus 3\%$	$\frac{80\%}{\sqrt{E_T}}$	$\frac{16\%}{\sqrt{E_T}} \oplus 1\%$	$\frac{80\%}{\sqrt{E_T}} \oplus 5\%$

Figure 3.11: Geometry, parameters and performance summary of CDF Calorimetric System. The position resolution is given in $r \cdot \phi \times z\ cm^2$ and is measured for a 50 GeV incident particle.

3.3 Calorimetric systems

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 as well as used to measure the missing transverse energy associated to neutrino production.

Particles hitting the calorimeter can be divided in two classes according to their interaction with matter: electromagnetically interacting particles, such as electrons and photons, and hadronically interacting particles, such as mesons or baryons 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, providing coverage up to $|\eta| < 3.6$. The calorimeter is also segmented in $\eta - \phi$ sections, called towers, projected towards the geometrical center of the detector, in order to supply information on particle positions. 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 calorimetric system is subdivided into three regions, central, wall and plug, in order of increasing pseudorapidity ranges, with the following naming convention: Central Electromagnetic (CEM), Central Hadronic (CHA), Wall Hadronic (WHA), Plug Electromagnetic (PEM) and Plug Hadronic (PHA); an inner commented view of the detector is shown in Fig. 3.12. Table in Fig. 3.11 summarizes the most important characteristics of each part of the calorimeter.

3.3.1 The Central Calorimeter

The Central Electro-Magnetic calorimeter (CEM) [1] is segmented in $\Delta\eta \times \Delta\phi$=0.11×15° projective towers consisting of alternate layers of lead and scintillator, while the Central and End Wall Hadronic calorimeters (CHA and WHA respectively), whose geometric tower segmentation matches the CEM one, use iron layers as radiators. A perspective view of a central electromagnetic calorimeter module, a *wedge*, is shown in Figure 3.13.

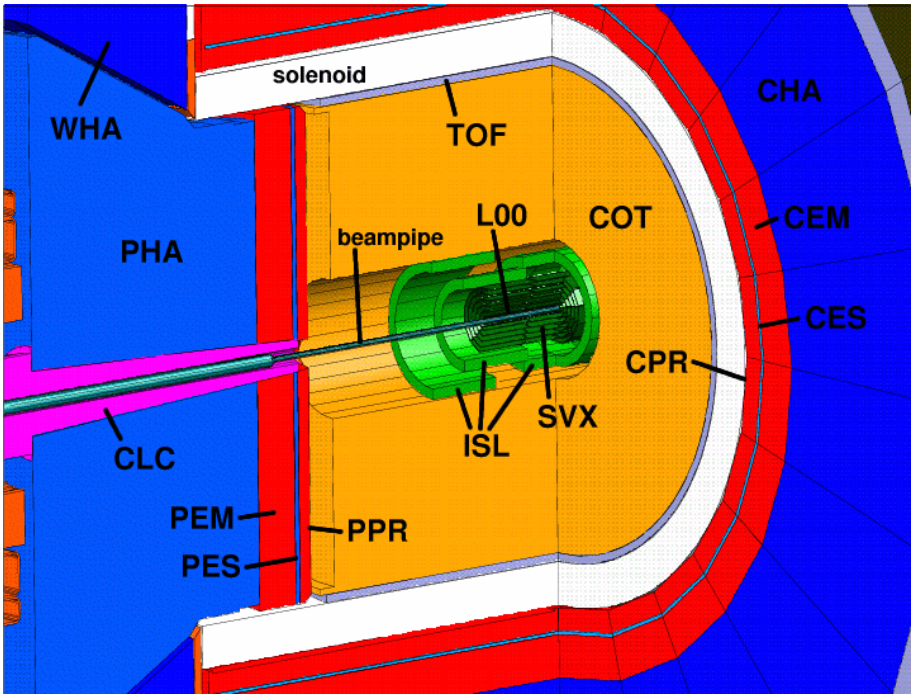


Figure 3.12: Schematic view of the inner parts of CDF detector.

Two position detectors are embedded in each wedge of the CEM:

- The Central Electromagnetic Strip chamber (CES) (see Fig. 3.14) is a two-dimensional strip/wire chamber located at the radial distance 184 *cm*. It measures the charge deposition of the electromagnetic showers, providing information on their pulse-height and position with a finer azimuthal segmentation than the calorimeter towers. This results in an increased purity of electromagnetic object reconstruction.

- The Central Pre-Radiator (CPR) consists of two wire chamber modules placed immediately in front of the calorimeter. It acts as pre-shower detector and with its 3072 channels collects charge deposit by showers originated by interaction of particles with tracking system and solenoid material. It can help in discriminating pions from electron and photons, because the latter deposit a greater amount of energy in the chamber.

3.3.2 The plug calorimeter

The plug calorimeter, shown in Fig. 3.15, covers the η region from 1.1 to 3.6. 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 segmented in 48 or 24 (for $|\eta| < 2.1$ or $|\eta| > 2.1$ respectively) projective towers.

As in the central calorimeter, there is a front electromagnetic compartment and a rear hadronic compartment (PEM and PHA). Projective towers consist of alternating layers of absorbing material (lead and iron for electromagnetic and hadronic

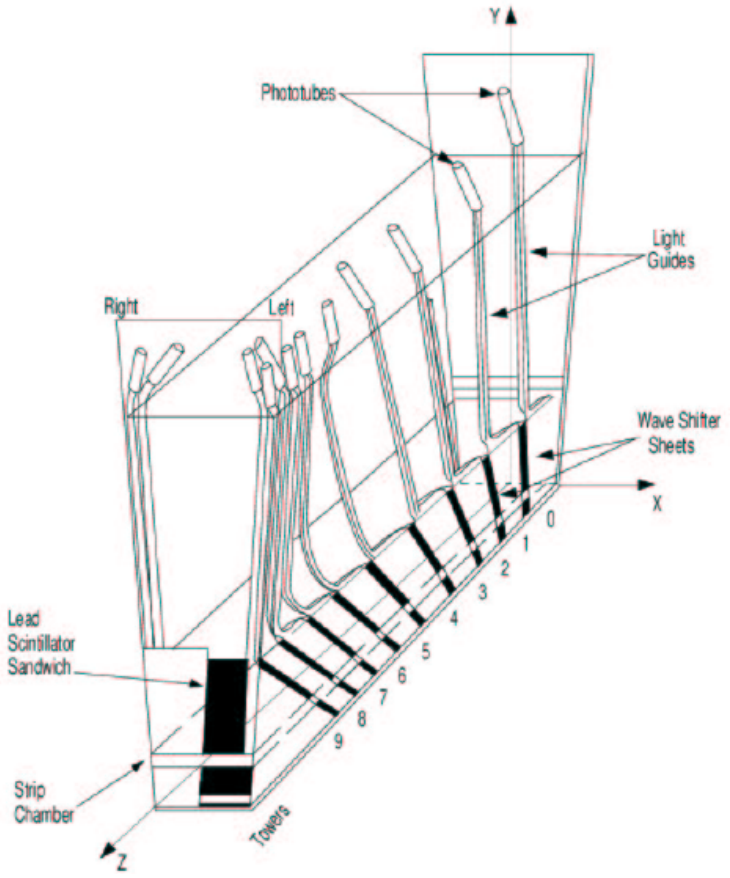


Figure 3.13: Perspective view of a CEM module.

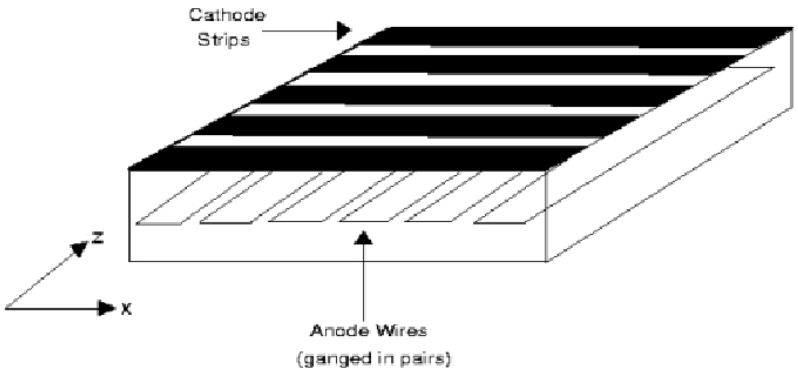


Figure 3.14: The CES detector in CEM. The cathode strips run in the x direction and the anode wires run in the z direction providing x and $(r \cdot \phi)$ measurements.

sectors respectively) and scintillator tiles. The first layer of the electromagnetic calorimeter acts as a pre-shower detector; to this scope, the first scintillator tile is

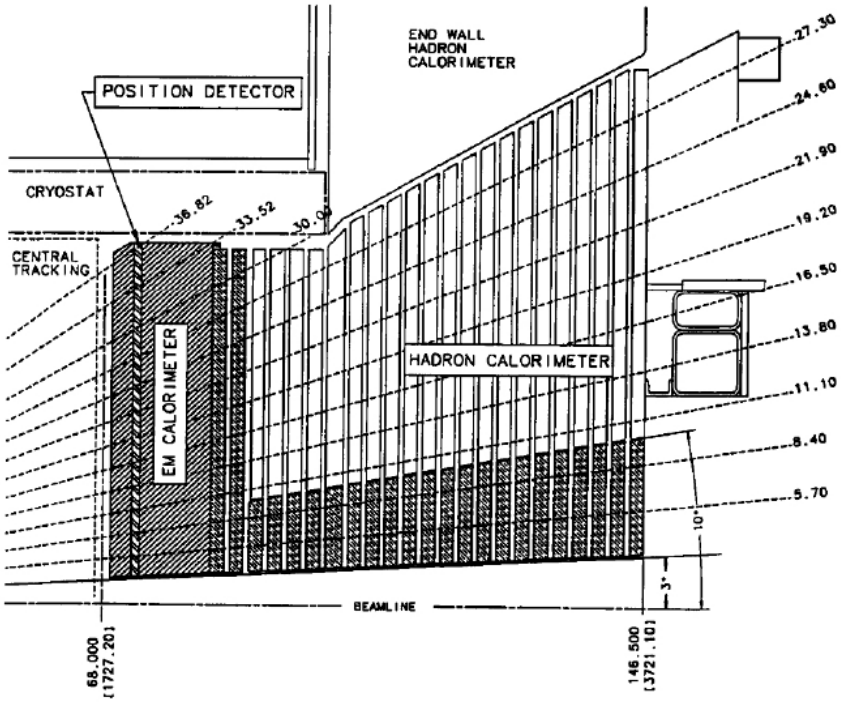


Figure 3.15: Plug Calorimeter (PEM and PHA) inserted in the Hadron End Wall calorimeter WHA and into the solenoid.

thicker (10 mm instead of 6 mm) and made of a brighter material.

As in the central calorimeter, a shower maximum detector (SMD) is also included in the plug electromagnetic calorimeter (PES). The PES consists of two layers of 200 scintillating bars each, oriented at crossed relative angles of 45^o ($\pm 22.5^o$ with respect to the radial direction). The position of a shower on the transverse plane is measured with a resolution of ~ 1 mm.

3.4 Muon detectors

Muons are highly penetrating, so they are separated from charged hadrons by the calorimeter, that acts as a shield for strongly and electromagnetic interacting particles.

Muon identification can then be performed by extrapolating the tracks outside the calorimeter and matching them to tracks segments (called *stubs*) reconstructed in an external muon detector.
Figure 3.16 provides an overview of the muon detectors coverage, that goes up to $|\eta| < 2.0$. Muon systems are divided in muon *chambers* and muon *scintillators*, see Fig. 3.17:

- Central MUon detector (CMU) consists of a set of 144 modules, each containing four layers of rectangular drift cells, operating in proportional mode. It

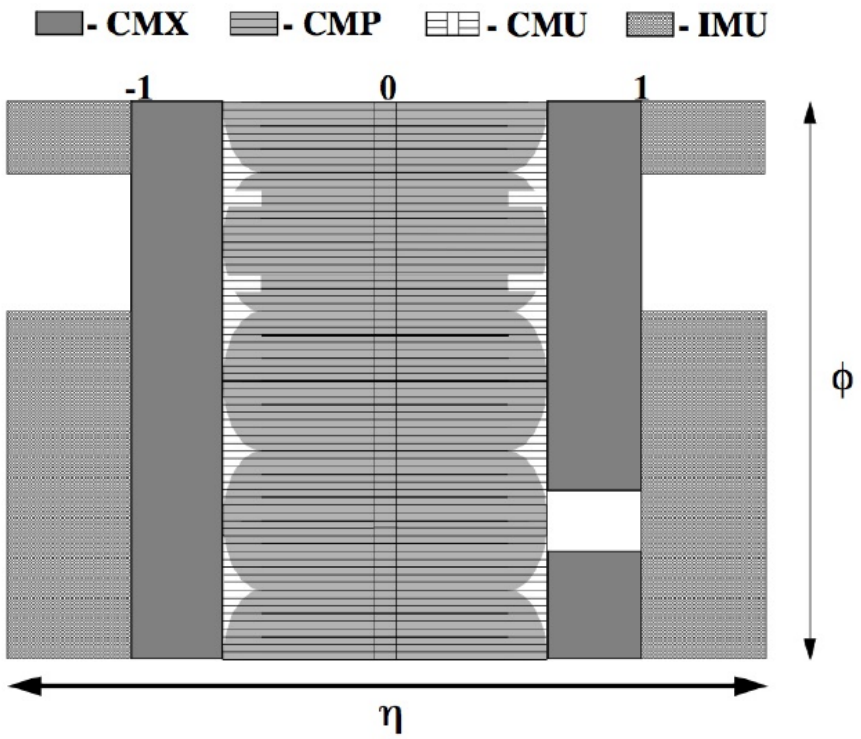


Figure 3.16: $\eta - \phi$ coverage of the Run II muon detector system. The shape is irregular because of the obstruction by systems such as cryo pipes or structural elements.

is placed immediately outside the calorimeter and supplies a global coverage up to $|\eta|$ <0.6.

- Central Muon uPgrade (CMP) consists of four layers of single-wire proportional drift tubes stagged by half cell per layer and shielded by an additional 60 *cm* steel layer. It is arranged in a square box around the CMU, providing a ϕ-dependent η coverage (see Figure 3.16).

- Central Scintillator uPgrade (CSP) is a layer of rectangular scintillator counters placed on the outer surface of CMP.

- Central Muon eXtension (CMX) consists of a stack of eight proportional drift tubes, arranged in conical sections to extend the CMU/CMP coverage in the $0.6 < |\eta| < 1$ region.

- Central Scintillator eXtension (CSX) consists of a layer of scintillator counters on both side of CMX. Thanks to scintillator timing, this device completes with z information the measurement of the muon position provided by CMX (ϕ).

- Intermediate MUon detector (IMU) consists of four staggered layers of proportional drift tubes and two layers of scintillator tiles, arranged as for the

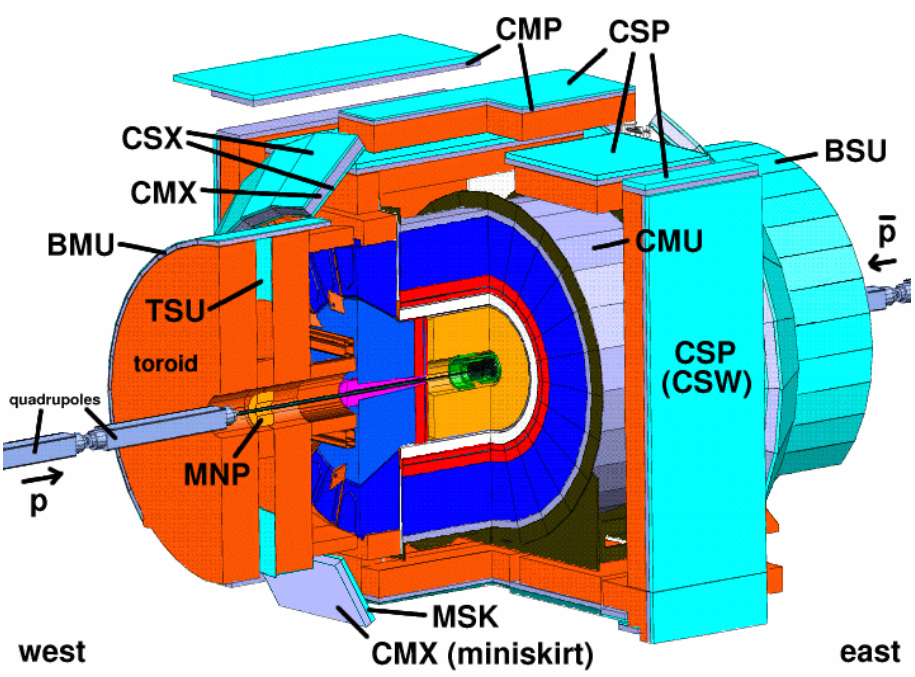


Figure 3.17: Schematic view of the whole CDF detector.

CMP/CSP system to extend triggering and identification of muons up to $|\eta| \leq 1.5$ and $|\eta| \leq 2$, respectively.

3.5 Cherenkov luminosity counters

CDF measures the collider luminosity with a coincidence between two arrays of Cherenkov counters, the CLC, placed around the beam pipes on the two detector sides [7]. The counters measure the average number of interactions per bunch crossing μ, which is used to provide a measurement of the instantaneous luminosity $\mathcal{L}$, by means of the following relation:

$$\mu \cdot f_{bc} = \sigma_{p\bar{p}} \cdot \mathcal{L}, \tag{3.3}$$

where $\sigma_{p\bar{p}}$ is the total $p\bar{p}$ cross section at $\sqrt{s} = 1.96\ TeV$ ($\sigma_{p\bar{p}} = 60.7 \pm 2.4\ mb$) and f_{bc} is the bunch crossing rate in the Tevatron. This method measures the luminosity with about the 6% systematic uncertainty. Each CLC module contains 48 gas Cherenkov counters of conical shape projecting to the nominal interaction point, organized in concentric layers. It utilizes Cherenkov radiation: particles traversing a medium at a speed higher than the speed of the light in the medium radiate light into a cone around the particle direction; the cone opening angle depends on the ratio of the two speeds and the refraction index of the medium. Taking into account that light produced by any particle originated at the collision point is collected with much higher efficiency than for background stray particles, the CLC signal is thus approximately proportional to the number of traversing particles produced in the collision.

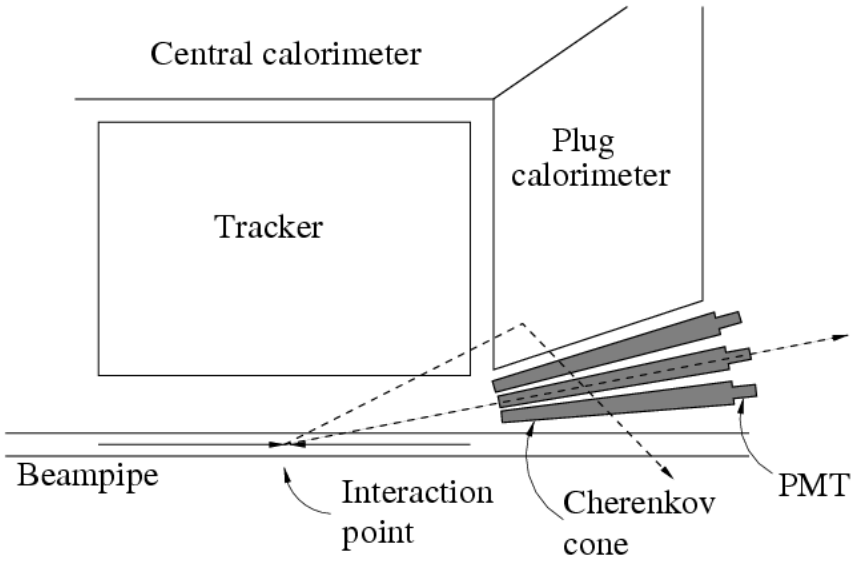


Figure 3.18: Schematic view of the luminosity monitor inside a quadrant of CDF.

3.6 Forward Detectors

CDF Forward Detectors (see scheme in Fig. 3.19) include the Roman Pots detectors (RPS), beam shower counters (BSC) and two forward Mini Plug Calorimeters (MP). These detectors enhance CDF sensitivity to production processes where the primary beam particles scatter inelastically in large impact parameter interactions.

The Tevatron complex allowed to arrange a proper spectrometer making use of the Tevatron bending magnets only on the antiproton side. On this side, at appropriate locations, scintillating fiber hodoscopes inside three RPS measure the momentum of the inelastically scattered antiproton. Only the direction of the scattered proton is measured on the opposite side. The BSC counters at $5.5 < |\eta| < 7.5$ measure the rate of charged particles around the scattered primaries.

The MiniPlugs calorimeters at $3.5 < |\eta| < 5.1$ measure the very forward energy flow. MiniPlugs are a single compartment integrating calorimeter, consisting of alternate layers of lead and liquid scintillator read by longitudinal wavelenght shifting fibers (WLS) pointing to the interaction vertex. Although the miniplug is not physically split into projective towers, its response can be split into solid angle bins in the off-line analysis. The MiniPlug energy resolution is about $\frac{\sigma}{E} = \frac{18\%}{\sqrt{E}}$ for single electrons.

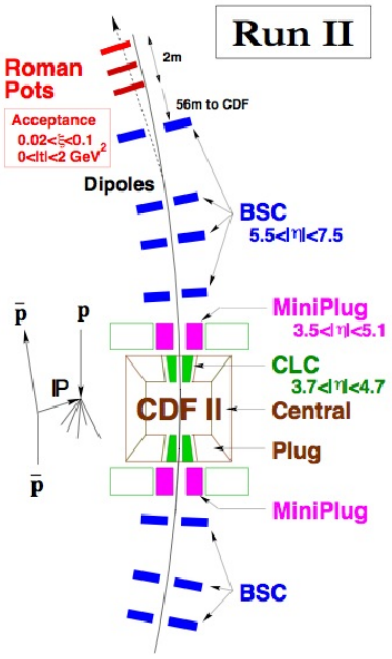
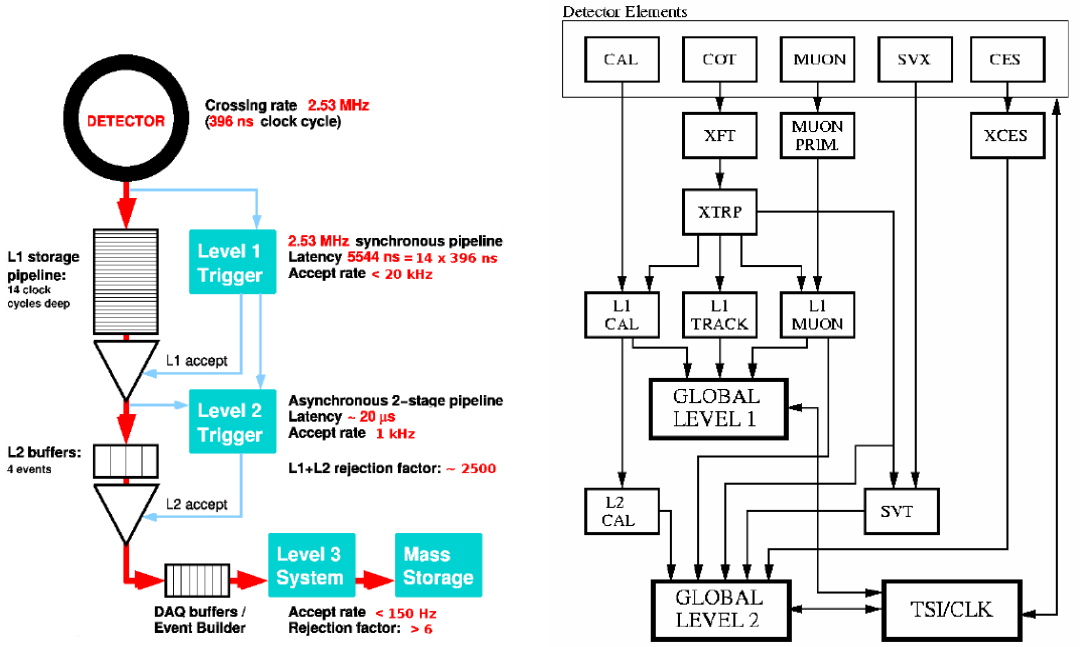


Figure 3.19: The forward detectors system in CDF, as arranged for Run II.

3.7 Trigger and data acquisition system

At hadron collider experiments the collision rate is much higher than the rate at which data can be stored on tape. At CDF II the predicted inelastic cross section for $p\bar{p}$ scattering is $60.7 \pm 2.4\ mb$, which, considering an instantaneous luminosity of order $10^{32}\ cm^{-2}s^{-1}$, results in a collision rate of about 6 MHz, while the tape writing speed is only of ~ 100 events per second. The role of the trigger is to efficiently select the most interesting physics events from the large number of *minimum bias* events. Events selected by the trigger system are saved permanently on a mass storage and subsequently fully reconstructed offline.

The CDF trigger system has a three-level architecture providing a rate reduction sufficient to allow more sophisticated event processing one level after another with minimal deadtime (see Fig. 3.20). The front-end electronics of all detectors is interfaced to a syncronous pipeline where up to 42 subsequent events can be stored for 5.5 μs while the hardware is taking a decision. If by this time no decision is made, the event is lost. Level 1 (L1) always occurs at a fixed time $< 4\ \mu s$ so that it doesn't cause any dead time. Using a custom designed hardware, L1 makes a raw reconstruction of physical objects and takes a decision after counting them. Events passing the L1 trigger requirements are then moved to one of four on-board Level 2 (L2) buffers. Each separate L2 buffer is connected to a two-step pipeline, each step having a latency time of 10 μs: in step one, single detector signals are analyzed, while in step two the combination of the outcome of step one are merged and trigger decisions are made. The data acquisition system allows a L2 trigger accept rate of $\sim 1\ kHz$ and a $L1 + L2$ rejection factor of about 2500. Events



(a) CDF readout functional diagram. (b) Block diagram of the CDF trigger system.

Figure 3.20: CDF trigger system.

satisfying both L1 and L2 requirements are transferred to the Level 3 (L3) trigger processor farm where they are reconstructed and filtered using the complete event information, with an accept rate $< 150\ Hz$ and a rejection factor > 6, and then finally written to permanent storage.

According to the signal one wants to isolate, specific sets of requirements are established by exploiting the physics objects (*primitives*) available for each trigger level. Successively, links across different levels are established by defining *trigger paths*: a trigger path identifies a unique combination of a L1, a L2, and a L3 trigger; datasets (or data streams) are then finally formed by merging the data samples collected via different trigger paths.

3.7.1 Level 1 primitives

Tracks

The most significant tool for L1 trigger is the possibility of track finding by means of a hardwired algorithm named e**X**tremely **F**ast **T**racker (XFT). The XFT has been designed to work with COT signals at high collision rates, returning track P_T and ϕ_0 by means of a fast r-ϕ reconstruction. These tracks are then extrapolated to the central calorimeter wedges and to the muon chambers (CMU and CMX), allowing a track to be matched to an electromagnetic calorimeter cluster for a first electron identification, or to a stub on the muon detectors for improved muon reconstruction, and tracks to be used alone for specific triggers.

Calorimetric primitives

At L1 calorimetric towers are merged in pairs along η to define *trigger towers*, which are the basis for two types of primitives:

- object primitives: electromagnetic and hadronic transverse energy contributions are used to define electron/photon and jet primitives respectively;
- global primitives: transverse energy deposits in all trigger towers above 1 GeV are summed to compute event ΣE_T and $\not\!\!E_T$.

Correspondingly, object and global triggers can be defined by applying a threshold to the respective primitives.

Leptons

As already mentioned above, L1 muon and electron triggers are obtained by matching a XFT track to a corresponding primitive: for electrons, primitives are essentially the calorimetric trigger towers described above, while for muons they are obtained from clusters of hits in the muon chambers.

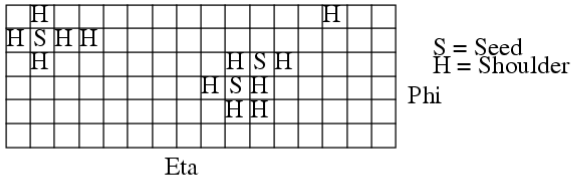
3.7.2 Level 2 primitives

L2 trigger takes a decision on a partially reconstructed event, exploiting data collected from L1 and from the calorimeter shower maximum detectors. Simultaneously a hardware cluster finder processes data from calorimeters while a track processor finds tracks in the silicon vertex detector.

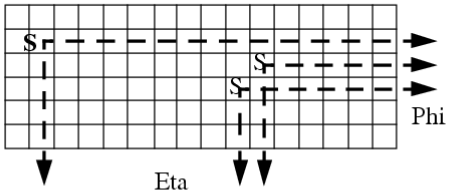
Calorimeter clusters

Since jets are expected not to be fully contained into a single calorimeter trigger tower, the energy threshold on L1 jet primitives must be set much lower than the typical jet energy in order to maintain high selection efficiency. As a consequence, jet trigger rates are too high to be fed directly into L3. An effective rate reduction can be obtained at L2 by triggering both on multiplicity and transverse energy of trigger tower clusters. The algorithm for cluster finding is based on the four-step procedure described in Fig. 3.21:

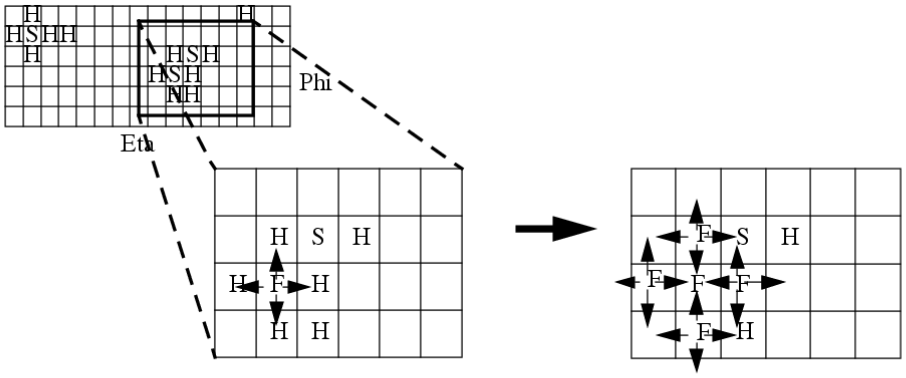
- electromagnetic and hadronic transverse energy of the trigger towers are checked to see if they are above predetermined *seed* and *shoulder* thresholds;
- all trigger towers whose energy has been found above the seed threshold are ordered according to increasing ϕ and η values.
- Cluster finding begins with the first seed tower. The four orthogonal nearest towers are considered: if their energy is above the shoulder threshold, they are merged to the cluster and their orthogonal neighbors are in turn considered.



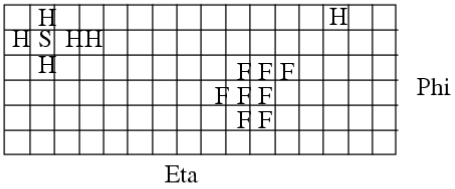
1. The inputs from DIRAC are checked to see if they are over the seed or shoulder threshold.



2. The Cluster Control board tells the DCAS's to enable their Phi and Eta address lines for all seeds. The Control board then selects the seed with the lowest Eta address, and then within that Eta, the lowest Phi address to begin cluster finding.



3. Cluster finding then begins with the selected seed sending a "found" bit to its four orthogonal neighbors. If the neighbors are over the shoulder threshold and receive a "found" bit, they in turn send out the signal to their four orthogonal neighbors, and so on, and so on.



4. When a tower is "found" it is disabled from being found in another cluster. i.e. if the tower is a seed and is found in a cluster initiated by another seed, it will not enable its address lines when the control board prompts for seed addresses. After one cluster is found, the tower energy values are enabled into an adder tree. Steps two through four are repeated until all clusters are found.

Figure 3.21: Level 2 calorimeter cluster finding procedure.

- Towers merged in the cluster are disabled from being merged into another cluster. When no other tower is found to be added to the cluster, tower energy values are summed to define cluster E_T and a new clustering procedure starts with the successive seed tower.

L2 clusters can be used to build object triggers by applying a cut on their transverse energy and position (provided from η-ϕ address of the seed towers), and global

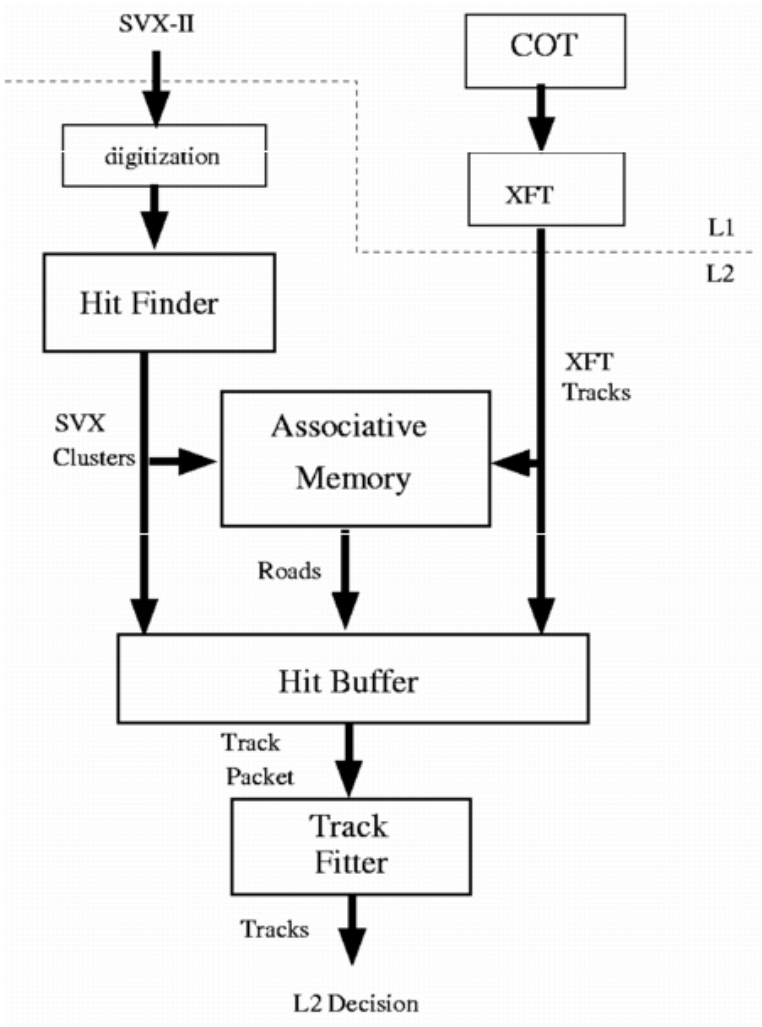


Figure 3.22: The SVT architecture.

triggers by selecting on the number and $\sum E_T$ of clusters.

SVT tracks

One of the most significant tools for the L2 trigger system is the **S**ilicon **V**ertex **T**racker (SVT) [8] which exploits the potential of a high precision silicon vertex detector to trigger on tracks with large impact parameter: this can allow to detect secondary vertexes and to study a large number of processes involving decays of b-hadrons with a long lifetime.

The architecture of SVT is shown in Fig. 3.22. Its inputs are the list of axial COT tracks found by XFT and the data from SVXII. First SVXII hits are found by a Hit Finder algorithm and stored in hit buffers; then association between XFT and SVXII tracks is performed by *Associative Memory* (AM), a massive parallel mechanism based on the search of *roads* among the list of SVXII hits and XFT tracks; a road is a coincidence between hits on four of the silicon layers and XFT tracks. Upon receiving a list of hits and tracks, each AM chip checks if all the components of one of its roads are present in the list of hits and XFT tracks. When

AM has determined that a road might contain a track, hits belonging to that road are retrieved from the input buffer and passed to a track fitter to compute track parameters.

Leptons

L2 muon primitives are essentially unchanged with respect to L1, the only difference consists in an improved ϕ-matching (within 1.25^o) between XFT tracks and stubs. In the case of electrons, a finer ϕ-matching can be instead performed at L2 thanks to the information from central and plug shower maximum detectors.

3.7.3 Level 3 primitives

The L3 trigger is a software trigger that runs on a Linux PC farm where all events are almost fully reconstructed using C++ codes and object-oriented techniques. In particular jets, COT tracks and leptons are identified. The algorithms used for the reconstruction are the same used in offline analysis. Events coming from L2 are addressed to the *Event Builder* (EVB), which associates information on the same event from different detector parts. Some variables, like global kinematic event observables, cannot be computed due to the long processing time required. Other tasks, like a full track reconstruction, could be possible only on subsets of data passing low-rate triggers. The final decision to accept an event is made on the basis of its features of interest (large E_T leptons, large missing E_T, large energy jets and a combination of such) for a physics process under study, as defined by the trigger path tables containing up to about 150 entries. Events exit L3 at a rate up to about 100 Hz and are permanently stored on tapes for further offline analisys. Each stored event is about 250 kB large on tape. Further offline processing is then performed on the selected events.

3.7.4 Trigger Upgrades

CDF has recently undergone two major trigger upgrades in order to deal with high trigger rates with increasing luminosity and to augment signal acceptance: an XFT upgrade and an upgrade in L2CAL system [9, 10].

XFT upgrade regards both Level 1 (L1) and Level 2 (L2) trigger systems. At L1 it rejects fake axial tracks by requiring the association with *stereo* segments, with a rejection factor of about 7. Moreover XFT segments of finer granularity can be sent to L2 where a 3D-track reconstruction can be performed with a good resolution on $cot\theta$ ($\sigma_{cot\theta}$= 0.12) and z_0 (σ_{z_0} = 11 cm).

The upgraded L2CAL system uses a *fixed cone* cluster finding algorithm which prevents fake cluster formation and exploits full 10-bit trigger tower energy information for $\not\!\!E_T$ and ΣE_T calculation (the old system, due to hardware limitations, used only 8-bit tower information). A jet is formed starting from a *seed* tower above a 3 GeV threshold and adding all the towers inside a fixed cone centered at the seed tower and having a radius $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.7$ units in the azimuth-pseudorapidity space. Jet position is calculated weighting each tower inside the

cone according to its transverse energy. This upgrade has reduced L2 trigger rate and has provided at L2 jets nearly equivalent to offline ones.

3.8 Offline data processing

The raw data flow from L3 triggers, segmented into streams according to trigger sets tuned to a specific physics process, is then stored on fast-access disks in real time (on-line), as the data are collected. All other manipulations with data are referred to as off-line data handling. The most important of these operations is the so-called "production" which stands for the complete reconstruction of the collected data. At this stage raw data banks are unpacked and physics objects suitable for analysis, such as tracks, vertices, leptons and jets are generated. The procedure is similar to what is done at L3, except that it is done in a much more elaborate fashion, applying the most up-to-date detector calibrations, using the best measured beamlines, etc. The output of the production is further categorized into datasets which are used as input to physics analyses. Occasionally, if more detailed calibrations or significantly improved codes become available, data are re-processed. Re-processing is an heavy computer time-consuming operation which is performed only when a significant gain in reconstructed event quality is expected. For the analysis performed in the present work, the reconstruction code versions 5.3.3_nt5 and 6.1.4 were used.

Bibliography

[1] F. Abe *et al.* [The CDF II Collaboration], *The CDF II Technical Design Report*, FERMILAB-PUB-96-390-E (1996), http://www-cdf.fnal.gov/upgrades/tdr/tdr.html.

[2] *An Ace's Guide to the CDF Run II Detector*, http://www-cdfonline.fnal.gov/ops/ace2help/detector_guide/detector_guide.html.

[3] *Lecture Series on CDF Detectors and Offline*, http://www-cdf.fnal.gov/internal/spokes/lecture_detector.html.

[4] The CDF II Collaboration, *Proposal for Enhancement of the CDF II Detector: An Inner Silicon Layer and A Time of Flight Detector*, Fermilab-Proposal-909 (1998).

[5] T.K. Nelson, *The CDF Layer 00 Detector*, CDF Public Note 5780, FERMILAB-CONF-01/357-E.

[6] C. Grozis *et al.*, *A Time-Of-Flight Detector for CDF*, International Journal of Modern Physics A Vol.**16**, Suppl.1C (2001).

[7] D. Acosta *et al.* [The CDF II Collaboration], *The CDF Cherenkov luminosity monitor*, Nucl. Instrum. Meth. A **461** (2001) 540.

[8] S. Belforte *et al.*, *SVT - Silicon Vertex Tracker - Technical Design Report*, CDF Internal Note 3108.

[9] C.Cox *et al.*, *XFTL2 Stereo Upgrade Overview*, CDF Internal Note 8975.

[10] A.Canepa *et al.*, *L2CAL - The L2 Calorimeter Trigger Upgrade*, CDF Internal Note 8940.

Chapter 4

Reconstruction of Physical Objects

In this chapter we will describe how particles produced in a $p\bar{p}$ collision are reconstructed starting from the raw outputs of the different parts of the detector. First we will see how information from silicon detectors and COT are used to reconstruct charged particle trajectories. Then we will move to the reconstruction of jets of hadronic particles, based on calorimeters, analyzing the corrections of jet energies for different error sources introduced by calorimeters and reconstruction algorithms. Then we will give a brief description of the identification procedures for leptons and photons, and of the method used at CDF to identify a jet of particles originated from a b quark.

4.1 Track reconstruction

Track reconstruction is performed using data from silicon tracking system and COT. The reconstruction is based on the position of the hits leaved by charged particles on detector components. Several algorithms have been developed in order to reconstruct tracks: a tracking algorithm can use either COT or silicon detector only information, or can rely on information provided by the complete tracking systems.

In any case, track reconstruction requires an excellent alignment between COT and silicon detectors, since the global CDF II coordinate system is anchored to the center of the COT. Positions of other detector components are measured with respect to COT reference frame and encoded in so-called *alignments tables*.

We remind that the whole tracking system is immersed in a 1.4 T magnetic field, causing charged particles moving trough it to describe a helix trajectory, whose axis is parallel to the magnetic field. Measuring the radius of curvature of the helix, one can obtain the particle's transverse momentum P_T, while the longitudinal momentum is related to the helix pitch. Particle trajectories can be completely described by the following parameters [1]:

- z_0 : the z coordinate of the closest point to the z axis;
- d_0 : the impact parameter, defined as the distance between the point of closest approach to z axis and the z axis;

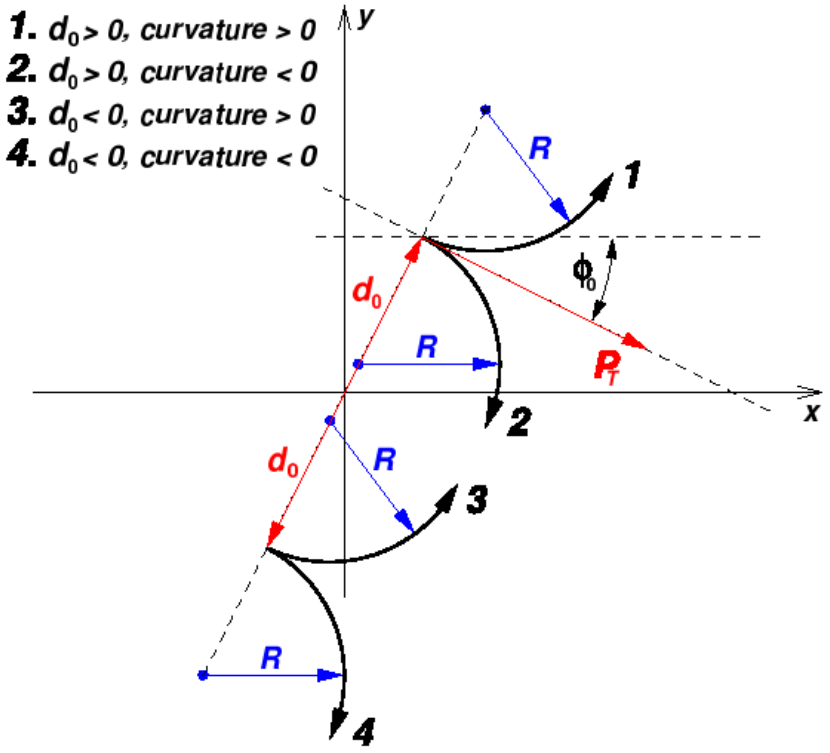


Figure 4.1: Illustration of track helix parameterization.

- ϕ_0 : the ϕ direction of the transverse momentum of the particle (tangential to the helix) at the point of the closest approach to the z axis;
- $\cot\theta$: the helix pitch, defined as the ratio of the helix step to its parameter;
- C : the helix curvature, defined as $C = \frac{q}{2R}$, where q is the charge of the particle and R is the radius of the helix.

Starting from helix parameters, particle transverse and longitudinal momenta can be calculated as:

$$\begin{aligned} P_T &= \frac{cB}{2|C|} \\ P_z &= P_T \cot\theta \end{aligned} \tag{4.1}$$

Track parameters and the relation between particle charge sign and impact parameter are illustrated in Fig. 4.1.

4.1.1 Outside-In tracking

The standard CDF track reconstruction is performed by the so called *Outside-In* algorithm [2], that exploits information from both COT and silicon detectors. The process starts by considering tracks reconstructed with information provided by the drift chamber (COT) alone, and by extrapolating them through the Silicon

Detector, where additional hits can be used for the final determination of track parameters.

Track reconstruction in the COT begins by finding track segments or just individual hits in the axial superlayers; matched segments and hits are then used to produce a track candidate.

Tracking in the COT starts translating the measured drift times in hits positions; once all COT hit candidates in the event are known, the eight superlayers are scanned looking for line segments. A line segment is defined as a triplet of aligned hits which belong to consecutive layers. A list of candidate segments is formed and ordered by increasing slope of the segment with respect to the radial direction so that high transverse momentum tracks will be given precedence. Once segments are available, the tracking algorithm tries to assemble them into tracks. At first, axial segments are joined in a 2D track and then stereo segments and individual stereo hits are attached to each axial track. Outside-In algorithm takes COT tracks and extrapolates them into the silicon detectors, adding hits via a progressive fit.

As extrapolation proceeds from the outermost SVX layer towards the beampipe (going from the outside in), the track error matrix is updated to reflect the amount of scattering material traversed. At each SVX layer, hits that are within a certain radius are appended to the track which is then re-fitted. A new track candidate is generated for each of the newly appended hits, but only the best two candidates (in terms of the fit quality and the number of hits) are considered for the next reconstruction steps. Each of these candidates is extrapolated further in, where the process is repeated. In the end there may still be several candidates associated to the original COT-only track. In this case the best one in terms of the number of hits and in terms of fit quality is retained.

4.1.2 Inside-Out algorithm

Although the *Outside-In* algorithm can achieve high performance in the central detector region, it looses efficiency in the forward region. For this region another tracking algorithm, named *Inside-Out* [3], has been developed.

This algorithm essentially works in a reverse mode with respect to the Outside-In one: it uses silicon stand alone reconstructed tracks to define a search road through the COT chamber.

Standalone tracking consists in finding triplets of aligned 3D hits, extrapolating them and adding matching 3D hits on other layers. This technique is called standalone because it doesn't require any input from outside: it performs tracking completely inside the silicon detector. First the algorithm builds 3D hits from all possible couples of intersecting axial and stereo strips on each layer. Once a list of such hits is available, the algorithm searches for triplets of aligned hits. This search is performed fixing a layer and doing a loop on all hits in the inner and outer layer with respect to the fixed one. For each hit pair - one in the inner and one in the outer layer - a straight line in the $r - z$ plane is drawn. Next step consists in examining the layer in the middle: each of its hits is used to build a helix together with the two hits of the inner and outer layers.

The triplets found so far are track candidates. Once the list of candidates is

complete, each of them is extrapolated to all silicon layers looking for new hits in the proximity of the intersection between candidate and layer. If there is more than one hit, the candidate is cloned and a different hit is attached to each clone. Full helix fits are performed on all the candidates. The best candidate in a clone group is kept, the others are rejected.

Inside-Out tracks can be used in conjunction to the standard *Outside-In* tracks to increases the detector tracking capabilities.

Precise determination of track parameters allows to discern which track come from what vertex and thereby to distinguish the *primary* vertex (PV) from the possible *secondary* vertices (SV) originated by long-lived particle decays, such as B hadrons.

4.2 Primary vertex reconstruction

The position of the interaction point of the $p\bar{p}$ collision (primary vertex) is of fundamental importance for event reconstruction. At CDF two algorithms can be used for primary vertex reconstruction. One is called *PrimVtx* [4] and is used, as an example, in b quark identification. PrimVtx starts from the beamline z-position (seed vertex) measured during collisions and then proceeds through an iterative algorithm that combines all the information on the reconstructed tracks.

The following cuts (with respect to the seed vertex position) are applied to the tracks:

- $|z_{trk} - z_{vertex}| < 1.0\ cm$;
- $|d_0| < 1.0\ cm$, where d_0 denotes the track impact parameter;
- $\frac{d_0}{\sigma} < 3.0$, where σ is the error on d_0.

Tracks surviving the cuts are ordered in decreasing P_T and used in a P_T-weighted fit to a common vertex. Tracks with χ^2 relative to the vertex greater than 10 are removed and the remaining ones are fit again to a common point. This procedure is iterated until no tracks have $\chi^2 > 10$ relative to the vertex.

The resulting resolution on the primary vertex position in the transverse plane ranges from 6 to 26 μm, depending on the topology of the event and on the number of tracks used in the fit. It is a significant improvement over the beam spot ($\sim 35\mu m$) information alone, and it provides the benchmark to secondary vertex searches for heavy flavour jets tagging. Finally, the z coordinate of the primary vertex is used to define the actual pseudorapidity of each physics object reconstructed in the event.

The second vertex finding algorithm developed in CDF is *ZVertexColl* [5]. This algorithm starts from pre-tracking vertices, i.e. vertices obtained from tracks passing minimal quality requirements. Among these, a lot of fake vertices are present: ZVertexColl cleans up these vertices requiring a certain number of tracks with $P_T > 300\ MeV$ be associated to them. A track is associated to a vertex if it is within 1 cm from a silicon standalone vertex (or 5 cm from a COT standalone vertex), where a vertex is considered "standalone" if it is reconstructed completely

inside a single detector - silicon detector or COT - without any input from other detectors.

The vertex position z is calculated from tracks positions z_0 weighted by the error σ:

$$z = \frac{\sum_i \frac{z_{0_i}}{\delta_i^2}}{\sum_i \frac{1}{\delta_i^2}} \tag{4.2}$$

Vertices found by ZVertexColl are classified by quality flags according to the number of tracks with silicon/COT tracks associated to the vertex. Associated COT tracks have shown to reduce the fake rate of vertices thus higher quality is given to vertices with COT tracks associated:

- Quality 0: all vertices;
- Quality 4: ≥ 1 track with COT hits;
- Quality 7: ≥ 6 tracks with silicon hits, ≥ 1 track with COT hits;
- Quality 12: ≥ 2 tracks with COT hits;
- Quality 28: ≥ 4 tracks with COT hits;
- Quality 60: ≥ 6 tracks with COT hits.

4.3 Jet reconstruction

In general jets are the results of the fragmentation process of partons outcoming from $p\bar{p}$ collision, see Fig. 4.2. The fragmentation yields a stream of energetic, colorless, spatially collimated particles along the original parton direction.

Jets are observed as clusters of energy located in adjacent calorimetric towers. Depending on the nature of the particles contained in a jet, energy deposit can be detected in the electromagnetic and/or hadronic sectors of the calorimeters.

The reconstruction procedure, named *jet clustering*, is based on the algorithm *JetClu* [6]; it starts with *preclustering* by identifying a list of seed towers (i.e. towers having $E_T \geq 1\ GeV$) and assigning a vector in the (r, η, ϕ)-space whose module is defined by the tower transverse energy content.

The vector origin is set in the interaction point, while its direction points towards the energy barycenter of the tower. The barycenter is defined assuming that all energy has been released at the average depth computed for CDF calorimeter (6 radiation lengths, X_0, and 1.5 interaction lengths, λ, for electromagnetic and hadronic sectors respectively).

Preclusters are created by combining adjacent seed towers within a preselected window in the $\eta - \phi$ plane. Starting from the highest E_T seed, the algorithm incorporates into the precluster the adjacent seed towers within the window and removes them from the list. The process is iterated by adding the seeds adjacent to the previous ones until no new such seeds are found in the window (see Fig. 4.3).

The jet reconstruction algorithm at CDF continues using the energy depositions in the calorimetric towers in a *fixed opening cone*. The opening of the cone is usually defined in terms of a radius in the $\eta - \phi$ plane, $R_{cone} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, and

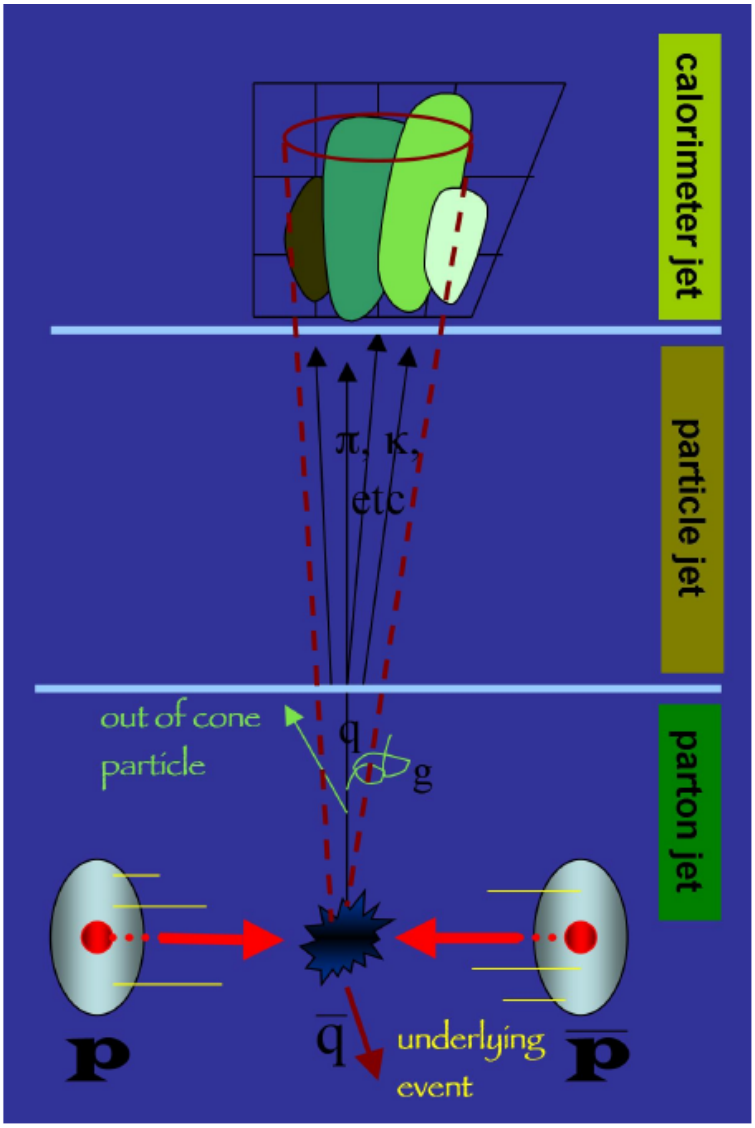


Figure 4.2: The scheme shows the development of a jet from parton level to particle level and to detector level.

has to be chosen according to the topology and the characteristics of the physical process to be studied: in high jet multiplicity events, a small cone radius (typically 0.4) is preferred, in order to avoid jet overlapping, on the contrary higher cone radii are chosen for the reconstruction of low jet multiplicity events in order to ensure the most of the energy flow to be contained therein. A fixed radius cone is drawn around each pre-cluster in the $\eta-\phi$ plane, whose axis is the vector with maximum module. All vectors falling inside a cone are grouped together, their energies summed up and the pre-cluster axis is re-estimated. This step is repeated until all vectors with $E_T > 100\ MeV$ are assigned to a cone.

Then E_T is calculated by assigning a massless four-vector with magnitude equal to the energy deposited in the tower, with a direction defined by a unit vector pointing from the center of the detector to the center of the calorimetric tower.

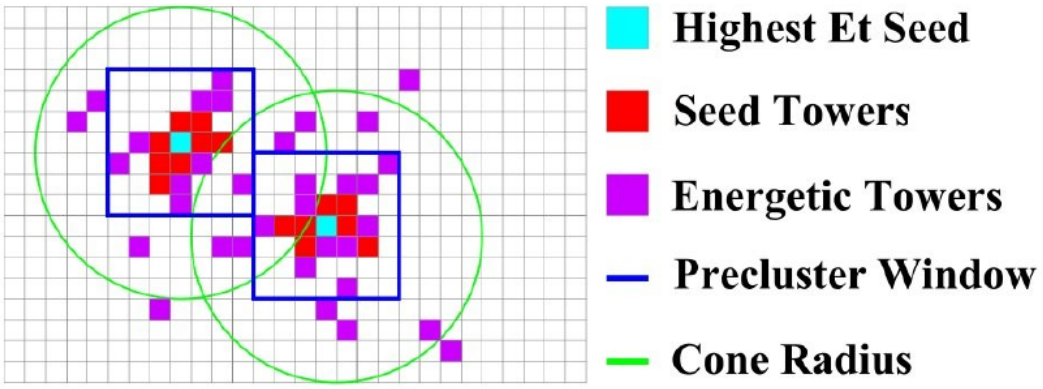


Figure 4.3: Prelustering of two jets on the $\eta-\phi$ plane. The green circles are the projections of the jet cone on the $\eta-\phi$ plane.

The center of the cluster is calculated according to the following definitions:

$$\begin{aligned} E_T^{jet} &= \sum_{i=1}^{N} E_T^i \\ \eta^{jet} &= \sum_{i=1}^{N} \frac{E_T^i \eta_i}{E_T^i} \\ \phi^{jet} &= \sum_{i=1}^{N} \frac{E_T^i \phi_i}{E_T^i} \end{aligned} \tag{4.3}$$

where N is the number of towers associated to the cluster and $E_T^i = E^i \sin\theta^i$ is the transverse energy of the i-th tower with respect to the z-position of the $p\bar{p}$ interaction.

This procedure is repeated iteratively with the jet E_T and direction being recalculated until the list of towers assigned to the clusters is stable. If two jets overlap, a decision has to be taken: if more than 50% of the transverse energy of the less energetic one is common, the two cones are replaced by a single one, centered around the sum of their resultants. Otherwise, the two jets are kept distinguished, and common vectors are assigned to the closest cone in the η-ϕ plane.

At the end of the procedure, the jet four-momentum ($E_{T,jet}$, $P_{x,jet}$, $P_{y,jet}$, $P_{z,jet}$),

is calculated using the final list of towers associated to the cluster:

$$
\begin{aligned}
E_{jet} &= \sum_{i=1}^{N} E_i \\
P_{x,jet} &= \sum_{i=1}^{N} E_i \sin\theta_i \cos\phi_i \\
P_{y,jet} &= \sum_{i=1}^{N} E_i \sin\theta_i \sin\phi_i \\
P_{z,jet} &= \sum_{i=1}^{N} E_i \cos\theta_i \\
P_{T,jet} &= \sqrt{P_{x,jet}^2 + P_{y,jet}^2} \\
\phi_{jet} &= \tan\frac{P_{y,jet}}{P_{x,jet}} \\
\sin\theta_{jet} &= \frac{P_{T,jet}}{\sqrt{P_{x,jet}^2 + P_{y,jet}^2 + P_{z,jet}^2}} \\
E_{T,jet} &= E_{jet}\sin\theta_{jet}
\end{aligned}
\tag{4.4}
$$

Jet quadrimomentum explained so far is computed starting from raw calorimetric energies.

4.3.1 Jet corrections

Jet energies measured in calorimeters suffer from the intrinsic limits of both calorimeters and jet reconstruction algorithm. Raw energies differ from real deposited energies, thus jet four-momenta need to be corrected, as we will discuss in this section.

A lot of factors can contribute to mis-measurements of the real parton energies:

- Some particles can fall outside the cone of the reconstructed jet causing an under-estimation of the energy measurement (*out-of-cone energy*).
- Particles like muons, whose energy is not completely detected, or neutrinos, which escape from the calorimeter, can be present in the jet, causing energy mismeasurements.
- The calorimeter coverage of the detector is imperfect, and there are some un-instrumented detector regions (so-called *cracks*) that can contribute to the degradation of the energy measurement.
- Calorimeter response can be non-homogeneous for particles hitting different regions of the detector.
- Strong interactions involving beam remnants (*underlying event*) or due to *multiple interactions* in the same bunch crossing can produce soft hadrons interfering with the jet clustering procedure.

Level	Type of jet correction
Level 0	Calorimeter energy scale setting
Level 1	η-dependent correction, f_η
Level 2	Time dependent corrections (already included into Level 0)
Level 3	Not in use
Level 4	Multiple $p\bar{p}$ interactions correction, $Mp\bar{p}I$
Level 5	Absolute energy scale ($P_T^{calo} \rightarrow P_T^{particle}$), f_{jes}
Level 6	Underlying Event correction, UE
Level 7	Out-of-Cone correction, OOC

Table 4.1: Naming convention for the different jet corrections.

For all these reasons, a set of correction algorithms have been developed [7], whose input variables are E_T and η of the jet, in order to scale measured jet energy back to the energy of the particle originating the jet. Tab. 4.1 shows the current naming convention for the different type of corrections.

Level 0 correction

These corrections are applied in the CEM to set the overall energy scale with electrons resulting from the Z^0 boson decay. The same calibration is performed in CHA and WHA via J/Ψ electrons about every 40 pb^{-1} of collected data. ^{60}Co radioactive sources and laser beams allow to transport the relative calibration to the entire calorimeter volume.

η-dependent correction

Even after the calorimetric absolute scale calibrations, the response of the calorimeter is not uniform in pseudorapidity. The differences are due to uninstrumented regions, different amount of material in the tracking volume and in the calorimeters, different responses by detectors built with different technologies. The response dependencies on η arise from the separation of calorimeter components at $\eta = 0$, where the two halves of the central calorimeter join, and at $\eta \sim 1.1$, where the plug and central calorimeter are merged.

The η-dependent corrections are obtained by requiring P_T balance between the two leading jets in dijet events (*dijet balancing method*). The corrections are determined based on the fact that the two leading jets in dijet events should be balanced in P_T in absence of hard QCD radiation. To determine the corrections, events with exactly two jets are selected, one of which is called *trigger* and is in the region $0.2 < |\eta_{jet}| < 0.6$ where the response of the calorimeter is well understood, while the other one is called probe. If both jets in an event are within $0.2 < |\eta_{jet}| < 0.6$, trigger and probe jets are assigned randomly. The correction consists in modifying the probe jet transverse energy in order to balance the transverse energy of the trigger. The P_T balancing fraction, $\Delta P_T f$, is then

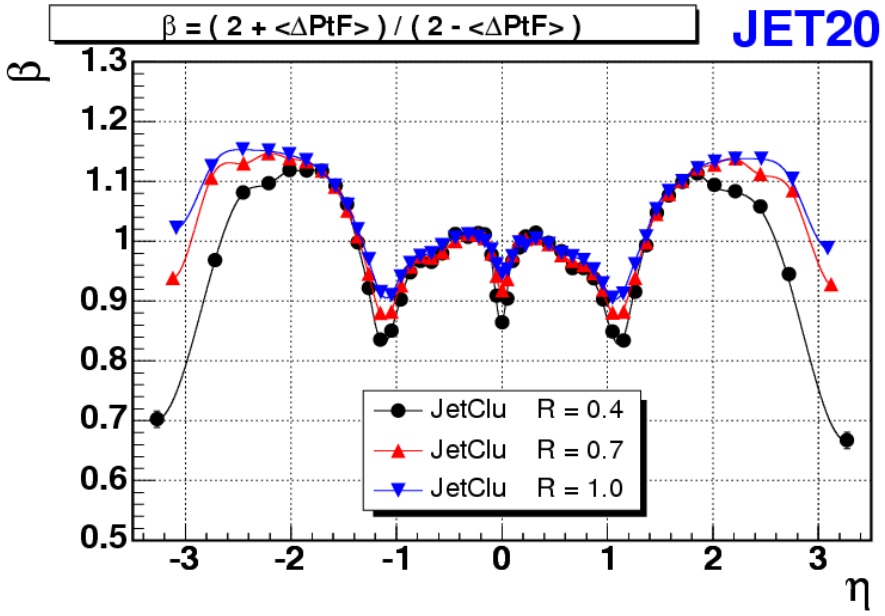


Figure 4.4: Relative energy scale correction factor as a function of η for three different values of cone radii. *Jet*20 is the name of the sample on which correction was calculated: it is a sample of events collected with a trigger requiring at Level 1 one calorimetric tower with energy above 20 GeV.

defined as:

$$\Delta P_T f = \frac{\Delta P_T}{P_T^{ave}} = \frac{P_T^{probe} - P_T^{trigger}}{\left(P_T^{probe} + P_T^{trigger}\right)/2} \tag{4.5}$$

With the above definition, the correction factor required to correct the probe jet can be inferred as:

$$\beta_{dijet} = \frac{2 + < \Delta P_T f >}{2 - < \Delta P_T f >} \tag{4.6}$$

In Fig. 4.4 we show the correction factor as a function of η. The η-depended corrections also include time dependence corrections for the calorimeter response and P_T dependence.

Multiple $p\bar{p}$ interaction

At high instantaneous luminosity more than one $p\bar{p}$ interaction may occur in the same bunch crossing due to the large $p\bar{p}$ cross section at the Tevatron center-of-mass energy. Given the Tevatron characteristics, the average number of interactions is one for $\mathcal{L} = 0.4 \times 10^{32}\ cm^{-2}s^{-1}$, and increases to 3 and 8 for $\mathcal{L} = 1 \times 10^{32}\ cm^{-2}s^{-1}$, and $\mathcal{L} = 3 \times 10^{32}\ cm^{-2}s^{-1}$, respectively.

Energy from these non overlapping minimum bias events may fall into the jet clustering cone of the hard interaction thus causing a mismeasurement of jet energy.

In order to compute the corrections, the number of primary vertices of quality 12 (see Sec. 4.2 for definitions) in the event N_{vtx} is taken into account. Indeed, N_{vtx}

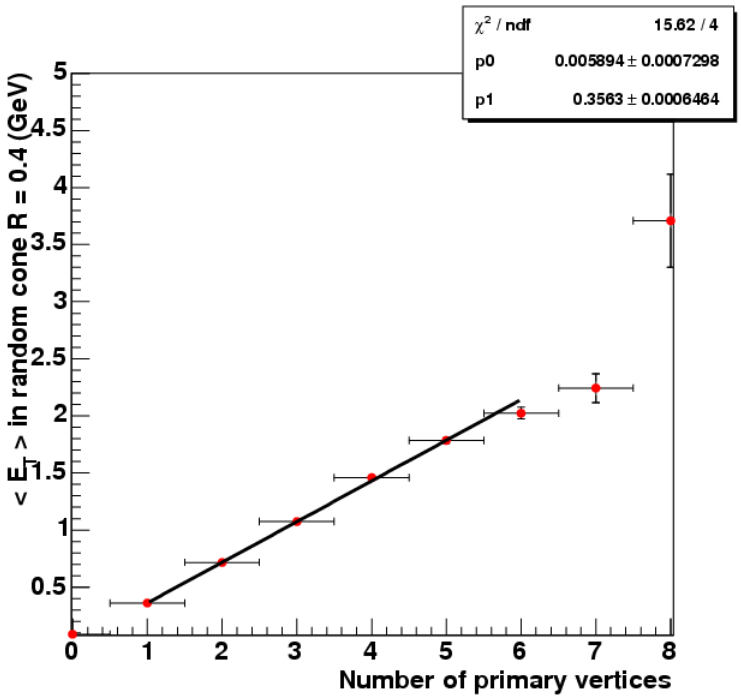


Figure 4.5: Average transverse energy as a function of the number of primary vertices in the event: a correction factor for multiple interaction is extracted from the slope of the fitting line.

is a good indicator of additional interactions occurring in the same bunch crossing. For each event, the transverse energy inside cones of different radii (0.4, 0.7 and 1.0) is measured in a region far away from cracks ($0.1 \leq |\eta| \leq 0.7$) in a minimum bias data sample [8]. Then the distribution of average E_T as a function of the number of quality 12 vertices is fitted with a straight line and the slope of the fitting line is taken as a correction factor (see Fig. 4.5). This procedure allows to extract the average energy each extra vertex in the event is adding, and then to correct jet energies accordingly.

Absolute jet energy scale

A jet contains different types of particles with wide momentum spectra. As calorimeter response to a particle depends on its momentum, position, incident angle and type of particle, the jet momentum at hadron level is in general different from its momentum measured at calorimeter level [9]. Absolute energy scale correction converts the calorimeter cluster transverse momentum P_T to the sum of transverse momenta of the particles in the jet cone: calorimeter energy is converted to particle energy. After this correction the energy scale of a jet becomes independent from the CDF II detector. The procedure to extract a calorimeter-to-hadron correction factor is based on the following steps :

1. Generate a large sample of MC events with full CDF simulation to cover the P_T range $[0, 600]$ GeV;

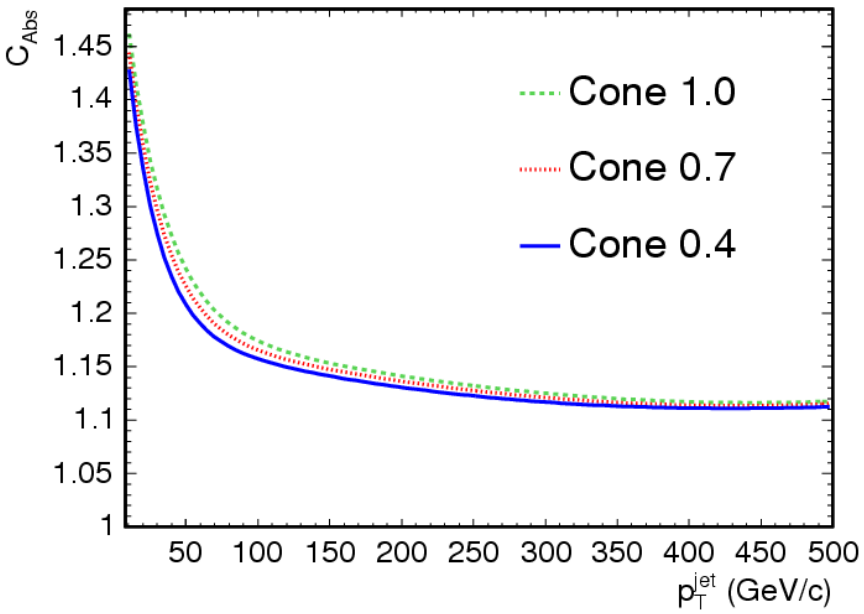


Figure 4.6: Absolute corrections for different cone sizes as a function of calorimeter jet momentum.

2. Create clusters of calorimeter towers and of HEPG particles using the same CDF standard cluster algorithm;

3. Associate calorimeter-level jets with hadron-level jets;

4. Parameterize the mapping between calorimeter and hadron-level jets as a function of hadron-level jets;

5. The absolute correction is defined maximizing the probability $d\mathcal{P}(P_T^{particle}, P_T^{calo})$ of measuring a jet with P_T^{cal} given a jet with a fixed value of P_T^{had}.

Absolute corrections as a function of calorimeter-level jet momentum are shown in Fig. 4.6 for different cone sizes.

Underlying event

In a hadron-hadron collision, in addition to the hard interaction that produces the jets in the final state, there is also an underlying event, originating mostly from soft spectator interactions. In some of the events, the spectator interaction may be hard enough to produce soft jets. Energy from the underlying event can fall in the jet cones of the hard scattering process thus biasing jet energy measurements. A correction factor for such effect has been calculated using a sample of minimum bias events as for multiple interaction correction, but selecting only those events with one vertex [10]. For each event, transverse energy E_T inside cones of different radii (0.4, 0.7 and 1.0) is measured in a region far away from cracks ($0.1 \leq |\eta| \leq 0.7$). The correction factor is extracted from the mean values of E_T distribution.

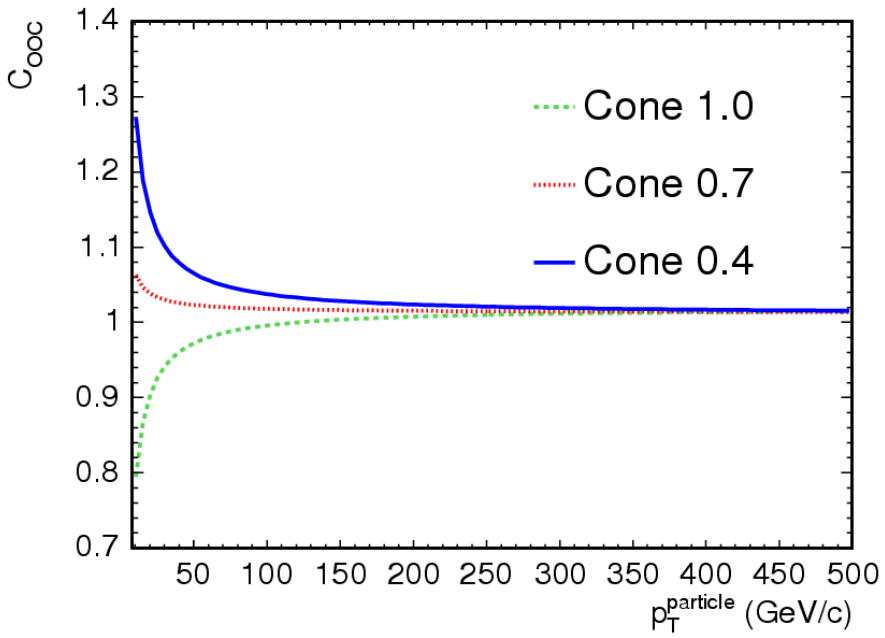


Figure 4.7: Out-of-cone correction factor as a function of jet momentum for different cone sizes.

Out-of-cone energy

The jet clustering may not include all the energy from the initiating partons. Some of the partons generated during fragmentation may fall outside the cone chosen for the clustering algorithm. This energy must be added to the jet to get the parton level energy. A correction factor is obtained using MC events [11]: hadron-level jets are matched to partons if their distance in the $\eta - \phi$ plane is less than 0.1. Then the difference in energy between hadron and parton jet is parameterized using the same method as for absolute corrections (see Fig. 4.7).

Depending on the physics analysis, all of the reviewed corrections or just a subset of them can be applied.

Corrections are applied to the raw measured jet momentum according to the following equation [7]:

$$\begin{aligned} P_T(R, P_T, \eta) &= [P_T^{raw}(R) \times f_\eta(R, P_T^{raw}, \eta) - Mp\bar{p}I(R)] \times f_{jes}(R, P_T^{raw}) - \\ &\quad -UE(R) + OOC(R, P_T^{raw}) \end{aligned} \tag{4.7}$$

where R is the clustering cone radius, P_T^{raw} is the raw (i.e. measured) energy, and η is the pseudorapidity of the jet with respect to the center of the detector. On the other hand, f_η refers to the η-dependent correction, $Mp\bar{p}I$ stands for multiple interaction correction; f_{jes} is the jet scale energy correction, and finally, UE and OOC indicate the underlying event and out-of-cone correction factors, respectively.

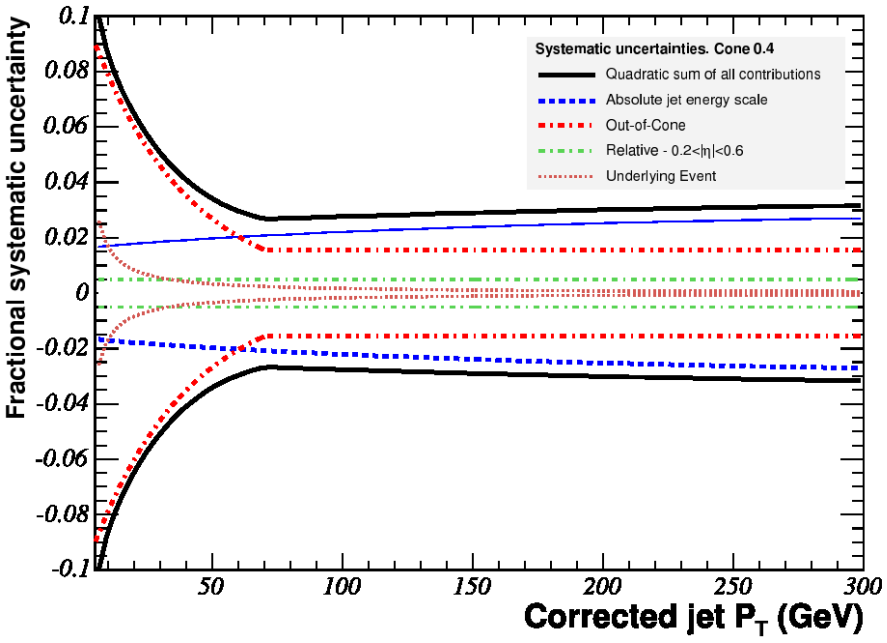


Figure 4.8: Total systematic uncertainties to corrected jet P_T.

Jet energy measurement systematic uncertainties

The application of jet corrections is subjected to systematic uncertainties whose origin can be either related to the method used for their calculation or to discrepancies in the jet modelling between data and Monte Carlo. The systematic uncertainties associated to the jet energy response are found to be largely independent of the correction applied and mostly arising from the jet description provided by the Monte Carlo simulation.

The total systematic uncertainty to the jet corrected P_T is shown in Fig. 4.8, and it results from the sum in quadrature of several contributions coming from the systematics associated with each level of correction described previously. For high P_T the largest contribution arises from the absolute energy scale which is limited by the uncertainty of the calorimeter response to charged hadrons. On the other hand, at low P_T the main contribution to the total uncertainty arises from the modelling comparison of the energy flow around the jet cone between data and Monte Carlo samples.

4.4 Missing energy measurement

Neutrinos cannot be directly detected however their production can be inferred by the presence of imbalance in the calorimeter energy. The longitudinal component of the colliding particle momenta is not accessible, but the transverse component can be measured and it is subjected to conservation. From the transverse energy measured in the calorimeter, the transverse component of the neutrino momenta can be calculated.

The missing transverse energy $\vec{\not{E}}_T$ is a two component vector $(\not{E}_{Tx}, \not{E}_{Ty})$ whose

raw value is defined by the negative vector sum of the transverse energy of all calorimeter towers:

$$\vec{\not\!\!E}_T^{\,raw} = -\sum_{towers} (E_T^i)\vec{n}_i \tag{4.8}$$

where E_T^i is the transverse energy of the i-th calorimeter tower, and $\vec{n}_i$ is a transverse unit vector pointing from the center of the detector to the center of the tower. The sum extends up to $|\eta| < 3.6$.

The value $\vec{\not\!\!E}_T^{\,raw}$ needs to be corrected for the actual primary vertex position, for escaping muons and for energy mismeasurements. Muons do not deposit substantial energy in the calorimeter, but may carry out a significant amount of energy. The sum of transverse momenta of escaping muons $\sum \vec{P}_{T,\mu}$ measured in the COT has to be accounted for in the calculation of $\vec{\not\!\!E}_T$. On the other hand, the energy corrections to jets must be taken into account too.

Uncertainties on $\not\!\!E_T^{corr} = \sqrt{\not\!\!E_{Tx}^{\,2} + \not\!\!E_{Ty}^{\,2}}$ are dominated by uncertainties related to jet energy response (Sec. 4.3.1).

The resolution of the $\not\!\!E_T$ generally depends on the response of the calorimeter to the total transverse energy deposited in the event. It is parameterized in terms of the total scalar transverse energy $\sum E_T$, which is defined as:

$$\sum E_T = \sum_{towers} E_T^i. \tag{4.9}$$

The $\not\!\!E_T$ resolution in the data is measured using minimum bias events [12], dominated by inelastic $p\bar{p}$ collisions. Since minimum bias events are spherically distributed, no large energy imbalance is expected.

The $\not\!\!E_T$ resolution is defined by $\Delta = \sqrt{< \not\!\!E_T{}^2 >}$. For minimum bias events both the x and y component of the missing energy are distributed according to a Gaussian distribution with zero mean and $\sigma_x = \sigma_y = \sigma$ so that:

$$\begin{aligned} \frac{dN}{d\not\!\!E_{Tx}} &\sim e^{-\frac{\not\!\!E_{Tx}^2}{2\sigma^2}} \\ \frac{dN}{d\not\!\!E_{Ty}} &\sim e^{-\frac{\not\!\!E_{Ty}^2}{2\sigma^2}} \end{aligned} \tag{4.10}$$

Consequently, $\Delta = \sqrt{2}\sigma = \sqrt{< \not\!\!E_T{}^2 >}$. The $\not\!\!E_T$ resolution, Δ, is observed to scale as the square root of the total transverse energy, $\sum E_T$. From minimum bias studies it is found to be $\Delta \sim 0.64 \sum E_T$ [12], as shown in Fig. 4.9.

4.5 b-jet identification

The high position resolution provided by the silicon vertex detector can be exploited to identify secondary vertices originated inside a jet by decays of long lifetime particles produced in heavy quark hadronization. For this purpose, the SECondary VerTeX (SECVTX) tagging algorithm [13, 14] has been developed.

The B hadrons produced by bottom quark hadronization have a lifetime of the order of a picosecond and at the typical energy of the bottom quark originating

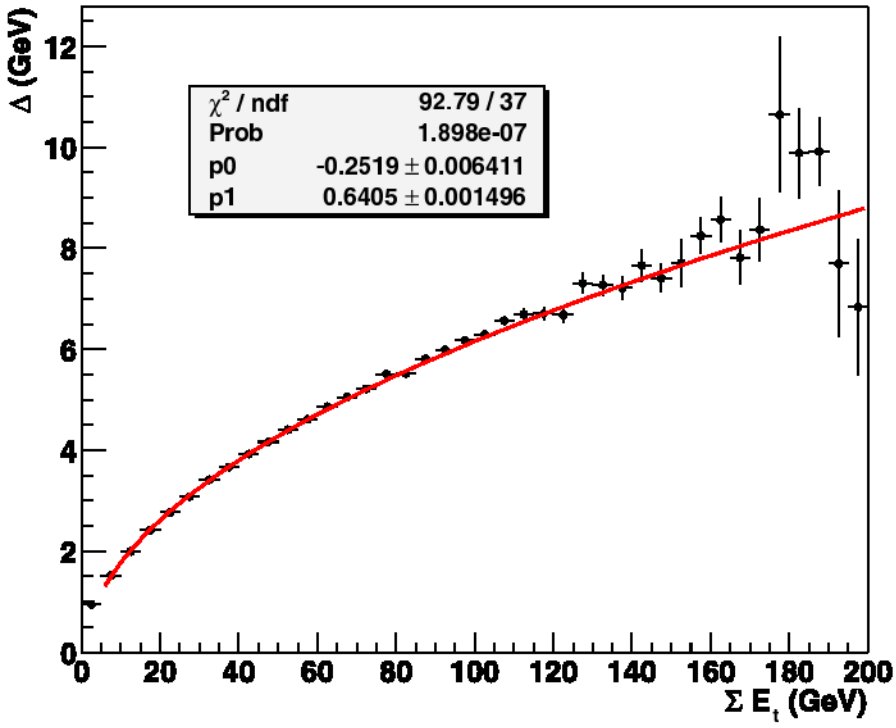


Figure 4.9: $\not\!\!E_T$ resolution as a function of $\sum E_T$ measured in minimum bias events [12].

by top quark they travel some millimeters before decaying. This provides a way to discriminate high P_T b-jets from jets originated by light quarks or gluons: the SECVTX algorithm relies on the displacement of secondary vertices relative to the primary event vertex to identify b hadron decays. In the following the SECVTX algorithm will be described.

The SecVtx algorithm

The secondary vertex tagging algorithm operates on a per-jet basis, where only tracks within the jet cone are considered for each jet in the event. A set of cuts involving the transverse momentum, the number of silicon hits attached to the track, the quality of those hits and the χ^2/n.d.f. of the final track fit are applied to reject poorly reconstructed tracks.

Only jets with at least two of these tracks can produce a displaced vertex; a jet is defined as "taggable" if it has at least two good tracks. Displaced tracks in the jet are selected on the basis of the significance of their impact parameter with respect to the primary vertex and are used as input to the SECVTX algorithm (Fig. 4.10). Tracks identified as K_S or Λ daughters, or consistent with primary vertex or too far from it are removed.

SECVTX uses a two-pass approach to find secondary vertices: in the first pass, using tracks with $P_T > 0.5\ GeV$ and $d_0/\sigma_{d_0} > 2.0$, it attempts to reconstruct a secondary vertex which includes at least three tracks. If the first pass is unsuccessful, it performs a second pass which makes tighter track requirements ($P_T > 1\ GeV$

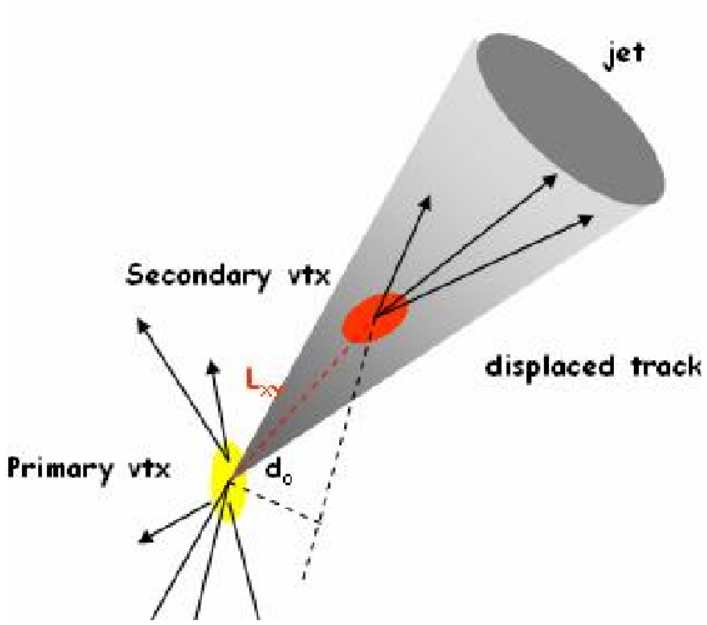


Figure 4.10: Reconstruction of the primary and secondary vertices in the r-ϕ plane. The impact parameter d and the distance L_{xy} (or L_{2d}) between the vertices in the transverse plane are shown.

and $d_0/\sigma_{d_0} > 3.5$) and attempts to reconstruct a two-track vertex.

Once a secondary vertex is found in a jet, the two-dimensional decay length of the secondary vertex L_{2d} is calculated as the projection onto the jet axis in the $r-\phi$ view only of the vector pointing from the primary vertex to the secondary vertex. To reduce the background from false secondary vertices (mistags), a good secondary vertex is required to have $|L_{2d}/\sigma_{L_{2d}}| > 7.5$. A tagged jet is defined to be a jet containing a good secondary vertex. Secondary vertices corresponding to the decay of b and c hadrons are expected to have large positive L_{2d} while the secondary vertices from random mis-measured tracks are expected to be less displaced from the primary vertex. The tags are classified depending on where the secondary vertex is located with respect to the jet cone axis.

Secondary vertices on the same side of the interaction point as the jet cone axis are positive tags, otherwise they are classified as negative tags. Negative tags can arise from tracks mismeasurements as illustrated in Fig. 4.11.

4.6 Electron identification

Electrons resulting from electroweak W and Z production or from top decays are generally highly energetic and can be identified as high-P_T tracks in the drift chamber accompanying large energy deposition in the electromagnetic calorimeters. Electron identification relies on the combination of tracking and calorimetric information. Electrons and photons leave a characteristic signature in the calorime-

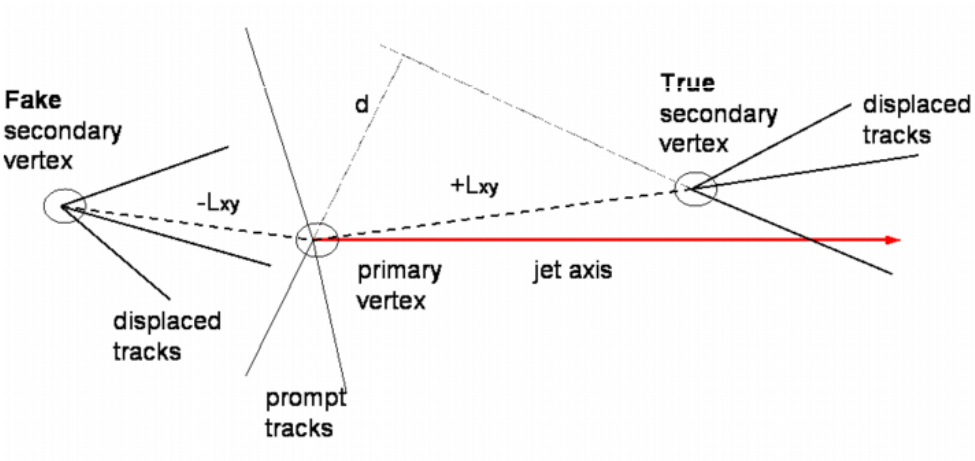


Figure 4.11: Real and fake secondary vertices as seen in the transverse plane.

ter, their electromagnetic shower. Electrons can be distinguished from photons in part by the slight difference of the shape of the electromagnetic shower, but mostly by requiring a track to point to the calorimetric cluster produced by the shower; photons, being neutral, do not leave any trace in the tracking systems [15].

4.7 Muon reconstruction

Unlike electrons, muons do not initiate an electromagnetic shower in the calorimeters due to their larger mass (105 MeV compared to 0.511 MeV). Moreover, unlike hadrons, muons do not interact strongly and hence do not shower in the hadronic calorimeter either. As a result, muons with a transverse energy of few GeV or more deposit only a small fraction of their energy in the calorimeters due to ionization, and escape the detector. Muons are thus identified by matching hits in the muon chambers with a well reconstructed track in the drift chamber and requiring little energy to be deposited in the calorimeter along the particle trajectory. In each muon system (CMU, CMP, CMX) the scintillator layers provide the reconstruction of muon track segments (stubs). A muon candidate is reconstructed if such a stub is found in one of the muon systems and if an extrapolated COT track matches with the stub [16].

4.8 Tau reconstruction

Tau lepton can decay leptonically into electron or muon (and the corresponding neutrinos) or semileptonically into charged and neutral pions; the first case is not distinguishable from a leptonic decay from W, while the second has a precise signature: tau decays preferably into 1 or 3 charged pions (One/Three prong event) and in most cases also neutral pions are present. So a well isolated jet with low track multiplicity and neutral pions is a good tau candidate. Tau reconstruction

procedure exploits information from the calorimeter and tracking systems: the algorithm searches for an isolated narrow cluster above a certain energy treshold and then matches it to COT tracks [17].

4.9 Photon identification

A photon traversing CDF detector interacts only with electromagnetic calorimeter and shower maximum detector. Thus photon identification starts by looking for clusters around a seed tower with energy $\geq 3\ GeV$. Total energy of the hadronic towers located behind the photon cluster has to be negligible with respect to the photon cluster energy. Additionally, photon cluster isolation is required: the difference between photon energy and the energy in a 0.4 cone around the seed tower has to be less tan 15% of the photon energy. Moreover the sum of transverse momenta of all tracks pointing to the 0.4 cone is required to be less than 2 GeV. Electromagnetic shower shape shall be transverse and no matching tracks have to be present. The line connecting the primary event vertex to the CES shower position determines the photon's direction [18].

Bibliography

[1] H. Wenzel, *Tracking in the SVX*, CDF Internal Note 1790.

[2] W.M. Yao and K. Bloom, *Outside-In silicon Tracking at CDF*, CDF Internal Note 5991.

[3] C. Hays *et al.*, *Inside-Out tracking*, CDF Internal Note 6707.

[4] H. Studie *et al.*, *Vxprim in Run II*, CDF Internal Note 6047.

[5] J.F. Arguin *et al.*, *The z-Vertex Algorithm in Run II*, CDF Internal Note 6238.

[6] F. Abe *et al.* [CDF collaboration], *Topology of three-jet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.8\ TeV$*, Phys. Rev. D **45**, 1448 (1992).

[7] A. Bhatti *et al.* [CDF collaboration - Jet Energy Group], *Determination of the jet energy scale at the Collider Detector at Fermilab*, Nucl. Instrum. Meth. A **566** 375 (2006), [arXiv:hep-ex/0510047].

[8] B.Cooper *et al.*, *Multiple interaction corrections*, CDF Internal Note 7365.

[9] A. Bhatti and F. Canelli, *Absolute corrections and their systematic uncertainties*, CDF Internal Note 5456.

[10] J.F. Arguin and B. Heinemann, *Underlying event corrections for Run II*, CDF Internal Note 6293.

[11] A. Bhatti *et al.*, *Out-of-Cone Corrections and their Systematics Uncertainties*, CDF Internal Note 7449.

[12] D. Tsybychev *et al.*, *A study of missing E_T in Run II minimum bias data*, CDF Internal Note 6112.

[13] D. Glenzinsky *et al.*, *A detailed study of the SECVTX algorithm*, CDF Internal Note 2925.

[14] J. Guimares *et al.*, *SecVtx Optimization Studies for 5.3.3 Analyses*, CDF Internal Note 7578.

[15] R.G. Wagner, *Electron identification for Run II: algorithms*, CDF Internal Note 5456.

[16] J.N. Bellinger *et al.*, *A guide to muon reconstruction and software for Run II*, CDF Internal Note 5870.

[17] A. Anastassov *et al.*, *Tau reconstruction efficiency and QCD fake rate for Run II*, CDF Public Note 6308.

[18] *Baseline Analysis Cuts for High Pt Photons V2.3*, http://www-cdf.fnal.gov/internal/physics/photon/docs/cuts.html.

Chapter 5

Neural Networks

The main goal of our analysis will be to extract the $t\bar{t} \to \not{E}_T + jets$ signal events from our complete data sample and to be able to discriminate between top-like and background events. We will rely heavily on the features provided by Artificial Neural Networks, having as inputs the kinematical variables that best discriminate signal among backgrounds. In the following we will give a brief description of the main concepts about neural networks, describing in detail the learning strategy we used in this work.

5.1 Introduction

Artificial Neural Networks, more precisely *Feed Forward Neural Networks*, belong to the multivariate analysis branch of statistics; they may be defined as a computing system of Von Neumann type aiming at approximating efficiently a given mapping from a subset D of $\mathcal{R}^n$ into $\mathcal{R}^m$ with $m \leq n$ on the basis of a set of known examples, often called *training set*. In particular, in this work we will restrict ourselves to the case $m = 1$ so that the network will be mapping a vector of variables into a single scalar variable; this will allow the use of the FFNN as a simple classifier between signal and background events by searching for a mapping that will assign 0 to all background and 1 to all signal events. The mapping is defined as a function of a number of parameters, called *weights*, and organized in a particular hierarchical structure, called *architecture*, whose smallest unit is the *perceptron*.

5.2 Perceptrons and Neural Networks

A perceptron is a mathematical abstraction of a biological neuron, see Fig. 5.1. Given a set $\vec{x} = (x_1, \ldots, x_N)$ of N input variables, the perceptron output value y is given by the following expression:

$$y = \theta\left(\frac{1}{N}\sum_{i=1}^{N}\omega_i x_i - \phi\right) \tag{5.1}$$

where $\{\omega_i\}$ are the weights of the connections entering the perceptron, $\theta(\zeta)$ is a transfer function (among the many available choices, the most common are Heav-

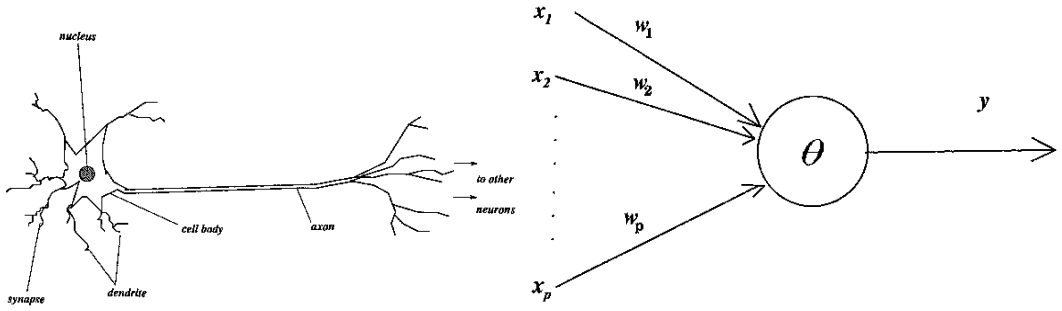


Figure 5.1: Perceptron as a mathematical abstraction of a neuron.

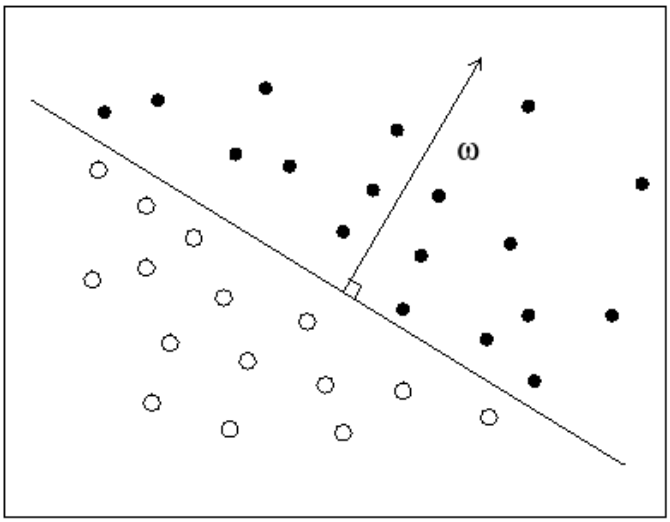


Figure 5.2: Example of two classes separated by perceptron weights $\{\omega_i\}$.

iside's step function, heath bath function $\theta(\zeta) = \tanh(\beta\zeta)$ or any smooth variant of the step function) and ϕ is a bias. The operating mode of a perceptron has an easy geometrical interpretation: basically it provides two-class classification, as illustrated in Fig. 5.2. In fact, if we have a mapping into two linearly separated subsets A and B with $A, B \subset \mathcal{I}$ (i.e. it is possible to find an hyperplane that separates the two subsets), then a single perceptron is sufficient to reproduce the mapping, since there exists a vector $\vec{\omega}$ such that the two conditions:

- $A = \{\vec{x} \in \mathcal{I} : \vec{\omega} \cdot \vec{x} \geq 0\}$
- $B = \{\vec{x} \in \mathcal{I} : \vec{\omega} \cdot \vec{x} < 0\}$

are sufficient to define the two subsets; in this case the components of the vector $\{\omega_i\}$ will be the perceptron's weights.

Unfortunately only linearly separable sets can be classified using a single perceptron: for example two dimensional *AND* and *OR* logic operators can be implemented using a single perceptron, while an exclusive OR, *XOR*, cannot; to overcome this limitation one can combine multiple perceptrons in such a way that the output

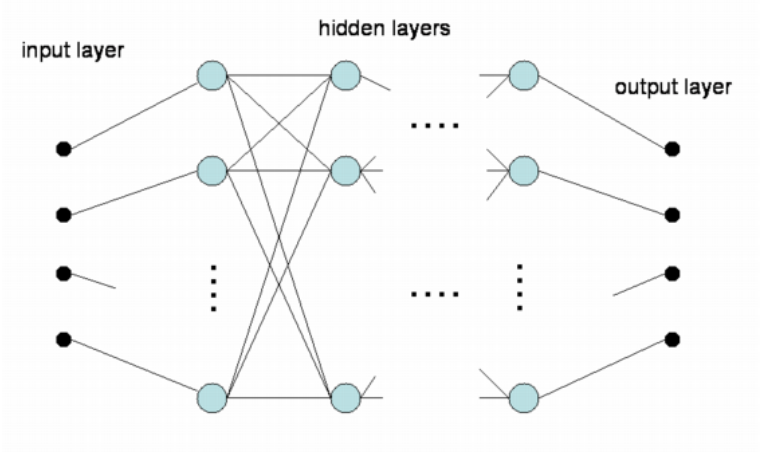


Figure 5.3: Example of a multi-layer percepetron.

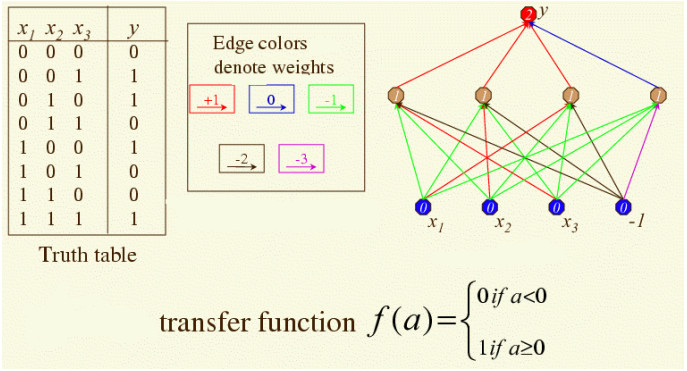


Figure 5.4: A simple neural network implementing the XOR logic operator.

of one becomes the input of another, building an architecture with multiple layers, as depicted in Fig. 5.3. Then it is understandable that in this kind of network perceptrons could cooperate to build some sort of set of contiguous "pieces" of flat hypersurfaces capable of approximating the generally curved surface of separation between the two sets. As an example, Fig. 5.4 illustrates a simple multi layer network implementing the *XOR* logic operator.

Mappings that separate their definition sets into multiple subsets are typical in classification problems through pattern recognition and in high-energy physics analysis. What makes Neural Networks particularly suitable in these tasks and better performing than the usual "sequential-cuts" attack to the problem is that a cut on their output for classification pourposes may be completely impossible to reach using simple sequential cuts on any of the projections of the definition sets on the available axis: this is visualized in Fig. 5.5 for a simple two-dimensional case.

Neural Networks architectures are usually identified by the number of *layers* they are made of, each composed by a definite number of *neurons*, and by the activation function used in those neurons. Typically, in software implementations

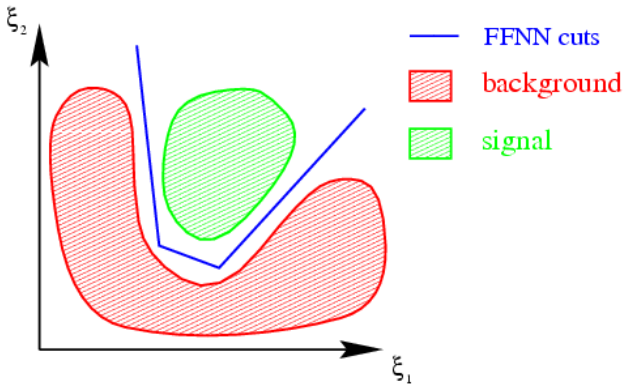


Figure 5.5: Example of a two-dimensional classification: while the two definition classes are not separable using any set of sequential cuts on the two axis, the blue line could be an hypersurface of constant output for a suitable Neural Network, thus allowing a cut separating the two classes.

the *sigmoid* function is used:

$$S(x) = \frac{1 - e^{-1}}{1 + e^{-1}} \tag{5.2}$$

The set of input nodes is called *input layer*, the set of output nodes *output layer*, while all the remaining layers are called *hidden layers*. The specific class of networks we will use in the following are called "feed forward" networks because the informations proceeds from input to output along successive layers.

It is possible to prove that any continous functional mapping from a finite-dimensional space to a finite-dimensional space can be approximated aribitrarily well using a two-layer network, if a sufficient number of hidden perceptrons is provided; a complete discussion of this important feature can be found in [1, 2, 3] and references. What is particularly interesting is that in the context of classification problems, networks with sigmoidal nonlinearities of two layers can approximate any decision boundary with arbitrary accuracy.

5.3 Training

Once we choose a topology, in order to use the desired network as a classifier first we need to determine the weights to be associated to each perceptron. This task is performed using a set of *a priori* known samples belonging to the classes we want to separate, and the whole process goes under the name of "supervised learning". The procedure of creating the approximate mapping (known as *training*) consists in finding the set of weights and biases that minimizes the difference between the desired outputs $\{y\}$ and the outputs $\{o\}$ obtained by the neural network on the training samples. Usually the function to minimize is the following quadratic error

function:

$$f(\{o\},\{y\}) = \frac{1}{\#}\sum_{i=1}^{\#}(o_i - y_i)^2 \tag{5.3}$$

where # is the number of elements in the training sample.

General theorems ensure that absolute minima of this kind of function do exist (see for example [4]), but finding them is obviously complex and computing intensive due to the high dimensionality of the configuration space; doing an exhaustive search exploring all possible values of the weights is not an option in any real case scenario. That's why many different algorithms have been created in order to implement optimal strategies to find minima; we will review in the following some aspects of a particular *search method* based on the so called "Reactive Taboo Search" strategy and a review of a method in the *steepest gradient* class known as "BFGS": both of them were important during the preparation of this work.

Before we proceed any further, it is also useful to stress another important issue, related to the choice of training samples. The fact that the training samples have a finite number of elements implies that the mapping function implemented by the network will by definition have some *noise*, so that there will be an overlapping in its output for elements belonging to different classes in the input variable space. It is then crucial to find some training patterns that are good enough representative of the classes we want to separate.

Once a set of weights is chosen, the next step is to proceed with the *testing* or *generalization* phase and to classify some new known elements in order to test the performances of the network and check the value of the error function on the test sample. The choice of a set of weights and the successive testing phase constitute an *epoch* of the training process. A training can continue through several epochs before reaching a minimum of the error function.

When using networks trained with a single output neuron used to separate two classes (this will be our case throughout this work, where we will try to discriminate signal and background in our decay channel) it can be shown that the output of the network may be interpreted as the probability that an element belongs to a particular class. In this two-dimensional problem the performances of the network are usually evaluated using an *efficiency* vs. *purity* curve where efficiency $\epsilon(cut)$ and purity $\eta(cut)$ are defined as follows:

$$\epsilon(cut) = \frac{N_s^{pass}(cut)}{N_s},\ \eta(cut) = \frac{N_s^{pass}(cut)}{N_s^{pass}(cut) + N_b^{pass}(cut)} \tag{5.4}$$

and $N_s^{pass}(cut)$ ($N_b^{pass}(cut)$) is the number of signal (background) events passing the cut on the neural network output (i.e. with $NNout >= cut$), and N_s is the total number of signal events in the test sample. Basically, purity describes how well a neural network can discriminate between signal and background, while efficiency is a measure of the neural network capability in recognizing signal events. An ideal neural network should have infinite precision in discriminating signal from background, so $\epsilon \approx 1$ and $\eta \approx 1$ and the efficiency vs. purity plot would be in this case a step function: the more the plot obtained after the training approaches the ideal one, the better the performances of the neural network.

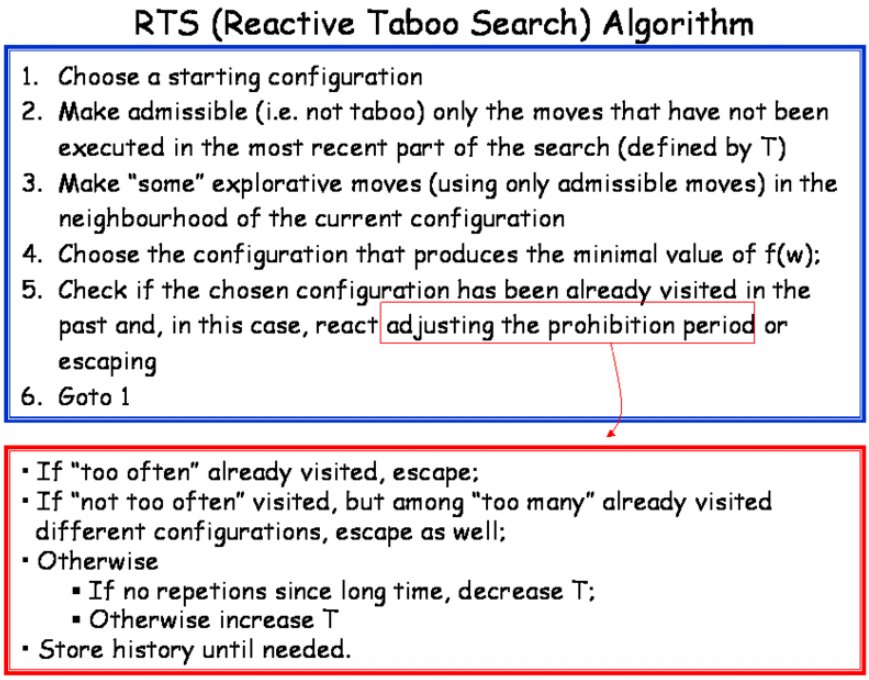


Figure 5.6: Basic operations of the RTS training algorithm.

5.3.1 Reactive Taboo Search training algorithm

In the early stages of this work we have been using this non-derivative based search strategy to train our network. The RTS training algorithm was developed by a joint INFN IRST effort in Trento to exploit the features of a custom hardware neural network chip called *TOTEM*. The algorithm (see [5, 6] for a detailed description) makes a combinatorial optimization of the squared error function by means of a heuristic operational method that will be briefly described. First of all the problem is translated from the weight space into a $\{0-1\}$ string using *Grey* encoding, to fully exploit the features of the chip hardware. The method is based on the construction of search paths in the string space, with the aim of locating an optimal minumum on the error surface by means of a sequence of elementary moves, each consisting in a single bit flip in the string of weights. Each visited configuration is recorded for future reference: an important feature of RTS is that it is intensively *history based*. When a move is done, its inverse is forbidden for a number of successive steps T called *prohibition period*, that can be dinamically adjusted if the configuration was already visited in the past. The path is built by choosing among the admissible elementary moves in the string space the one producing the minimal value of the error function and by iterating the process until the required precision is reached. Every time the same configuration is visited again T is increased, while it decreases if the moves are exploring new unknown configurations; if T grows too much, meaning that the same configuration is visited too often (or if its neighbours in terms of elementary moves are) then the algorithm escapes to a different random configuration. A summary of the steps involved in the training algorithm is shown in Fig. 5.6.

This allows some diversification in the training process, in a way that prevents to get trapped in local minima, reacting dinamically to the local shape of the error surface and avoiding attractive cycles by random escapes.

5.3.2 BFGS training algorithm

Another approach to the minimization problem is constituted by the so called steepest gradient methods: in this kind of approach the minimum on the error surface is searched starting from a random point in the weights space and then by moving along the steepest direction around that point; this is repeated until no further improvements are possible. Methods like these go usually under the name of *backpropagation* and require to compute the local slope of the error surface, usually a difficult task since it involves the calculation of many derivatives.

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is based on the quasi-Newtonian method developed around 1970 indipendently by the authors of [7], [8], [9] and [10] to solve an unconstrained nonlinear optimization problem. Following the Newtonian optimization method, one assumes that the error function can be approximated as quadratic in the region around its minumum and uses its first and second derivatives to find the stationary point; the iterative procedure to find the minumum starts from a random point x_0 in the weights space and for each step k, if f is the function we want to minimize, one would have to calculate in the point x_k the steepest direction $\mathbf{p}_k$ like:

$$H_k \mathbf{p}_k = -\nabla f(\mathbf{x}_k) \tag{5.5}$$

where H_k denotes the complete Hessian Matrix of the function f in that point:

$$H_k(\mathbf{x}_k)_{ij} = \frac{\partial^2 f(\mathbf{x}_k)}{\partial x_i \partial x_j} \tag{5.6}$$

Then a *line search* along $\mathbf{p}_k$ is used to find the next point $\mathbf{x}_{k+1}$, by *loosely* minimizing (i.e. requiring a sufficient decrease) the following function of the parameter α_k, $\phi(\alpha_k)$:

$$\phi(\alpha_k) = f(\mathbf{x}_k + \alpha_k \mathbf{p}_k),\ \alpha_k \in \mathbb{R} \tag{5.7}$$

In quasi-Newtonian methods, instead of computing the full Hessian matrix H_k of the function in Eq.5.5 at each iteration step, an approximated matrix B_k is defined and updated by analyzing successive gradient vectors. In particular, in the BFGS method the following approximation is used:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k,\ \mathbf{y}_k = \frac{\nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)}{\alpha_k} \tag{5.8}$$

$$B_{k+1} = B_k + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{p}_k} - \frac{B_k \mathbf{p}_k (B_k \mathbf{p}_k)^T}{\mathbf{p}_k^T B_k \mathbf{p}_k} \tag{5.9}$$

A complete review of the algorithm goes beyond the pourpose of this thesis, so we suggest the curious reader to check for example [11].

Even if in our experience RTS based training strategies have proven to give slightly better results than derivative based ones, during this analysis we decided to use a neural network training method based on BFGS optimization procedure for its fast and easy to use software implementation in the ROOT Analysis Framework [12], the program used for data access and analysis in the preparation of the work.

Bibliography

[1] R. Rojas, **Neural Networks, a systematic introduction**, *Springer-Verlag (1996).*

[2] C.M. Bishop, **Neural Networks for Pattern Recognition**, *Oxford University Press (1995).*

[3] B.D. Ripley, **Pattern Recognition and Neural Networks**, *Cambridge University Press (1996).*

[4] R. Hecht-Nielsen, **Neurocomputing**, *Addison-Wesley (1990).*

[5] R. Battiti, G. Tecchiolli *The Reactive Tabu Search*, ORSA J. Comp. **6**, 126 (1994).

[6] R. Battiti, G. Tecchiolli *Training neural nets with the reactive tabu search*, IEEE Transactions on Neural Networks **6**(5), 1185 (1995).

[7] C.G. Broyden, *The convergence of a class of double-rank minimization algorithms*, J. Inst. Math. Applcs., **6**, 76 (1970).

[8] R. Fletcher, *A new approach to variable metric algorithms* Computer J. **13**, 317 (1970).

[9] D. Goldfarb, *A family of variable metric methods derived by variational means*, Math. Computation **24**, 23 (1970).

[10] D.F. Shanno, *Conditioning of quasi-Newton methods for function minimization*, Math. Computation **24**, 647 (1970).

[11] J.E. Dennis, S. Schnabel, **Numerical Methods for Unconstrained Optimization and Nonlinear Equations**, *Prentice-Hall (1983).*

[12] R. Brun, F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Nucl. Inst. and Meth. in Phys. Res. A **389**, 81 (1997).
See also http://root.cern.ch/.

Chapter 6

The $t\bar{t} \rightarrow \not\!\!E_T + jets$ channel selection

In $p\bar{p}$ collisions at $\sqrt{s} = 1.96\ TeV$ top quark pairs are produced through $q\bar{q}$ annihilation ($\sim 85\%$) and gluon fusion ($\sim 15\%$). Since $|V_{tb}| \sim 1$ and $M_t > M_W + M_b$, the $t \rightarrow W^+ b$ decay is dominant (and has branching ratio $\sim 100\%$ in Standard Model); so we can classify the different top quark pairs search channels with respect to the W boson decay modes.

When both the produced W bosons decay into $e\bar{\nu}_e$ (*or c.c.*) or $\mu\bar{\nu}_\mu$ (*or c.c.*) we have the so called "di-lepton" channel; if both W bosons decay into quark pairs, the final state is instead called "all-hadronic". If one W decays hadronically and the other one leptonically, we have the "lepton+jets" channel. Finally, a so-called "tau dilepton" category was introduced to take into account $e\tau$ and $\mu\tau$ topologies studied in [1].

In this chapter we will describe an inclusive search of the $t\bar{t}$ production process in the $\not\!\!E_T + jets$ final state, using a Neural Network to isolate the decay channel. We will show how this choice grants a high acceptance to general leptonic W decays, with a sizeable presence of $\tau + jets$ top pair decays, that are very difficult to isolate by means of standard τ identification procedure.

Moreover $\not\!\!E_T + jets$ $t\bar{t}$ decays, that were already studied in previous CDF analyses in a lower statistics data sample (see [2, 3]), provide complementary results with respect to standard lepton+jets, di-lepton, and all-hadronic top pair searches: in fact the signal sample we will estract is by means of our choice of cuts orthogonal to the ones used by any other cross section analysis produced so far by the collaboration. This allows us to produce a measurement that will have a strong impact on the combination of the results produced by the CDF experiment.

In following we will review the analysis setup and the tools we used in our work.

6.1 Monte Carlo samples

The two software packages PYTHIA version v6.216 [4] and HERWIG v6.510 [5] are used for the simulation of $t\bar{t}$ events; they can calculate the hard process with leading order QCD matrix elements, and then use different parton showering algorithms to simulate gluon radiation and fragmentation starting from the chosen parton distribution functions.

After that, CDF II detector simulation reproduces the response of each subsystem to particles produced in the collision, for instance:

- Tracking of particles through the detector material is performed using the GEANT package [6];
- Charge deposition in the silicon detectors is calculated using the model in [7];
- The COT drift model uses GARFIELD [8];
- The calorimeter simulation uses GFLASH [9];
- The trigger simulation can be performed using TRIGSIM++ [10];

The main Monte Carlo data sample used in this work, *ttop75*, is a set of almost 4 millions inclusive $t\bar{t}$ events generated using PYTHIA with $M_{top} = 175\ GeV/c^2$, and with corresponding integrated luminosity of 594 fb^{-1} assuming $\sigma_{t\bar{t}} = 6.7\ pb$.

6.2 Data

Several among the available CDF datasets can contain a detectable amount of $\not\!\!E_T + jets\ t\bar{t}$ events and, in principle, many of the available trigger paths could be used to select a data sample in which to perform the analysis.

Our choice was to use the TOP_MULTI_JET trigger, which is specifically designed for the all hadronic $t\bar{t}$ decays, whose final state nominally consists of six hadronic jets. Trigger requirements, among the three-level trigger architecture of the CDF data acquisition system, are the following:

- at Level 1: at least one calorimetric tower with $E_T \geq 10\ GeV$;
- at Level 2: at least four calorimetric clusters with $E_T \geq 15\ GeV$ each plus a total $\sum E_T \geq 125\ GeV$;
- at Level 3: at least four jets with $E_T \geq 10\ GeV$ and $|\eta| \leq 2$.

Additionally, starting from Run 194328 the Level 2 requirements have been changed to cope with higher accelerator luminosity in:

- at Level 2: at least four calorimetric clusters with $E_T \geq 15\ GeV$ each plus a total $\sum E_T \geq 175\ GeV$;

This choice of trigger is mainly due to the analysis strategy we want to deploy: this "multijet" trigger contains the signal signature we are looking for and gives us the possibility of investigating a sample of events that are normally not used by other analyses, providing us a cross section determination uncorrelated with the remaining ones at CDF. Moreover, we will rely on the b-tagging algorithm to indentify heavy flavour jets due to top quark decay: for this reason, triggers using selections based on SVT tracks with large impact parameter are not suitable for our purpose, since they can enrich the heavy flavour fraction of the data sample at the

Dataset	Run Range	CDF code version	Lum. (nb^{-1})
gset0d	138425 - 186598	prod 5.3.1 - topCode 6.1.4	355,460
gset0h	190697 - 203799	prod 6.1.1 - topCode 6.1.4	418,122
gset0i	203819 - 212133	prod 6.1.1 - topCode 6.1.4	
	217990 - 222426	prod 6.1.1 - topCode 6.1.4	
	222529 - 228596	prod 6.1.1 - topCode 6.1.4	
	228664 - 233111	prod 6.1.1 - topCode 6.1.4	886,494
gset0j	233133 - 237795	prod 6.1.1 - topCode 6.1.4	246,742
Total			1,906,818

Table 6.1: CDF datasets used for this analysis. The table shows the available run range, the version of the production and reconstruction software and the corresponding integrated luminosity for each dataset.

cost of introducing a sizeable and difficult to model bias as far as the b-tagging algorithm is concerned. Additionally, triggers with explicit missing E_T requirements can reduce the initial background amount in the triggered data sample, but they enhance the EWK+jets component with respect to the QCD-dominated fraction of events, which is essential to parameterize background b-tagging rates, as will be described in Sec. 6.9.

For these reasons, our choice is to use the TOP_MULTI_JET trigger which provides, at the first order, a QCD-dominated sample in which background prediction tools can be developed and used to estimate the background to $\not\!\!E_T + jets\ t\bar{t}$ decays.

The results reported in this work are based on data collected from March 2002 to March 2007 by the Collider Detector at Fermilab using the TOP_MULTI_JET trigger. With the requirement of fully operational silicon detectors, calorimeters and muon systems, the total integrated luminosity used in the analysis and corresponding to this period is 1.9 fb^{-1}. Additional details about the datasets used in this analysis are reported in Tab. 6.1.

The main features of the decay channel we want to study are the following: first of all, the W boson from the t-quark decaying leptonically yields a considerable amount of missing transverse energy $\not\!\!E_T$, whose direction in the transverse plane $r-\phi$ is expected to be uncorrelated with respect to any jet direction in the event. Moreover, each $t\bar{t}$ event contains two b-jets whose presence can be established by using the SECVTX tagging algorithm.

6.3 $\not\!\!E_T$ and $\not\!\!E_T$ significance

We recall that the missing transverse energy, $\vec{\not\!\!E}_T$, is a two component vector $(\not\!\!E_{Tx}, \not\!\!E_{Ty})$ whose raw value is defined by the opposite of the vector sum of the transverse energy of all calorimetric towers:

$$\vec{\not\!\!E}_T^{\,raw} = -\sum_{towers} (E_T^i)\vec{n}_i \tag{6.1}$$

where E_T^i is the transverse energy of the i-th calorimeter tower, and $\vec{n}_i$ is a transverse unit vector pointing from the center of the detector to the center of the tower.

The $\not\!\!E_T$ is the only observable signature that genuine neutrinos from W leptonic decays leave in CDF II detector. However missing transverse energy can be also produced by jet energy mismeasurement and by b-quark semi-leptonic decays. The former is an instrumental effect that can be partly accounted for with the application of jet energy corrections; the second is due to possible decays of b hadrons into $\nu + X$, that yield some missing energy oriented along the jet direction.

The resolution on the $\not\!\!E_T$ measurement is observed to scale as the square root of the total transverse energy $\sum E_T$ [11], so for this reason the $\not\!\!E_T$ significance defined as:

$$\not\!\!E_T^{sig} = \frac{\not\!\!E_T}{\sqrt{\sum E_T}} \tag{6.2}$$

is expected to be more discriminant than the $\not\!\!E_T$ as an analysis cut. In our analysis $\sum E_T$ will be over all jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ in the event, where E_T^{L5} is the jet L5-corrected energy, and we will refer to them as to *tight jets*.

Corrections to the $\not\!\!E_T$

As seen in 4.4, several corrections have to be applied to the $\not\!\!E_T$ to account for the actual primary vertex location, as well as to correct for the presence of high-P_T muons, and finally to propagate the effect of the jet energy corrections to the missing E_T measurement. We can summarize the corrections as follows:

- Vertex correction: since the geometric center of the CDF detector is used for the raw $\not\!\!E_T$ evaluation, the $\not\!\!E_T$ is recalculated using the primary vertex of the interaction.
- Muon corrections are then applied to account for the low energy deposits in the calorimeter released by high-P_T muons.
- Jet corrections are propagated to the $\not\!\!E_T$ measurement: the $\not\!\!E_T$ is recomputed after previous corrections taking into account the corrections applied to jets.

Regarding the last item, in this analysis we will use *tight jets*, i.e. jets reconstructed within the pseudorapidity range $|\eta| < 2.0$ with $E_T^{L5} \geq 15\ GeV$, where E_T^{L5} denotes the jet L5-corrected energy. We note that this cut on the value of jets E_T^{L5} has been chosen in order to enforce a jet energy threshold of $E_T^{raw} \geq 10\ GeV$ acting at trigger level, according to the correlation between uncorrected jet energies and $L5$-corrected values already observed in [2] for multijet data.

On the basis of studies already available in [12] we decided to adopt $L5$ jet corrections. The application of jet energy corrections can in fact alter the shape and the characteristics of the $\not\!\!E_T$ and $\not\!\!E_T$ significance distributions both for $t\bar{t}$ and background data: Fig. 6.1, taken from [12], shows the comparison of $\not\!\!E_T$ and $\not\!\!E_T$ significance cuts discrimination power for jet corrections up to level 7. The conclusion of this aproach is that $L5$ corrections for jets, when accounted for in the $\not\!\!E_T$ and $\not\!\!E_T$ significance calculation, provide the best signal to noise discrimination, and will thus be adopted for this analysis.

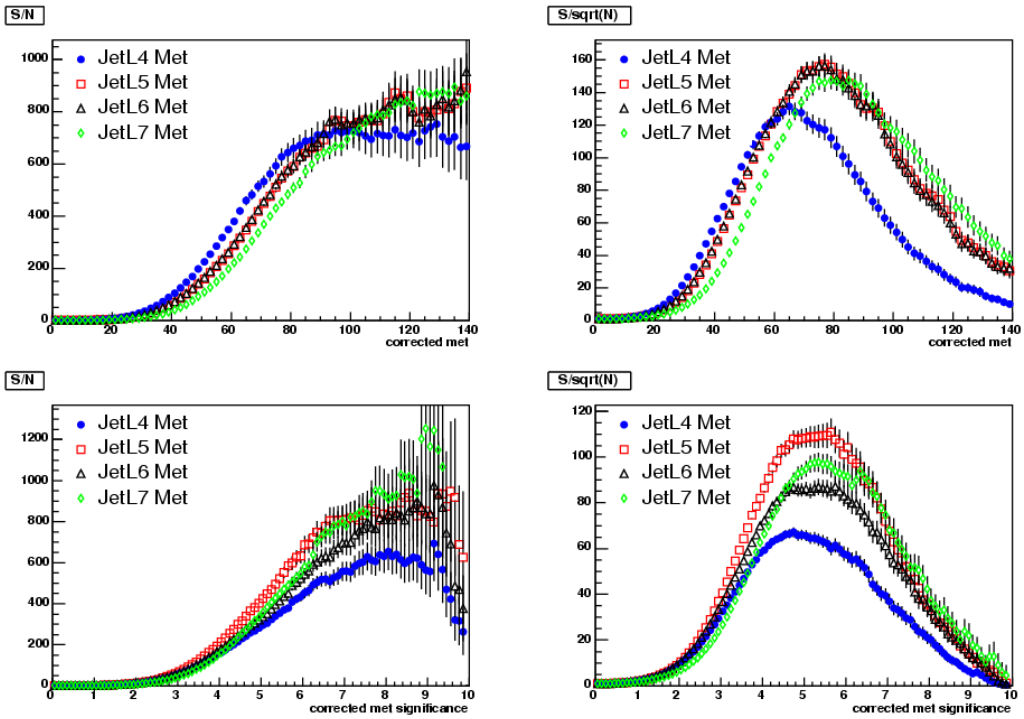


Figure 6.1: Cut S/B and $S/\sqrt{B}$ optimization studies performed using both $\not\!\!E_T$ and $\not\!\!E_T{}^{sig}$ distributions for Monte Carlo signal and multijet data as a function of the applied jet correction level. Figure is taken from [12].

All these corrections are very important, since a good knowledge of the $\not\!\!E_T$ of the event is essential to isolate the decay channel we are interested in; not only the $\not\!\!E_T$ absolute value is of great importance, but also the direction of the $\not\!\!E_T$ in the $r-\phi$ plane can provide an interesting handle to discriminate the possible sources of missing transverse energy on a geometrical basis. In fact, for example the $\not\!\!E_T$ due to neutrino production in leptonic W boson decays is generally uncorrelated with any jet direction in the event, so if we define the $DPhiMin = \min \Delta\phi(\not\!\!E_T, jet)$ as the minimum angular difference between $\not\!\!E_T$ and each jet in the event, we expect to observe large values of $DPhiMin$ in the cases of $W \rightarrow l\nu$ decays and of $t\bar{t} \rightarrow \not\!\!E_T + jets$ events. On the other hand, since for background events the main source of $\not\!\!E_T$ is represented by jet energy mis-measurement, the $\not\!\!E_T$ is expected to be aligned with the jet direction, thus providing values of $DPhiMin$ peaked around zero.

It is important to note that high $\not\!\!E_T$ significance uncorrelated with jet direction can still be produced by processes different from $t\bar{t}$ production: for example $W \rightarrow l\nu$ can be produced in association with jets giving the same missing energy signature as the $t\bar{t} \rightarrow \not\!\!E_T + jets$ decays. To further reject these kind of processes, we can rely on the additional requirement of at least one identified b-jet in the event using the SECVTX algorithm.

6.4 b-jet identification efficiency and scale factor

The b-jet identification is performed using the SECVTX algorithm described in Sec. 4.5.

SECVTX b-jet identification efficiency cannot be determined only on a Monte Carlo basis: imperfect detector descriptions, difficult to model tracks coming from underlying events, multiple interactions which are not modeled in the Monte Carlo, different heavy flavour contents of the various samples, raise the need to measure the b-tagging efficiency directly from data and then to introduce a data-to-Monte Carlo scale factor to account for the differences.

If we want to estimate the efficiency of the SECVTX tagger directly using data events, we need to identify a control sample made only of pure b-jets. Next we need to examine the ratio of the b-tagging efficiencies as measured in the data and in the Monte Carlo and to correct accordingly the Monte Carlo derived efficiency (i.e. applying the so-called SECVTX scale factor, or SF). By doing so, the efficiency of the SECVTX tagger in a given signal sample (such as the $t\bar{t}$ sample) is given by rescaling the measured Monte Carlo efficiency according to the scale factor estimate.

We can use dijet events which have a lepton within one jet ("lepton-jet" events) as a b-enriched control sample, and as an additional prerequisite on the sample we can require the presence of at least one tagged jet back-to-back with respect to the lepton jet (a so called "away-jet"). Using this selection, we end up with a heavy flavour enriched sample thanks to the requirement of a lepton within the jet, which is consistent with a semileptonic b-quark decay, and to the presence of a tagged away jet, which means that we are preferentially selecting $b\bar{b}$ events. Next step is to calculate the b-tagging rates in the selected sample in order to determine the b-jet identification efficiency. Additional complications can arise mainly because of the possible presence of a residual light flavour contamination to the lepton-jet tags. In order to account for this effect a combination of two methods, the electron and muon method, is adopted.

The electron method [13] makes use of conversions in order to calculate the residual light flavour contribution to the lepton-jet tags, by comparing the tag rates in jets where the electron is found to be part of a conversion with non-conversion jets, and attributing the enhancement to heavy-flavour processes. On the other hand, the muon method [14] uses a Monte Carlo template of the transverse momentum of the muon relative to the jet axis to fit to data distribution, and to determine the fraction of untagged and tagged jets attributable to b-quarks, thereby extracting the tagging efficiency for such jets.

Both methods rely on the following assumptions:

- the scale factor for tagging both jets in the event is the same as the scale factor for tagging only one of them, i.e. that the scale factor is the same for single and double tagged events;
- the tagging on the lepton side is uncorrelated with the tagging on the away side;
- the scale factor is the same for b- and c-jets.

Finally, a combination of the two scale factor measurements can be performed by maximizing a generalized likelihood that requires the knowledge of the correlation between the two scale factor measurements and the associated systematics [15, 16]. The combined result provides a SECVTX scale factor determination of $SF = 0.95 \pm 0.050$.

Misidentifications

We call *mistag* or fake tag a positive SECVTX tag on a jet that does not contain heavy flavour; this kind of misidentification by the b-tagging algorithm may be due to several reasons. For example, some false tags can arise from tracking resolution effects: when several tracks have large displacement significancies, they can combine to form a mistag. This effect can be reduced by selecting good-quality vertices with large L_{2d} displacement. Additionally, some mistags can be produced by long-lived particles, such as K_0^S and Λ, decaying into light-flavour jets. These can be reduced requiring the total mass of the tracks inside the tags to fall outside opportune mass windows around these particles. Finally, b-jet misidentifications can be due to material interactions or conversions on the beampipe or on inner silicon detector layers. These effects can be reduced by disallowing two-track vertices reconstructed within the region occupied by the detector material. Even if the amount of mis-identification can be partially reduced, any method is not 100% effective.

Since mistags due to limited detector resolution are expected to be symmetric in the signed $2D$ displacement L_{xy} of the vector separating the secondary and primary vertices, one can then use the ensemble of negative tagged jets ($L_{xy} < 0$) in order to estimate the residual light flavour jet contribution to the positive tag sample.

Tagging efficiency and mistag rate

The efficiency of the SECVTX algorithm is defined as the fraction of fiducial b-jets that possess a positive b-tag. Fiducial jets are defined according to the following requirements: $E_T^{raw} > 10\ GeV$ and $|\eta| < 2.0$. Figures 6.2(a) and 6.2(b) show the SECVTX efficiency times scale factor in $t\bar{t}$ events versus jet E_T and η, respectively; figures 6.2(c) and 6.2(d) show the SECVTX negative tag rates versus jet E_T and η, respectively. Performances for both the tight and loose versions of SECVTX are shown, even if only the tight (blue) version of the algorithm is used in this analysis. The error bands for the efficiency are derived from the b-tagging data-to-Monte Carlo scale factor (SF) uncertainties.

The efficiency curve rises as a function of jet E_T and then falls down. This is due to the imposed cuts on the maximum allowed vertex radius, and to the veto on vertices with 2 tracks within material regions. This affects the efficiency at high jet E_T where b-hadrons are more boosted, and have a higher probability of reaching large radii before decaying. The efficiency is flat in the $|\eta| < 1.0$ range, but then falls off due to reduced COT coverage for higher $|\eta|$ values. The negative tag rate also rises as a function of jet E_T, however it doesn't show the same drop-off as the efficiency. The negative tag rate also increases with jet $|\eta|$, and then falls off

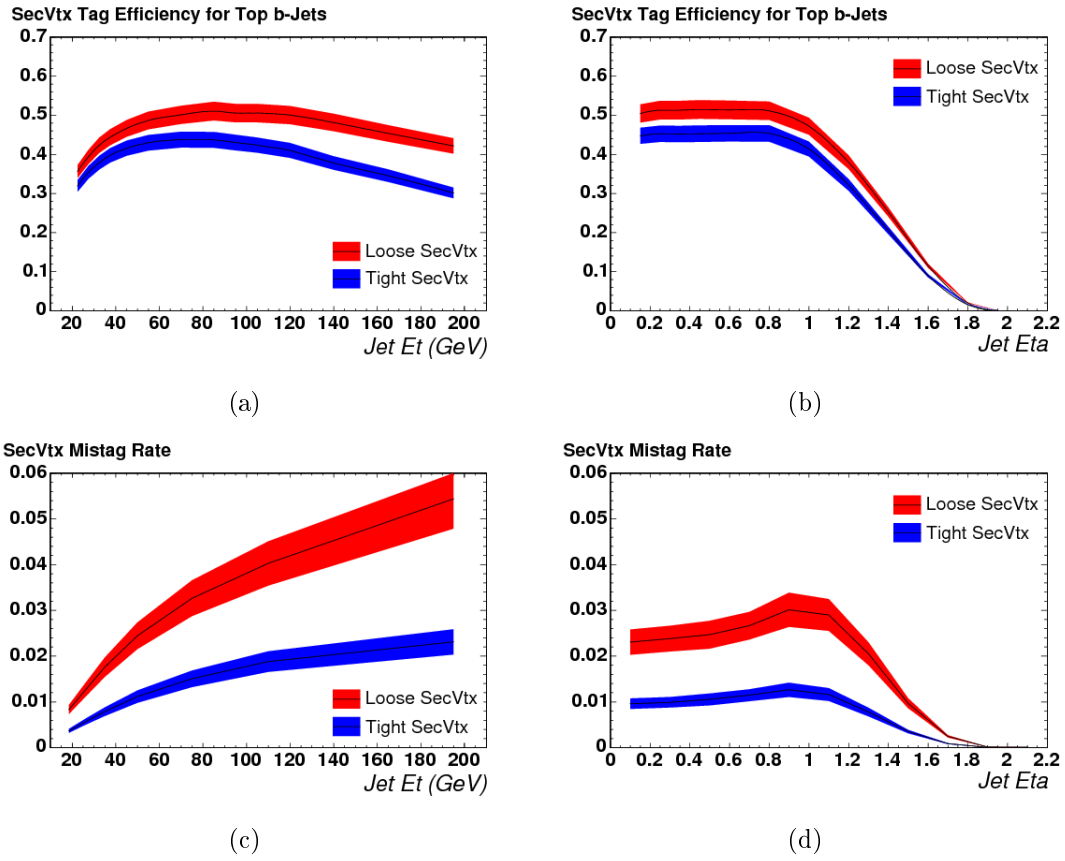


Figure 6.2: SECVTX tagging efficiency and mistag rate. Both tight (blue) and loose (red) options of the SECVTX algorithm are shown. See text for details.

as silicon coverage decreases. The initial increase is due to the fact that as jet $|\eta|$ increases, the tracks in the jet pass through more and more material, and the tracking algorithm becomes steadily worse due to multiple scattering. The result is an increase in the fake rate in that case.

In order to define our final sample to be used for the $t\bar{t}$ production cross section measurement, we will require the presence of at least one SECVTX-positive tagged jet in the selected events.

6.5 Additional kinematical variables

Besides the $\not\!\!E_T$, $DPhiMin$ and the b-tagging requirements, other kinematical variables related to the topology of the event or to its energy can be used to characterize the $t\bar{t}$ production with respect to background processes. In the following we will define some topological variables called Aplanarity, Centrality and Sphericity [17] in order to give a description of the jet activity in the event.

For each event we can define the following normalized momentum tensor M_{ab}:

$$M_{ab} = \frac{\sum_j P_{ja} P_{jb}}{\sum_j P_j^2} \tag{6.3}$$

where a, b run over the three space coordinates and P_j is the momentum of the jet j.

We are interested in finding the axis $\vec{n}$ such that the normalized sum of the square components of the jet momenta along it has a maximum

$$max \frac{\sum_j (\vec{P}_j \cdot \vec{n})^2}{\sum_j \vec{P}_j^2}. \tag{6.4}$$

This quantity can characterize the space direction distribution of the jet momenta, i.e. topologies where the jets are mostly along the direction $\vec{n}$ with respect to isotropical distributions. The ratio in 6.4 can be written as:

$$\frac{\sum_j (\vec{P}_j \cdot \vec{n})^2}{\sum_j \vec{P}_j^2} = \sum_{a,b=1}^{3} n_a \frac{\sum_j P_{ja} P_{jb}}{\sum_j \vec{P}_j^2} n_b = \sum_{a,b=1}^{3} n_a M_{ab} \tag{6.5}$$

M_{ab} is a symmetric and definite positive matrix, so it can be diagonalized. Its unit eigenvectors $\vec{n}_1$, $\vec{n}_2$, $\vec{n}_3$ have corresponding eigenvalues Q_j satisfying the relation $Q_1+Q_2+Q_3 = 1$ since the trace of M_{ab} is null. So, ordering the eigenvalues such as $0 \leq Q_1 \leq Q_2 \leq Q_3$ the axis $\vec{n}$ we are looking for is $\vec{n}_3$, the normalized eigenvector corresponding to the highest eigenvalue.

M_{ab} eigenvalues can be used to characterize the event shape. In particular for roughly spherical events, $Q_1 \approx Q_2 \approx Q_3$; for coplanar events, $Q_1 \ll Q_2$ and finally for collinear events $Q_2 \ll Q_3$. Particular combinations of the Q_j are used to define topological variables.

The *Sphericity* S is defined as:

$$\begin{aligned} S &= \frac{3}{2}(Q_1 + Q_2) = \frac{3}{2}(1 - Q_3) = \\ &= \frac{3}{2}\left(1 - \frac{\sum_j (\vec{P}_j \cdot \vec{n}_3)^2}{\sum_j \vec{P}_j^2}\right) = \frac{3}{2}\left(\frac{\sum_j \vec{P}_{jT3}^2}{\sum_j \vec{P}_j^2}\right) \end{aligned} \tag{6.6}$$

where subscript T denotes momentum component transverse to $\vec{n}_3$ axis. Sphericity values lie in the range $[0, 1]$: S is null in the limiting case where momenta are directed all exactly along $\vec{n}_3$, like a pair of back-to-back jets, while S approaches 1 for events with a perfectly isotropic jet momenta, when $Q_1 = Q_2 = Q_3 = \frac{1}{3}$, thus giving a spherical distribution.

The *Aplanarity* A is defined as

$$A = \frac{3}{2} Q_1 \tag{6.7}$$

and it is normalized to lie in the range $[0, 1/2]$. A is null when the sum of jet momenta has null component on $\vec{n}_1$ axis, and this is the case for coplanar or collinear events. On the contrary, when jet momenta have isotropic distribution $Q_1 = Q_2 = Q_3 = \frac{1}{3}$ and A reaches its maximum $\frac{1}{2}$, so that extremal values of A are reached in the case of two opposite jets and in the case of evenly distributed jets, respectively.

Jets emerging from a $t\bar{t}$ pair are expected to be uniformly distributed and as a consequence they will hardly lie on the same plane: thus we expect high aplanarity and sphericity values for $t\bar{t}$ events and this will give us a handle to discriminate them from the background.

In addition to kinematical variables describing the topology of the event, also distributions of energy-related variables, such as the Centrality, $\sum E_T$, $\sum E_T^3$ can be useful to give a discriminant for $t\bar{t}$ events over their background.

The *Centrality* C is defined as:

$$C = \frac{\sum E_T}{\sqrt{\hat{s}}} \tag{6.8}$$

where $\sqrt{\hat{s}}$ is the center of mass energy in the hard scattering reference frame: $\hat{s}$ is estimated as $\hat{s} = \sqrt{x_1 x_2}/1.96\ TeV$ where $x_1 = (\sum E + \sum P_z)/(1.96\ TeV)$ and $x_2 = (\sum E - \sum P_z)/(1.96\ TeV)$, $\sum E$ is the sum of the energy in the event and $\sum P_z$ is the sum of the z component of the momentum of all jets in the event. In the case of $t\bar{t}$ pairs decaying hadronically, jets are emitted preferably in the transverse plane ($r - \phi$ plane), so we expect to have a greater amount of energy emitted in this plane thus giving values of C closer to 1 with respect to background events.

We recall that the total transverse clustered energy $\sum E_T$ is defined as the jet E_T sum over all tight jets of the event, i.e. jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$.

On the other hand, the $\sum E_T^3$ is defined as the E_T sum over all tight jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ in the event except the two leading ones. In QCD events the two most energetic jets are produced by $q\bar{q}$ processes while the least energetic ones come from gluons *bremmsstrahlung*; on the contrary, in $t\bar{t}$ events up to 6 jets can be produced by hard processes, and as a consequence $\sum E_T^3$ can help us discriminating signal and background.

Another kinematical variable we will use is E_{T1}, the energy of the leading jet in the event.

6.6 Event Prerequisites

Before going into the details of Neural Network training, it is useful to define a set of clean-up cuts which will reject those events we are not interested in analyzing. First of all we will exclude events collected when the detector is not under optimal conditions (i.e. with partial functionality of the silicon, muon or calorimeter detectors) or reconstructed in regions not fully covered by the CDF II instrumentation. Moreover, we will preliminary reject events with well reconstructed high-P_T leptons in order to guarantee orthogonality with respect to other $t\bar{t}$ cross section analyses relying on the lepton+jets decay signature [18]. In addition to this, we will also reject events with low $\not\!\!E_T$ significance, enforcing the requirement $\not\!\!E_T{}^{sig} \geq 3\ GeV^{1/2}$: this will also assure the orthogonality of our cross section measurement with respect the all-hadronic one [19].

The following prerequisites will be applied both to data and Monte Carlo samples:

- A *Run* is a set of events collected under the same conditions of the detector and on a time window of $6 \div 12$ hours. We use the Good run list v17 provided by the CDF Data Quality Group requiring runs to have silicon, muon and calorimeter detectors fully operational [20].

- We discard events whose primary vertex location is not well centered in the CDF II detector, in particular:

  - In order to select well centered events, the z coordinate of the highest-$\sum P_T$ good quality vertex is required to be within ± 60 cm from the geometrical center of the detector: $|z_{vert}| < 60$ cm.
  - We require that the vertex used for jet reclustering and then for the secondary vertex search is close to be the primary vertex found in the event by means of the *PrimVtx* algorithm described in Sec. 4.2. So we require the distance between the event primary vertex and the vertex used for jet reclustering $|z_{jet} - z_{primvtx}|$ to be less than 5 cm, where z_{jet} denotes the z_0 of the good quality highest-P_T vertex.
  - A good quality vertex, by definition, is formed with at least three COT tracks [21]. We require the number of good quality vertices in the event to be greater than zero.

- We reject events with a good, $high - P_T$ reconstructed electron or muon to avoid overlaps with other top lepton+jets analyses.

- We clean up our sample by requiring events to have at least 3 *tight* jets, i.e. jets with $E_T^{L5} \geq 15$ GeV and $|\eta| \leq 2.0$.

- We reject events with low $\not\!\!E_T$ by requiring $\not\!\!E_T{}^{sig} \geq 3$ $GeV^{1/2}$, thus avoiding overlaps with top all-hadronic analyses.

- We simulate the new $L2$ trigger requirements (See Sec. 6.2) on data taken before run 194328, to have an homogenous sample to perform our analysis.

- When dealing with Monte Carlo events, we perform on each event a full simulation of the new TOP_MULTI_JET trigger path.

The impact of these preliminary selections on data and inclusive Monte Carlo $t\bar{t}$ is shown in Tab. 6.2 and Tab. 6.3. After prerequisites application we expect a signal to background ratio S/B of 1.33% in the sample with $NJets \geq 3$, of 0.12% in the sample with exactly 3 tight jets and of 1.83% in the sample with $NJets \geq 4$. In the following this negligible signal contamination will allow us to train a neural network using all data events after prerequisites as background and to determine an effective b tag parameterization to be used to predict the amount of background b tags in our network selected sample.

6.7 Neural Network Training

As previously discussed, in order to enhance the signal to background ratio in our final sample, we will use a neural network, trained to discriminate

N evts	gset0d	gset0h	gset0i	gset0j	tot.
Tot. Events	7219495	3802935	4018550	1222587	16263567
Good Run	4750786	3185795	3579217	1196129	12711927
Trigger	1243047	2012475	3579216	1196129	8030867
$\|Z_{vert}\| < 60\ cm$	1162209	1770306	3227545	1150029	7310089
$\|Z_{jet} - Z_{primvtx}\| < 5\ cm$,					
N_{vert} good quality ≥ 1	1127916	1665737	3077048	1051121	6921822
N tight leptons $= 0$	1126273	1663557	3072541	1049820	6912191
NJets ≥ 3	1088740	1562059	3001054	1013434	6665287
$\not\!\!E_T{}^{sig} \geq 3\ GeV^{1/2}$	14403	23808	41376	17652	97239
Out of which:					
with NJets$= 3$	4220	8884	10190	5166	28460
with NJets≥ 4	10183	14924	31186	12486	68779

Table 6.2: Events surviving the clean-up requirements for data, divided in each period of data taking.

N evts	MC_{incl}	eff.(%)	evts in 1.9 fb^{-1}
Tot. Events	4719385		
Good Run	4658603		
L2 Trigger	2786636	59.82	7642
L3 Trigger	2719975	97.61	7459
$\|Z_{vert}\| < 60\ cm$	2610396	95.97	7159
$\|Z_{jet} - Z_{primvtx}\| < 5\ cm$,			
N_{vert} good quality ≥ 1	2607087	99.87	7150
N tight leptons $= 0$	2333998	89.53	6401
NJets ≥ 3	2333351	99.97	6399
$\not\!\!E_T{}^{sig} \geq 3\ GeV^{1/2}$	464067	19.89	1273
Out of which:			
with NJets$= 3$	12058		33
with NJets≥ 4	452009		1240

Table 6.3: Events surviving the clean-up requirements for inclusive Monte Carlo $t\bar{t}$ samples. Last column shows the amount of $t\bar{t}$ events expected in 1.9 fb^{-1} of data.

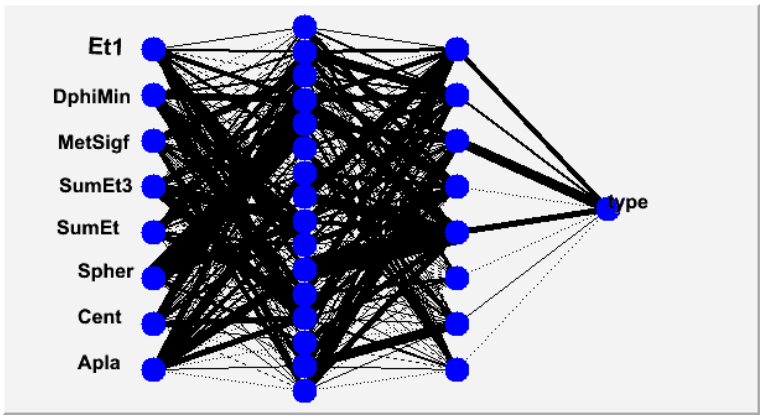


Figure 6.3: The 8-16-8-1 topology of the network used in the analysis: a feed forward neural network with 2 hidden layers, 8 input nodes and one single output for classification. The thickness of the black lines connecting each perceptron is proportional to the associated weight.

$t\bar{t} \rightarrow \not\!\!E_T + jets$ signal events from background. We will use the class *TMultiLayerPerceptron* available in ROOT to build a software abstraction of the network. For what concerns training samples, as background we will use all the data taken with the TOP_MULTI_JET trigger and passing the prerequisites previously discussed; additionally, we will require the presence of at least 4 tight jets in the event (i.e. jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$) to perform the training in a sample completely uncorrelated with the one we will use to determine a background parameterization. For signal we will use the same amount of events passing the same requirements of the data, taken randomly from the available Monte Carlo samples. As seen in the previous section, since S/B is negligible in the data sample with $NJets \geq 4$ obtained after prerequisites application, we can use all these data events for background in our neural network training without affecting its rejection power.

We used the topology depicted in Fig. 6.3, using as inputs for the network the following kinematical variables, normalized with respect to their maximum value:

- E_{T1}, the transverse energy of the leading jet;
- $DPhiMin$, already defined as $\min \Delta\phi(\not\!\!E_T, jet)$, the minimum difference between the $\not\!\!E_T$ and each jet in the event in the ϕ coordinates;
- $\not\!\!E_T{}^{sig}$, the $\not\!\!E_T$ significance of the event, defined as $\not\!\!E_T/\sqrt{\Sigma E_T}$;
- the energy-related variables $\sum E_T$, $\sum E_T^3$ and the Centrality;
- the topology-related variables Sphericity and Aplanarity.

Fig. 6.4 shows the signal versus background distributions of each input variable going into the network after the application of the previously discussed prerequisites. The obtained sample made of signal and background events will be split in two parts: half will be used for neural network training and the other half for the

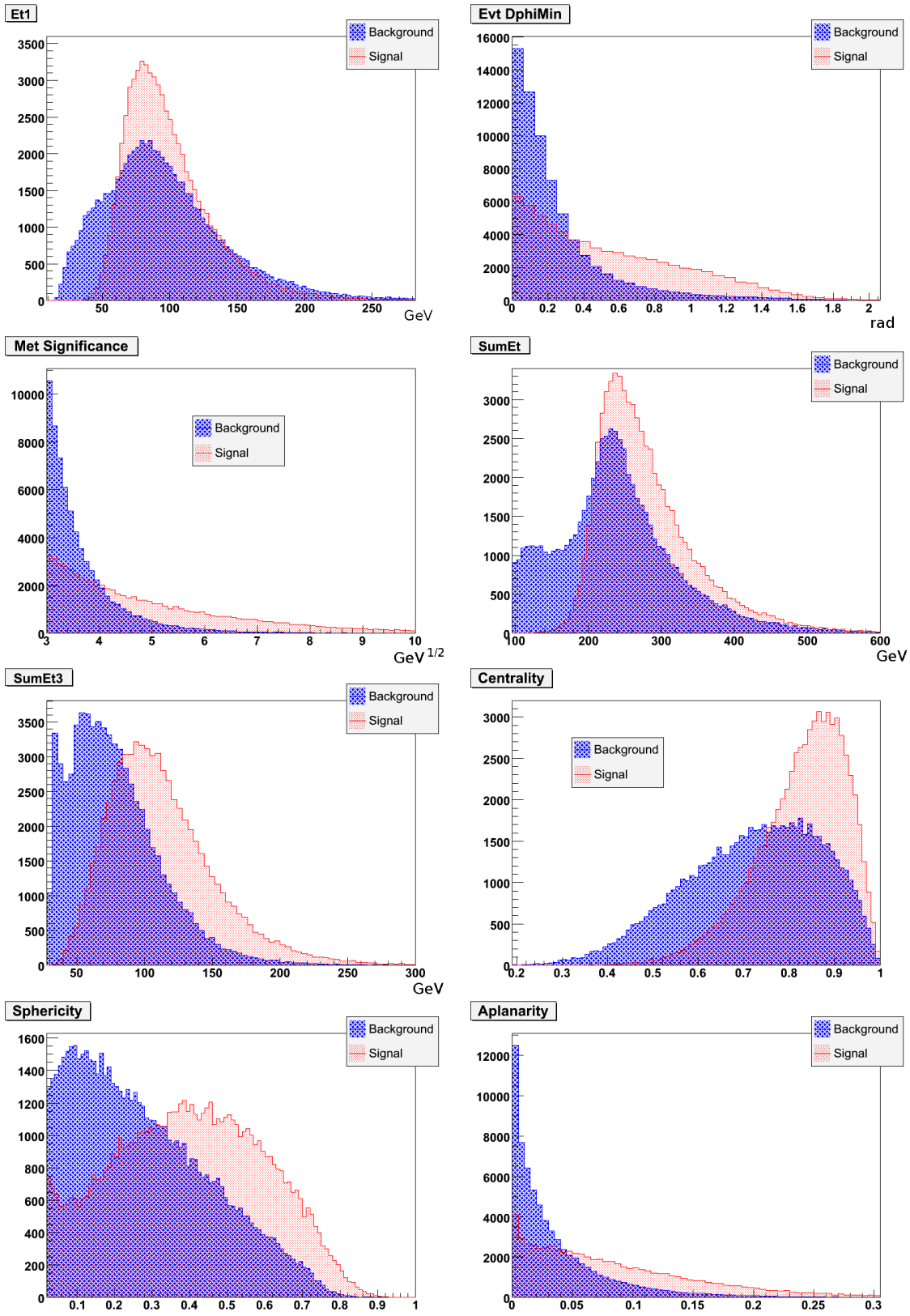


Figure 6.4: Distribution of neural network input variables for top multi jet data (background) and $t\bar{t}$ Monte Carlo (signal) samples, after prerequisites application (see text for details).

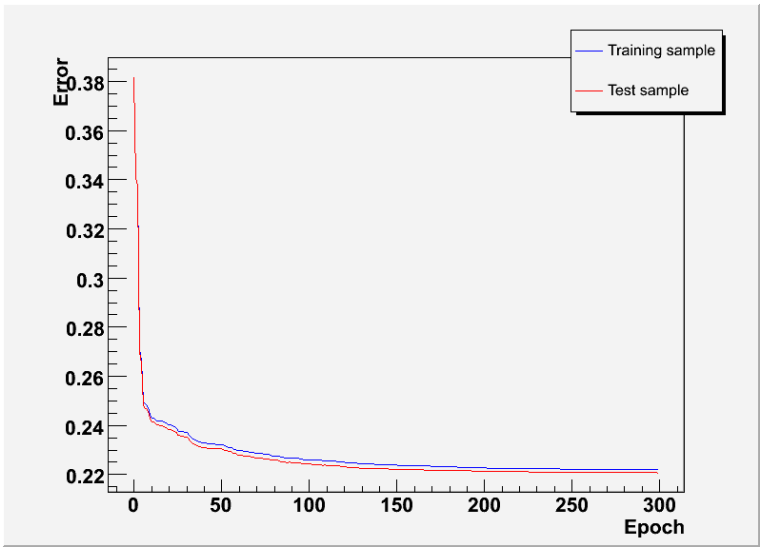


Figure 6.5: Average neural network error during training on training and test samples.

so called testing during each iteration of the training procedure; as previously discussed, we will use the ROOT implementation of the BFGS optimization method (see Sec. 5.3.2). A plot describing the "history" of the training is shown in Fig. 6.5: for each training epoch the average error made by the network in trying to discriminate events belonging to the signal or background class is calculated both for the events in the training sample and in the test one (see Sec. 5.3 for details).

We stop our training pocedure after 300 epochs, since after this number of iterations the network reaches the minimum of the error function for the chosen topology. Additionally, we want to avoid a situation of *overtraining*: overtraining happens when a neural network learns "too well" the details of the training set, getting stuck in the statistical fluctuations of its input variables, and looses the capability of generalizing its results on a different sample. The fact that errors on the training sample and on the test one are almost the same over all the training period tells us that the network has not been overtrained.

The neural network obtained after the training procedure is then applied to all the available events (training + test samples), its output is shown in Fig. 6.6: signal and background are well separated and their distributions are well peaked around their expected values. The performances of the neural network obtained will be briefly described using the quantities defined in Sec. 5.3: the efficiency of the network is good over all possible cuts on the output variable, while purity as a function of the cut on the output variable has a good trend, showing low background contamination for high cuts, as shown in Fig. 6.7. We recall that the purity parameter does not refer directly to the purity of the final sample we will use for the cross section measurement: in fact it is just a measure of the performances of the network, being calculated submitting to the network a sample made of the same number of signal and background events. Finally, the efficiency versus purity plot approaches quite well the ideal "step" one, as shown in Fig. 6.8.

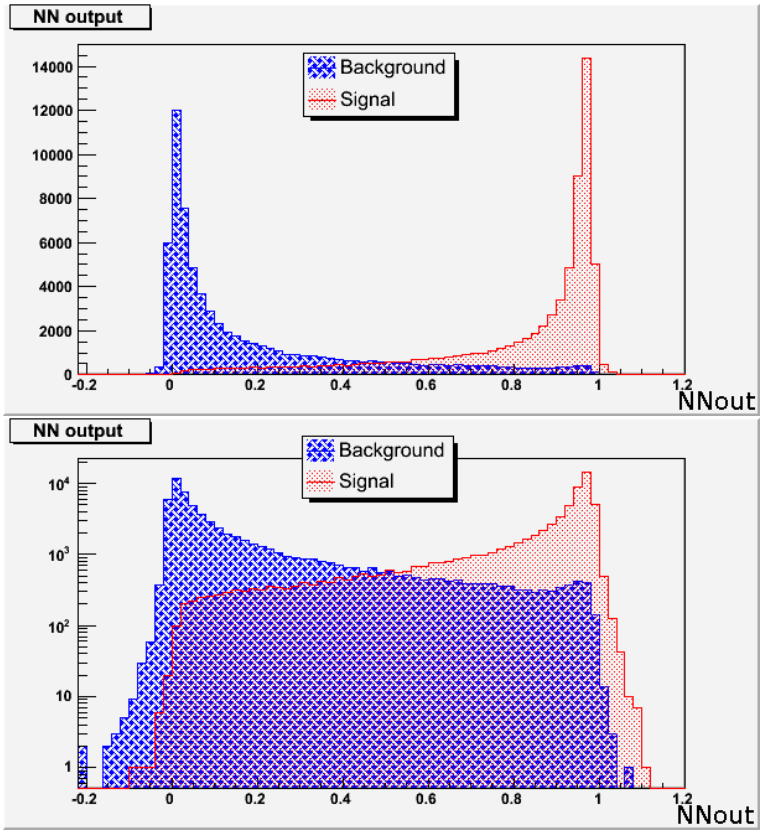


Figure 6.6: Output of the neural network after the training, bottom figure shows the same plot in log scale.

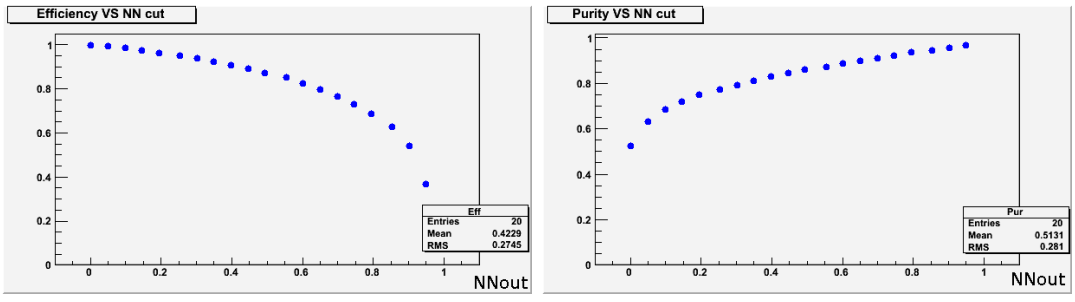


Figure 6.7: Performances of the Neural Network after training: efficiency vs cut on the output variable on top and purity vs cut on the output variable on the bottom.

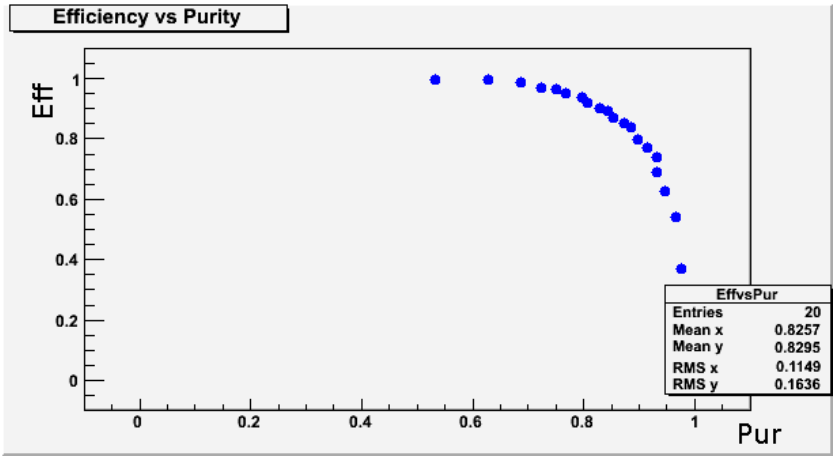


Figure 6.8: Efficiency versus Purity plot of the network obtained after the training.

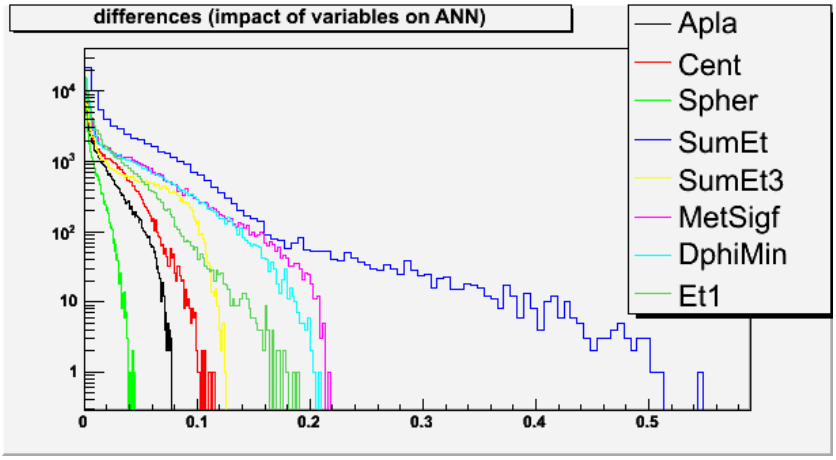


Figure 6.9: Impact of the input variables on the output of the neural network (see text for details).

Another important variable we can use to characterize the neural network obtained after the training is the impact of each different input variable on the output of the network itself. A way to estimate this quantity is the following: we choose a fixed input variable α and, for each event, while keeping all the other input variables untouched, we shift the value of the α_i input by $\pm\frac{1}{10} \cdot RMS$, where RMS denotes the root mean square ($\sqrt{\sum_i \alpha_i^2}$) of that input variable calculated over all events submitted to the network. The output of the network after the shift of this single input is calculated and then compared to the output of the network without the shift. Finally, the square root of the difference of the squares of the 2 outputs is calculated and then used to fill an histogram. This is repeated for every variable and for each event in the sample, and provides a way to quantify how the output of the network depends on the fluctuations of each single input variable. The result of this procedure is shown in Fig. 6.9: it is easy to notice how $\sum E_T$ and $\not{E}_T{}^{sig}$ variations have the most determinant impact on the output of the network.

6.8 Background estimation

In the following we will describe our background prediction method aimed at measuring the $t\bar{t} \to \not{E}_T + jets$ production cross section.

Our analysis setup is based on the idea, already developed for instance in [22], that it is possible to discriminate $t\bar{t}$ production from background processes in a given kinematically selected sample using their different b-jet identification rates, meaning that the SECVTX tagging probability for a b-jet produced by top quark decay is expected to be higher than the probability of identifying b-quark jets yielded by background processes.

The cross section measurement will then exploit the excess in the number of b-tagged jets over the background expectation:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{exp}}{\epsilon_{kin} \cdot \epsilon_{tag}^{ave} \cdot L} \tag{6.9}$$

where N_{obs} and N_{exp} are the number of b-tagged jets observed and expected from background parameterization, ϵ_{kin} is the combined trigger and kinematical selection efficiency on inclusive Monte Carlo $t\bar{t}$ events; ϵ_{tag}^{ave} is the average number of b-jets per $t\bar{t}$ event, and finally, L is the integrated luminosity of the TOP_MULTI_JET data sample.

In the following we will try to obtain a reliable prediction of the total amount of b-tags coming from background events, which will then be a part of the neural network selection optimization procedure on the data sample. Given our tight prerequisite cut on $\not{E}_T/\sqrt{\Sigma E_T}$ and the ≥ 1 positive b-tag requirement that will be enforced on the final sample, we expect the main background contributions to come from events like $b\bar{b} + jets$ and $Wb\bar{b} + jets$ [23].

In order to determine the background parameterization, the complete data sample obtained in the previous discussion can not be directly used since it has a sizeable signal contamination. Making the assumption that the per-jet positive tagging rate does not depend on the number of jets in the event, we will limit ourselves to the subsample of events with exactly 3 tight jets (i.e. jets with

$E_T \geq 15\ GeV$ and $|\eta| \leq 2.0$), where the $t\bar{t}$ fraction is totally negligible, and we will use this background-dominated sample to derive a per-jet b-tagging probability parameterization for events that are not top-like. We will then check the parameterization predictions for higher jet multiplicities and use it for the background determination.

6.9 Positive b-tagging rate parameterization

As previously discussed, the background rejection power provided by the neural network alone is not sufficient to isolate completely the $t\bar{t}$ events present in our data sample after the application of the prerequisites; the b-tagging algorithm is necessary to enhance the signal presence, but in order to derive a cross section measurement from the final tagged sample we need to find an estimate of the number of b-tagged jets yielded by background processes. Once obtained this estimation, we will use the amount of b-tagged jets expected from background processes in a given selected sample to optimize the cut on the neural network output, with the aim of minimizing the expected statistical uncertainty on the cross section measurement; we will rely on an estimate of both the amount of expected b-tagged jets from inclusive Monte Carlo $t\bar{t}$ and background events to perform such an optimization.

In the following we provide a description of the approach we adopted in order to estimate the background contribution in terms of b-tagged jets yielded by processes other than $t\bar{t}$ production.

The basic idea of our background prediction method rests on the assumption that b-tag rates for $t\bar{t}$ signal and background processes show differences that are due to the different properties of the b-jets produced by the top quark decays compared to the b-jets arising from QCD and vector boson plus heavy flavour production processes. In this hypothesis, parameterizing the b-tag rates as a function of some chosen jet characteristics, in events depleted of signal contamination, will allow to predict the number of b-tagged jets from background processes present in a given selected sample.

We summarize below the steps needed for this approach:

1. identify a subsample of data with negligible $t\bar{t}$ contamination;

2. in the identified sample, parameterize the b-tagging rate as a function of the N variables on which it mainly depends.

3. Build a N-dimensional b-tagging matrix in order to associate to a given jet a probability to be identified as a b-jet given its characteristics.

4. Predict the total amount of expected background tags in a given sample by summing b-tagging probabilities over all jets in the selected events.

5. In samples depleted of signal, check the matrix background prediction by comparing the number of expected and observed SECVTX tagged jets.

6. Use the tagging matrix to calculate the amount of background tags in the sample to be used for a cross section measurement (i.e. after neural network selection and the requirement of at least 1 SECVTX tag).

We remind that the use of this method based on tagging rate parameterizations rests on the assumption that the sample used for b-tag rates dependencies studies shows a negligible $t\bar{t}$ contamination: a $t\bar{t}$ presence in the sample used to parameterize the tagging rate may have a sizable impact in the amount of background tags prediction. For this reason, we need to choose as base sample a data region depleted as much as possible of signal: in our case, we decide to use for the background tagging rate parameterization the data sample obtained after the prerequisites application with exactly 3 tight jets (i.e. jets with $E_T^{L5} \geq 15\ GeV, |\eta| \leq 2.0$).

Fig. 6.10 and Tab. 6.4 show the number of events in the data sample and the $t\bar{t}$ contamination expected from Monte Carlo assuming the theoretical production cross section of 6.7 pb, corresponding to a top mass of $M_{top} = 175\ GeV/c^2$ for different tight jet multiplicities.

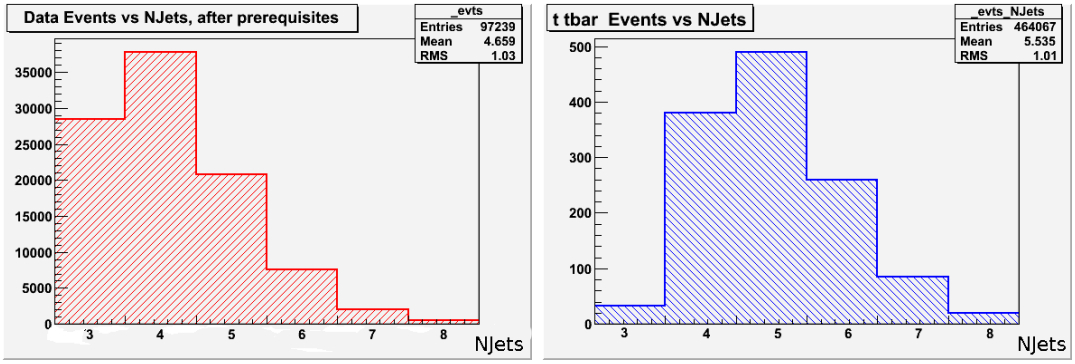


Figure 6.10: Data (left) and inclusive Monte Carlo $t\bar{t}$ (right) events versus number of tight jets (i.e. jets with $E_T^{L5} \geq 15\ GeV, |\eta| \leq 2.0$) in the event after prerequisites. The Monte Carlo expectation is rescaled according to the assumption of a theoretical production cross section of 6.7 pb, corresponding to a top mass of $M_{top} = 175\ GeV/c^2$.

Number of Events	3 jets	4 jets	5 jets	6 jets	7 jets	8 jets
Exp. Inclusive $t\bar{t}$	33	380	490	260	85	20
Data	28,460	37,796	20,743	7,529	2,051	475
Exp. Contamination (%)	0.12	1.01	2.36	3.45	4.14	4.21

Table 6.4: Expected signal contamination for different jet multiplicities. Number of events is also plotted in Fig. 6.10.

6.9.1 b-tagging rate parameterization

We can define the b-tagging probability as the ratio of the number of positive SECVTX tagged jets to the number of taggable jets in the sample of data events

after prerequisites with exactly 3 jets, where we define as taggable a tight jet (again, with $E_T^{L5} \geq 15\ GeV$, and $|\eta| < 2.0$) with at least two good SECVTX tracks (see Sec. 4.5 for details).

The per-jet b-tagging probability has been parameterized as a function of several jet and event variables in order to extract its main dependencies, and is found to depend mainly on jet characteristics such as E_T, the number of good quality tracks contained in the jet cone N_{trk}, and the $\not\!\!E_T$ projection along the jet direction $\not\!\!E_T{}^{prj}$, defined by:

$$\not\!\!E_T{}^{prj} = \not\!\!E_T \cos\Delta\phi(\not\!\!E_T, jet). \tag{6.10}$$

Figures 6.11, 6.12 and 6.13 show both the positive and negative tagging rates dependence on a set of event and jet variables.

Jet E_T and N_{trk} correlation with the tagging probability is expected due to the implementation details of the b-tagging algorithm. The $\not\!\!E_T$ projection along the jet direction is instead correlated with the heavy flavour component of the sample [12, 23] and with the geometrical properties of the event: in fact b-quarks can yield a considerable amount of missing transverse energy due to their semi-leptonic decays and in that case the $\not\!\!E_T$ is expected to be aligned with the jet direction; on the contrary, $\not\!\!E_T$ produced in W boson decays stands more likely away from jets, depending on the process-allowed regions of the phases space. By requiring the events to have large missing E_T significance ($\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$) as an analysis prerequisite, we reject those events whose missing E_T is mainly due to residual energy mis-measurement effects, and in turn concentrate our attention on physics-induced $\not\!\!E_T$.

These $\not\!\!E_T{}^{prj}$ features are depicted in Fig. 6.14. The upper left plot of Fig. 6.14 shows the $\not\!\!E_T{}^{prj}$ for taggable jets in 3-jet inclusive Monte Carlo $t\bar{t}$ events. On the other hand, in the upper right plot the corresponding distribution extracted from 3-jet events in multijet data is shown for comparison. On the second row, the missing transverse energy projection is drawn for SECVTX positive tagged jets, for both the samples.

In general, most of the dependencies observed on the variables in Fig. 6.12 and Fig. 6.13 are weaker than those on the jet E_T, N_{trk} and missing E_T projection: this is the case for the event luminosity, the aplanarity, centrality and sphericity. On the other hand, as far as the $\not\!\!E_T^{sig}$ and $DPhiMin$ dependences are concerned, they are already accounted for by the $\not\!\!E_T$ projection parameterization, so we decided to favour a per-jet variable instead of a per-event one in our matrix parameterization. The number of good quality vertices in the event N_{v12} is found to be discriminant for positive tagged jets but not much for negative ones, and additionally is a per-event variable; we decided not to include it in our parameterization. Finally, jet η is strongly correlated with the number of tracks in the jet (N_{trk}): the higher the track multiplicity the most central the jet is, so we can consider the η dependence to be hidden in the jet N_{trk} parameterization.

For the previous reasons, we decided not to include other variables except the jet E_T, N_{trk} and $\not\!\!E_T^{prj}$ for the b-tagging rate dependence description.

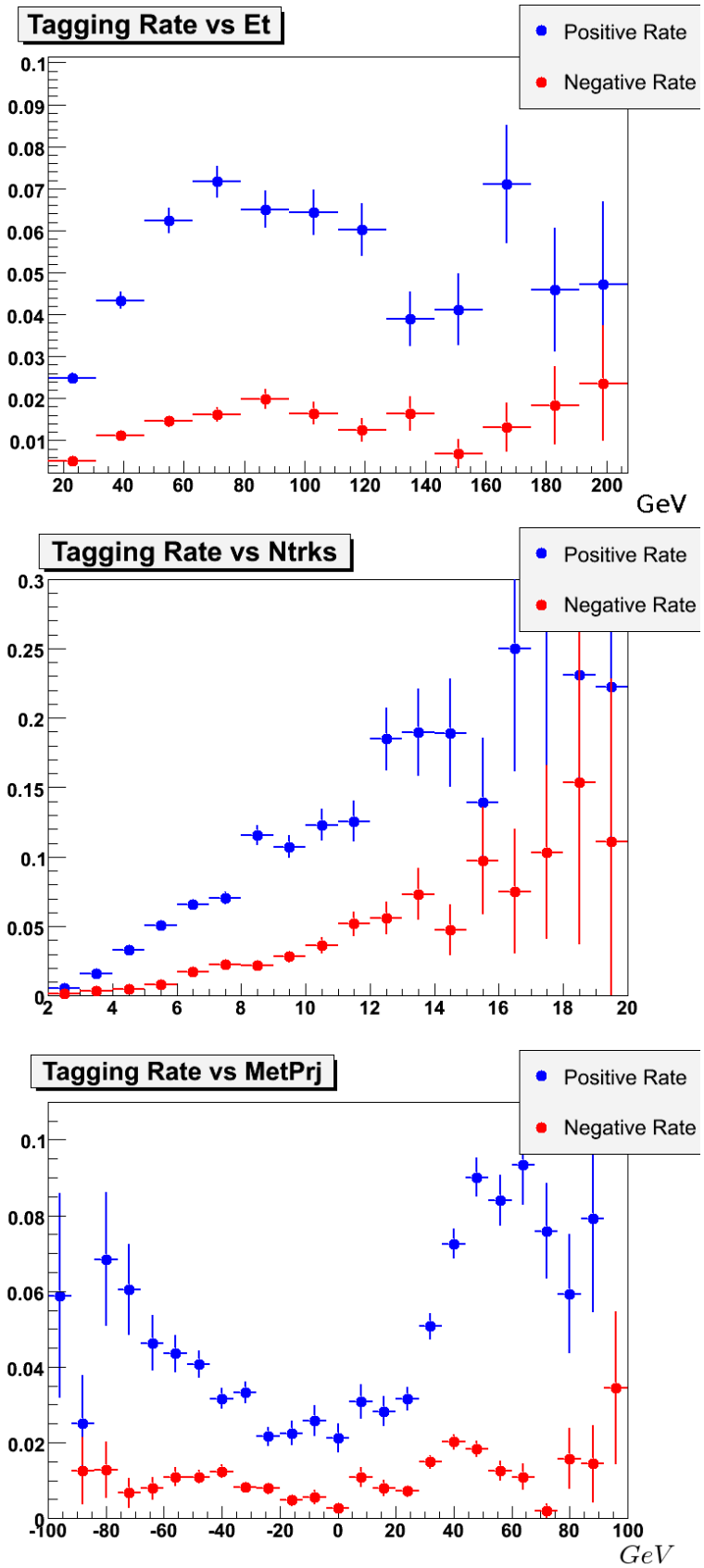


Figure 6.11: Positive and negative b-tagging rates as a function of E_T, N_{trk} and $\not{E}_T{}^{prj}$ for the data sample with exactly 3 tight jets in the event.

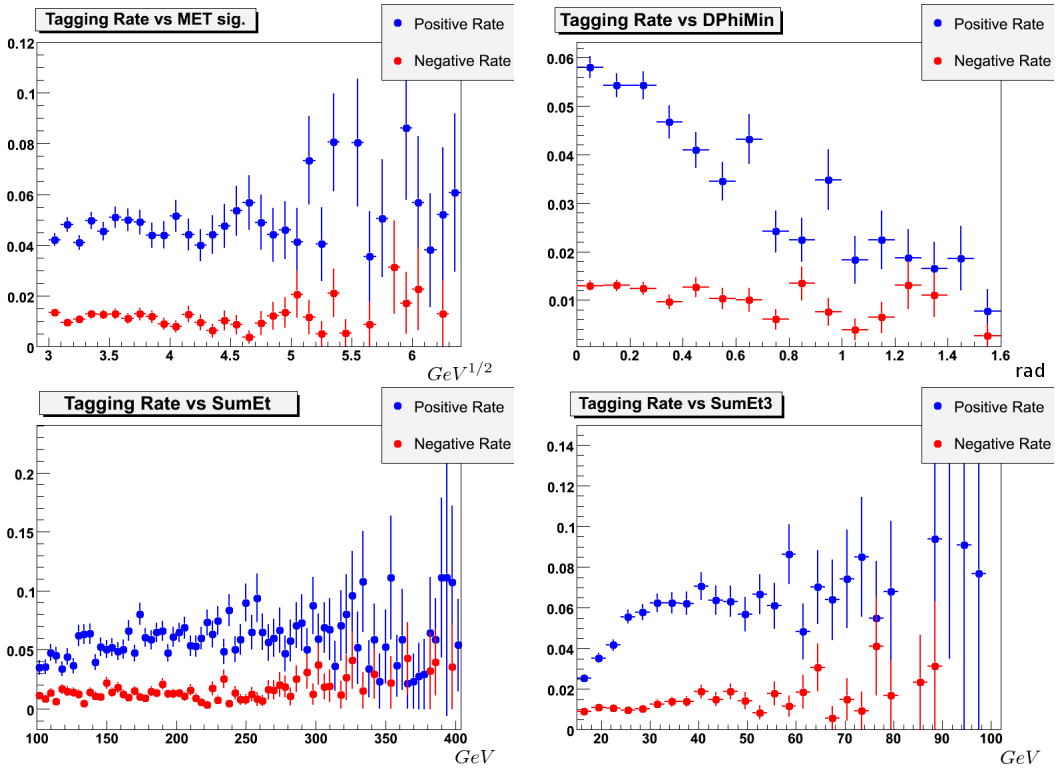


Figure 6.12: Positive and negative b-tagging rates as a function of (from top to bottom, from left to right): $\not\!\!E_T^{sig}$, $DPhiMin$, $\sum E_T$ and $\sum E_T^3$ for 3-jet events.

6.9.2 b-tagging matrix

Now we can define a so-called b-tagging matrix, using the per-jet b-tagging probability dependencies studied previously. The 3-dimensional matrix binning we decided to choose, according to the tagging rate dependencies shown in Fig. 6.11 and in order to minimize the number of low statistics or undefined matrix bins, is the one that was already successful in previous analyses:

- 3 bins in jet E_T: $[15,\ 40)$; $[40,\ 70)$; ≥ 70 GeV;
- 11 bins in jet N_{trk}: from $N_{trk} = 2$ to $N_{trk} \geq 12$;
- 10 bins in $\not\!\!E_T{}^{prj}$: < -40; $[-40,\ -30)$; $[-30,\ -20)$; $[-20,\ -10)$; $[-10, 0)$; $[0,\ 10)$; $[10,\ 20)$; $[20,\ 30)$; $[30,\ 40)$; and ≥ 40 GeV.

Each jet contained in the 3-jet events data sample will be classified according to the matrix bin it belongs to, in terms of the corresponding jet variables E_T, N_{trk} and $\not\!\!E_T{}^{prj}$. After the classification, for each matrix bin (x, y, z), with x, y, z integers in the range allowed by the chosen matrix binning, the total number of positive b-tagged jets $N_{jets}^{+}(x, y, z)$ and the total number of taggable jets $N_{jets}^{taggable}(x, y, z)$ falling in the (x, y, z) matrix bin will be used to calculate the following tagging rate

$$\mathcal{R}(x, y, z) = \frac{N_{jets}^{+}(x, y, z)}{N_{jets}^{taggable}(x, y, z)} \tag{6.11}$$

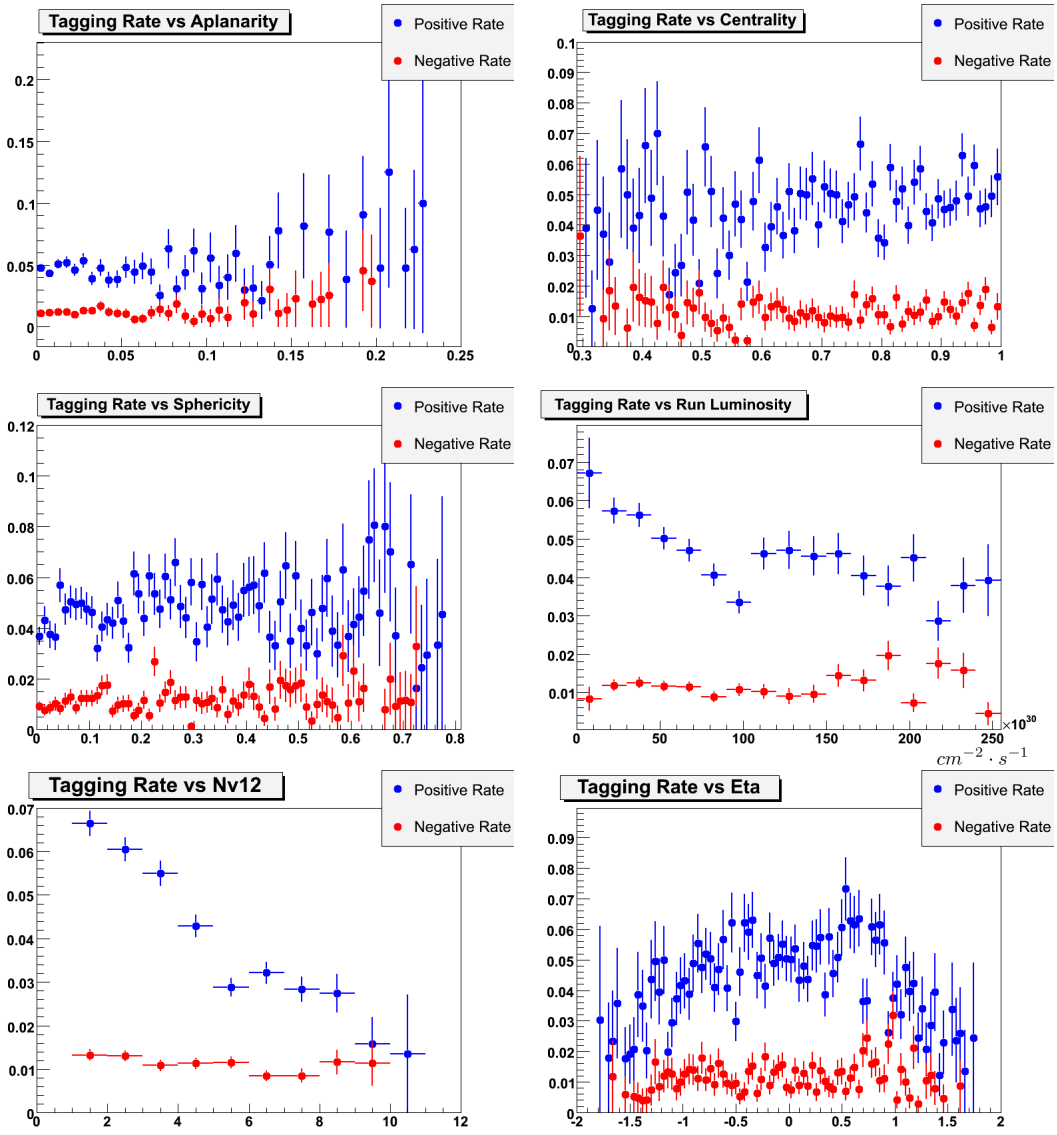


Figure 6.13: Positive and negative b-tagging rates as a function of (from top to bottom, from left to right): Aplanarity, Centrality and Sphericity, luminosity, N_{v12}, and jet η for 3-jet events.

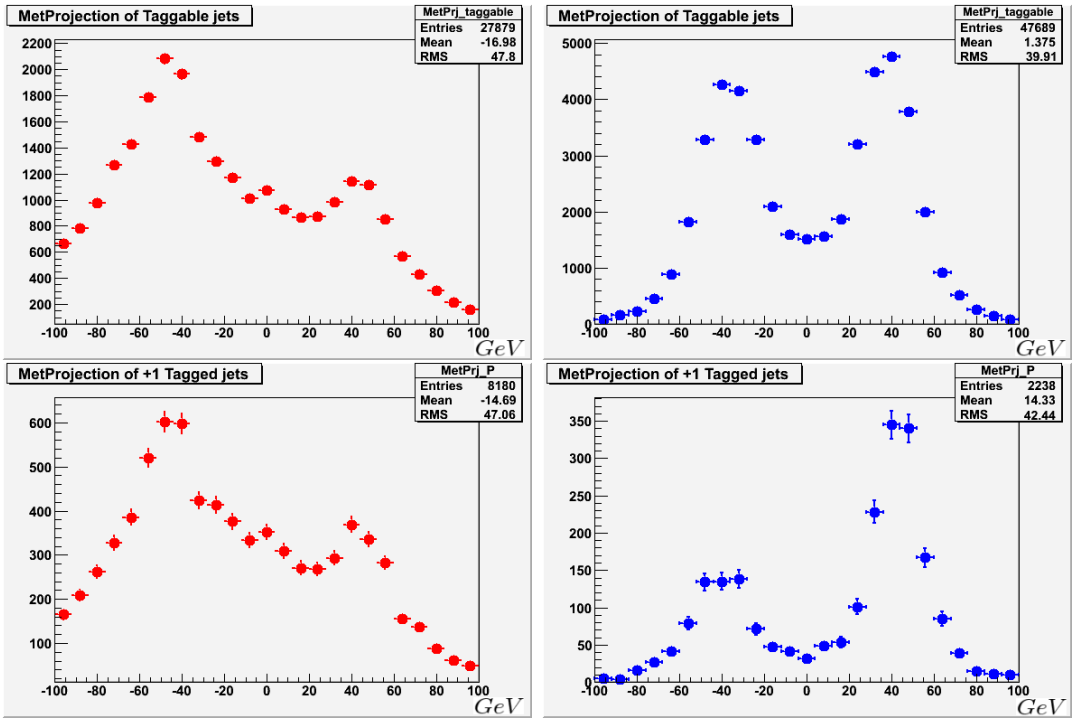


Figure 6.14: $\not\!\!E_T{}^{prj}$ distribution for inclusive Monte Carlo $t\bar{t}$ and data 3-jet events. Top row: left (right) $\not\!\!E_T{}^{prj}$ plot for taggable jets in $t\bar{t}$ (data). Second row: missing transverse energy projection for positive tagged jets for both $t\bar{t}$ (left) and data (right).

This allows us to associate to each $k-th$ jet in an event a $3-d$ b-tagging probability:

$$\mathcal{P}(E_T^k, N_{trk}^k, \not\!\!E_{T_{prj}}{}^k) = \mathcal{R}(x, y, z) \tag{6.12}$$

by finding the (x, y, z) matrix bin corresponding to the $(E_T^k, N_{trk}^k, \not\!\!E_{T_{prj}}{}^k)$ triplet of jet variables.

This per-jet probability will allow to calculate the number of background b-tags expected in a given data sample as follows: the number of expected background b-tags in the $i-th$ event in a given sample, is defined as:

$$N_{tags}^i = \sum_{k=1}^{n} \mathcal{P}(E_T^k, N_{trk}^k, \not\!\!E_{T_{prj}}{}^k) \tag{6.13}$$

where the sum on k is over all taggable jets in the event. The total number of tagged jets expected for a given data sample will then be the sum of the expected tags per each event.

In the next section we will check if this choice of parameterization and binning is satisfactory.

6.9.3 b-tagging matrix checks

Before applying the parameterization we found previously to estimate the number of background b-tagged jets in a given data sample, we first want to check that

it can predict the right kinematical distributions for b-tagged events in samples of data before any selection, where the $t\bar{t}$ signal contamination is quite small.

Kinematical distributions of matrix-predicted background

Once we chose our parameterization variables and built the tagging matrix, we can use the matrix definition in order to construct kinematical distributions and compare them with the observed data distributions for events with $N_{jet}(E_T^{L5} \geq 15\ GeV, |\eta| \leq 2.0) \geq 3$ and at least one b-tagged jet before any other kinematical requirements except the clean-up prerequisites selection.

The matrix-predicted kinematical distributions are obtained by weighting each jet according to its parameterized tagging probability.

Fig. 6.15 shows the observed and matrix-predicted distribution for kinematical variables such as jet E_T, N_{trk}, $\not\!\!E_T^{prj}$, η, ϕ, then global event variables Aplanarity, Centrality and Sphericity.

Fig. 6.16 shows the observed and matrix-predicted distribution for another set of kinematical variables such as $\not\!\!E_T$, $\not\!\!E_T^{sig}$, $\sum E_T$, $\sum E_T^3$, $DPhiMin$, the number of good quality vertices N_{v12}, luminosity and event run.

The insets at the bottom of each panel display the bin-by-bin ratio of observed to matrix-calculated distributions. In general, the observed to expected ratio is almost flat for all the variables here considered. Exceptions are for example the jet E_T and jet η spectra. For jet E_T the ratio shows some structure at low E_T, in the range $15 \div 40\ GeV$, where the b-tagging rate is parameterized with a single matrix bin. On the other hand, the jet η ratio presents some structure over all the η range, mainly due to a residual η dependence left by the jet N_{trk} b-tagging rate parameterization. Generally the ratio between observed and expected distributions behaves well, confirming the effectiveness of the tagging matrix in describing the kinematical distribution of tagged data.

b-tagging rate extrapolation at high jet multiplicities

Another important check consists in extrapolating the b-tagging rate dependencies at jet multiplicities higher than 3, where the matrix is parameterized, and compare the b-tags prediction from tagging matrix application to data to the observed number of b-tagged jets. This extrapolation is performed on the complete data sample obtained after the application of the prerequisites but before any additional kinematical requirement. As already discussed, our data sample has a sizeable content of $t\bar{t}$ events in jet multiplicities higher than 3: we thus expect the matrix predictions to be sistematically underestimating the number of observed tags in the sample.

Additionally, we have to take into account another problem: since the data sample before the tagging requirement is expected to contain a non-negligible $t\bar{t}$ component, the tagging rate parameterization procedure overestimates the background. In fact the expected number of b-tags provided by the positive tagging matrix parameterization does not refer only to background events, since it receives a contribution from $t\bar{t}$ events in the pre-tagging sample.

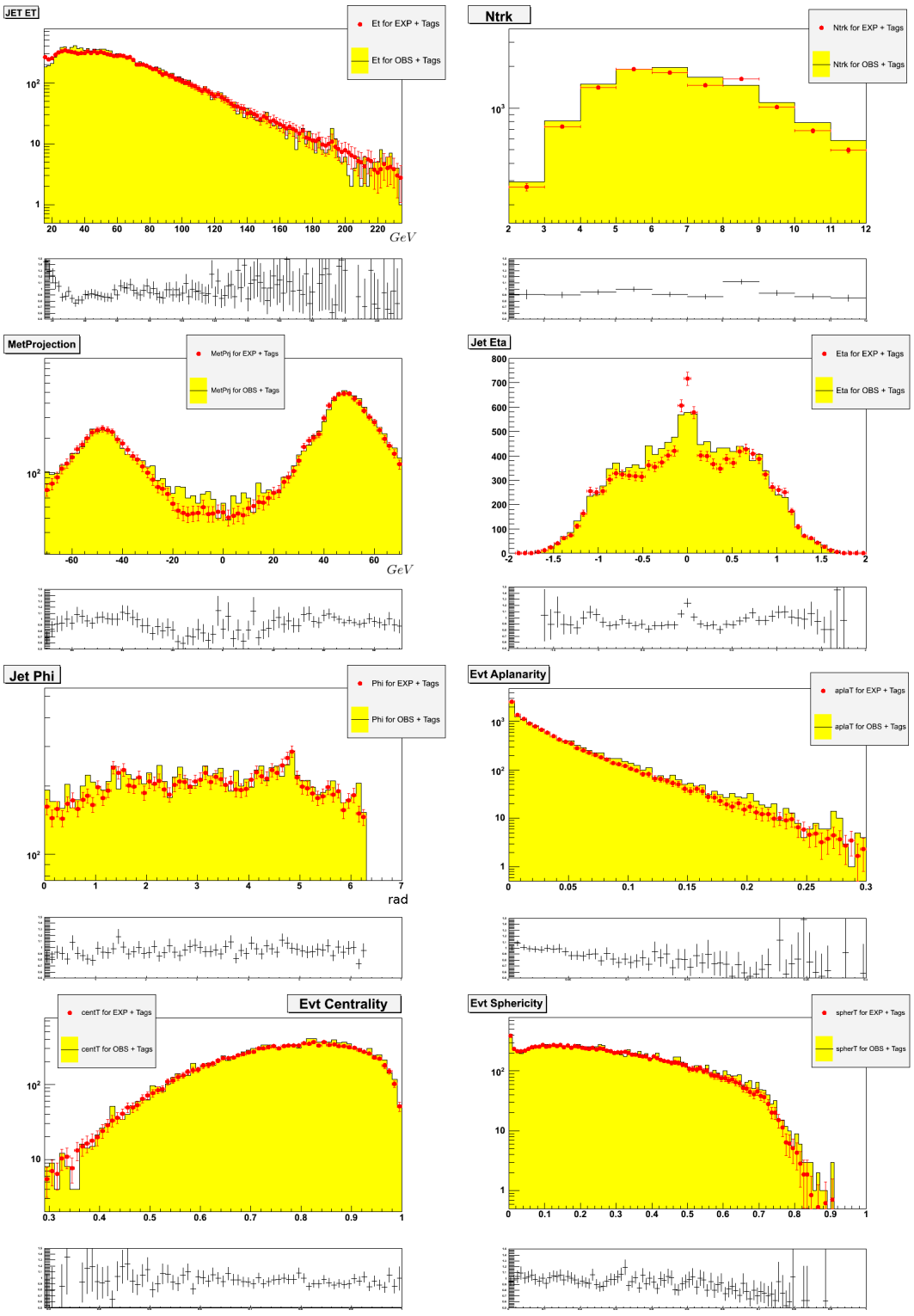


Figure 6.15: Checks of tagging matrix-based variables distributions in data events with at least three $E_T^{L5} \geq 15\,GeV$ and $|\eta| \leq 2.0$ jets. From top to bottom, from left to right: Jet E_T, N_{trk}, $\not\!\!E_T^{prj}$ η, ϕ; then global event variables Aplanarity, Centrality and Sphericity. All plots except the one for η are in log scale. The insets at the bottom of each panel display the bin-by-bin ratio of observed to matrix-calculated distributions.

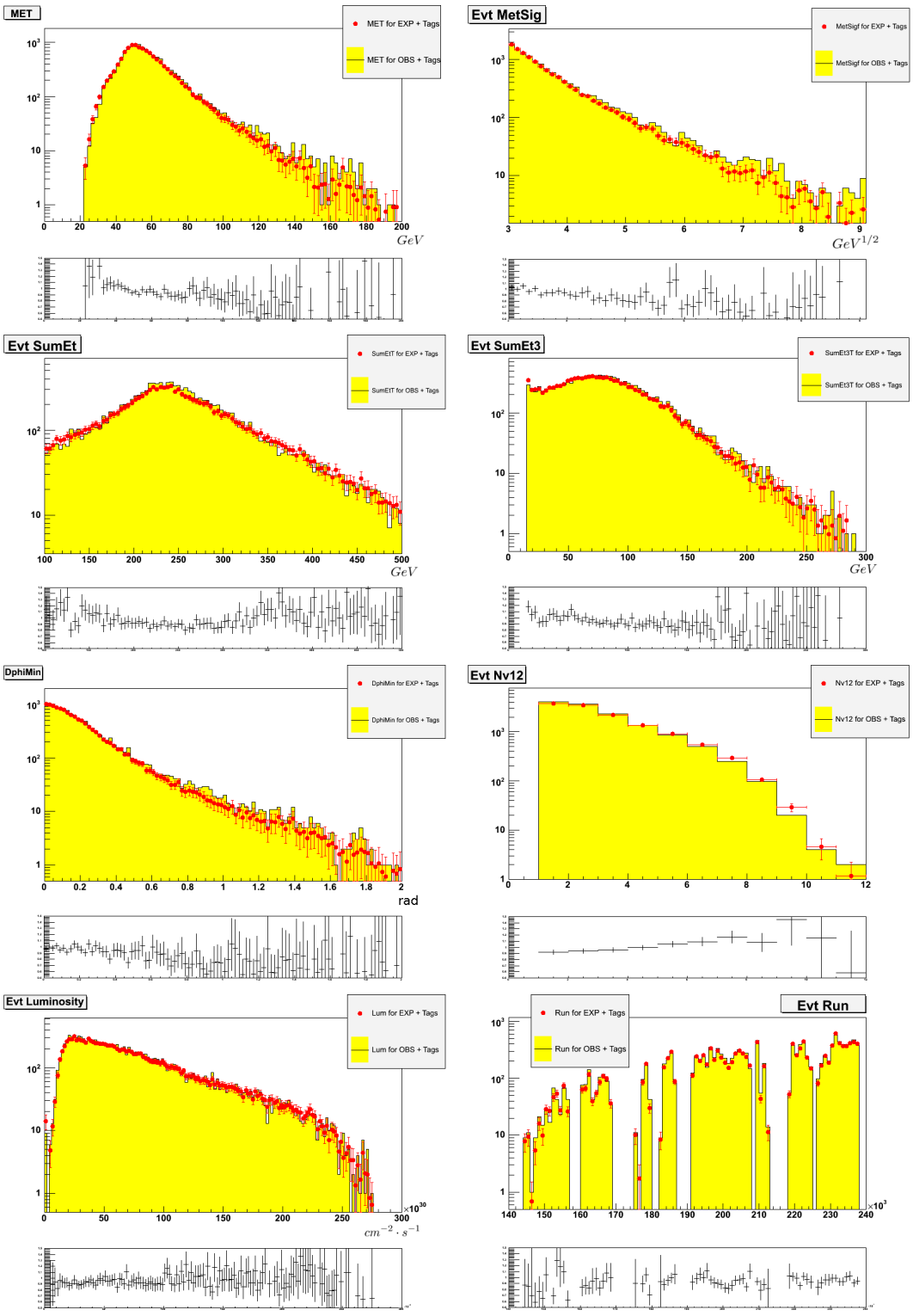


Figure 6.16: Checks of tagging matrix based event variables distributions in data events with at least three $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ jets. From top to bottom, from left to right: $\not\!\!E_T$, $\not\!\!E_T^{sig}$, $\sum E_T$, $\sum E_T^3$, $DPhiMin$, the number of good quality vertices N_{v12}, luminosity and event run. All plots are in log scale. The insets at the bottom of each panel display the bin-by-bin ratio of observed to matrix-calculated distributions.

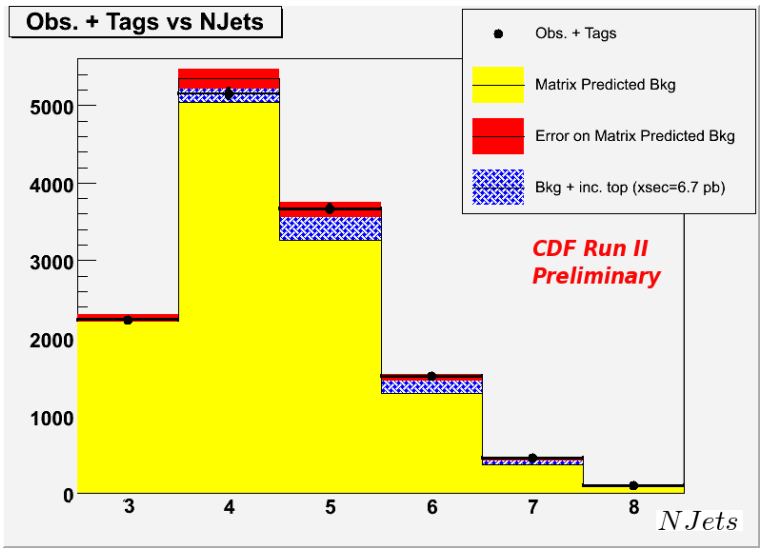


Figure 6.17: Tagging matrix check after prerequisites application and before any kinematical selection. Observed and predicted positive b-tags as a function of the jet multiplicity. The expected contribution coming from $t\bar{t}$ events is also shown, see text for details.

The consequence of this is that we need to remove the $t\bar{t}$ contribution in each jet bin in order to have a background-only determination of the number of expected b-tags. To do so, we iteratively correct the number of expected b-tags for each jet multiplicity as follows [22]:

$$N'_{exp} = N^{fix}_{exp}\frac{N_{evt} - N^{t\bar{t}}_{evt}}{N_{evt}} = N^{fix}_{exp}\frac{N_{evt} - \frac{N_{obs}-N_{exp}}{\epsilon^{ave}_{tag}}}{N_{evt}} \tag{6.14}$$

where, for a chosen jet multiplicity:

- N^{fix}_{exp} is the number of expected tags for that jet multiplicity coming from the tag rate parameterization before any correction; this number is fixed during the iterative procedure.
- N_{evt} is the number of events in the pre-tagging data sample of that jet multiplicity used to determine N^{fix}_{exp} through the tag matrix prediction;
- ϵ^{ave}_{tag} is the average tagging efficiency, defined as the Monte Carlo ratio between the number of positive b-tagged jets and the number of events in the pre-tag sample in the chosen jet multiplicity;
- $N^{t\bar{t}}_{evt}$ is the $t\bar{t}$ contamination in the pre-tagging sample of the chosen jet multiplicity, estimated as $\frac{N_{obs}-N_{exp}}{\epsilon^{ave}_{tag}}$.

The iterative procedure stops when the difference $|N'_{exp} - N_{exp}| \leq 1\%$.

The results of this approach are shown in Fig. 6.17, where we assumed a $t\bar{t}$ production cross section $\sigma_{t\bar{t}} = 6.7\ pb$ for the Monte Carlo. The red error bands in the plot are statistical only and come from the tag matrix application: we recall that for each matrix bin, the tag rate is calculated as $N^{+}_{bin}/N^{taggable}_{bin}$ with N^{+}_{bin}

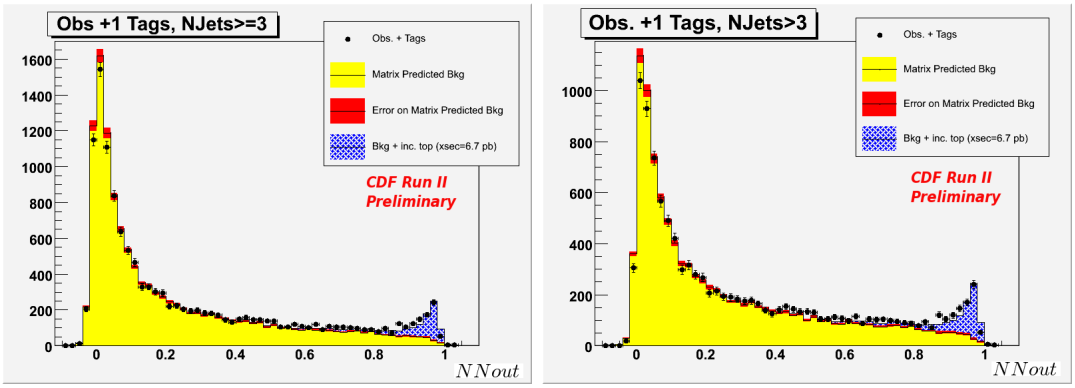


Figure 6.18: Tagging matrix check after prerequisites application and before any kinematical selection. Observed and predicted positive b-tags as a function of Neural Network output. Left plot shows the predictions for data events with at least three tight jets, right plot for at least four tight jets. The expected contribution coming from $t\bar{t}$ events is also shown, see text for details.

being the number of positive tagged jets and $N_{bin}^{taggable}$ the number of taggable jets in that matrix bin in the 3-jets sample used for matrix parameterization. We thus propagate the error associated with this ratio to the expected number of tags.

Once we take into account the $t\bar{t}$ signal contamination in the sample and its contribution to the number of observed b-tags, the agreement between the number of observed and predicted b-tags is good in all the jet multiplicity bins, being exactly the same by definition for 3-jet events, on which the matrix is calculated.

b-tagging rate extrapolation and Neural Network

An additional check we want to perform is related to the behaviour of the matrix predictions with respect to the output of the Neural Network we will use later for our kinematical selection; we want to verify that the prediction of the background works well over all the spectrum of the output of the neural network. Fig. 6.18 shows the output of the Neural Network and the corresponding background prediction from the tag matrix and the expected contribution from $t\bar{t}$ signal both for events with at least three tight jets and at least four tight jets. Matrix predicted tags for bins with a considerable amount of signal contamination have been corrected according to the iterative procedure described in Sec. 6.9.3.

Results are quite good over all the neural network spectrum, altough some discrepancies arise mainly in the low output region. In the high neural network output region we can see that the tagging matrix predictions are not sufficient to justify the number of observed tags, while the agreement is good if we add the amount of tags coming from the expected $t\bar{t}$ signal contribution. This is both a confirmation of the effectiveness of the method we used to estimate the background and an additional check of the correct behaviour of the neural network we trained.

Fig. 6.19 shows the same kind of plot for 3, 4 and 5 tight jet events. As expected, agreement is very good in the 3 jets sample and this provides an additional check of the fact that the matrix parameterization is not affected by the application of the neural network. Furthermore, since we don't expect a sizeable signal presence

in the sample, the fact that the vast majority of 3 jet events has a neural network output close to zero is again an indication of a well trained network. The behaviour of the matrix predictions in the 4 and 5 jets data samples is less accurate and more affected by statistical fluctuations mainly due to the fact that bins are low populated, but further highlights the desired features of the network, that classifies as expected the signal events.

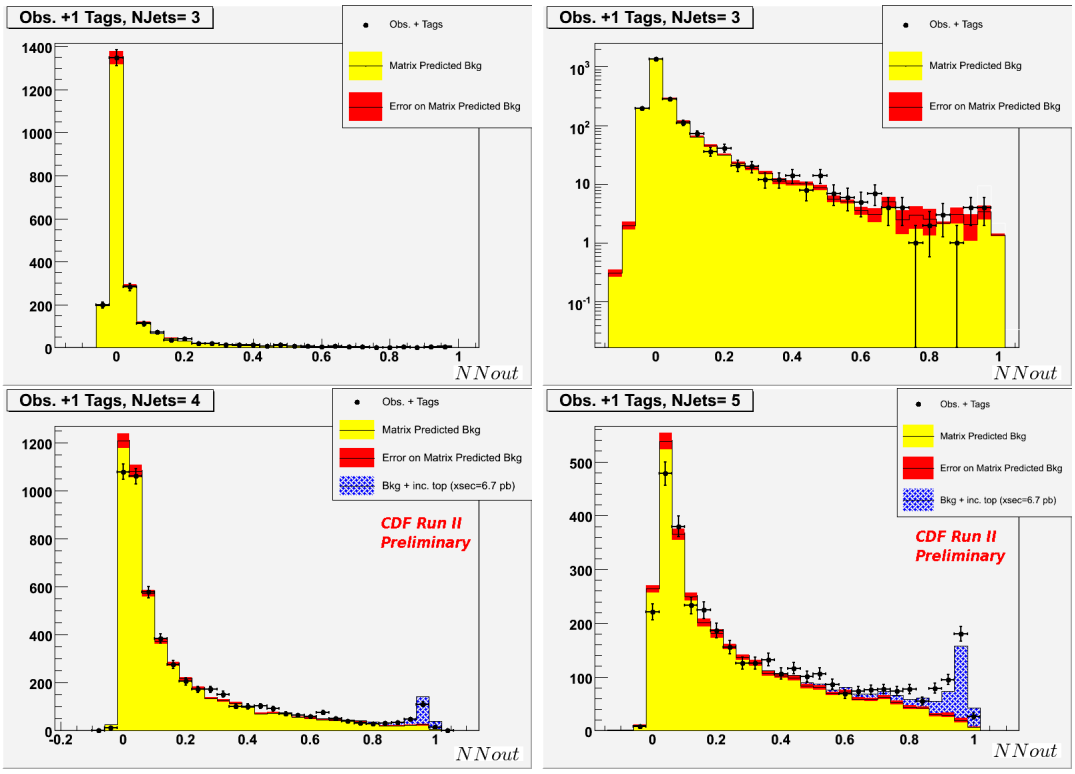


Figure 6.19: Tagging matrix check after prerequisites application and before any kinematical selection. Observed and predicted positive b-tags as a function of Neural Network output. Upper left plot shows the predictions for data events with exactly three tight jets, upper right plot is in log scale. Bottom left plot shows the predictions for data events with exactly four tight jets, bottom right plot for data events with exactly five tight jets. The expected contribution coming from $t\bar{t}$ events is also shown, see text for details.

6.10 Event selection

In this section the optimization procedure we adopted to set the best values of the neural network output cut will be described.

As previously described, b-jet identification provided by the SECVTX algorithm constitutes an effective handle to discriminate the $t\bar{t}$ production against background processes. The final data sample to be used for a cross section measurement will thus be obtained by applying, in addition to the neural network selection, the requirement of at least one positive SECVTX tagged jet.

The optimization procedure we seek is aimed at minimizing the statistical uncertainty on the cross section measurement, in order to optimize the measurement in terms of the expected number of b-tags over the background prediction. The former quantity is evaluated from inclusive Monte Carlo $t\bar{t}$ sample, the latter is derived from the b-tagging matrix application to data.

6.10.1 Optimization and Best Cut

After the clean up cuts described in Sec. 6.6, the analysis selection starts by asking for multijet data events with at least four jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$: 3-jet data events are not considered in the optimization procedure since they are used for the b-tagging rates parameterization and thus are intrinsically biased.

The optimization procedure for the event selection definition is performed after the $N_{jets} \geq 4$ requirement and scans different cuts on neural network output; among all possible cut configurations it chooses the one promising the minimum relative statistical error on the cross section measurement.

The central value of the production cross section we want to measure is given by:

$$\sigma(p\bar{p} \to t\bar{t}) \times BR(t\bar{t} \to \not\!\!E_T + jets) = \frac{N_{obs} - N_{exp}}{\epsilon_{kin} \cdot \epsilon_{tag}^{ave} \cdot L} \tag{6.15}$$

where N_{obs} and N_{exp} are the number of observed and matrix-predicted tagged jets in the selected sample, respectively; ϵ_{kin} is the trigger, prerequisites and neural network selection efficiency measured on inclusive Monte Carlo $t\bar{t}$ events; ϵ_{tag}^{ave}, defined as the ratio of the number of positive tagged jets to the number of pre-tagging events in the inclusive $t\bar{t}$ Monte Carlo sample, gives the average number of b-tags per $t\bar{t}$ event. Finally, L is the luminosity of the dataset used.

Using in input to Equation 6.15 the measured kinematical efficiency, the average number of b-tags per $t\bar{t}$ event, the actual integrated luminosity and the number of b-tagged jet expected from the tag rate parameterization in the selected sample, we can estimate the expected cross section value and its relative statistical uncertainty for each neural network cut. The only missing piece is N_{obs}. We cannot use the actual number of observed b-tags in the selected data, since it would bias our conclusion given its possible statistical fluctuations. For this reason, in order to obtain an *a priori* determination of the best cut, we substitute N_{obs} with the expression $N_{exp} + N_{MC}$, where N_{exp} and N_{MC} are the number of expected b-tagged jets from the tagging rate application and from inclusive $t\bar{t}$ Monte Carlo samples

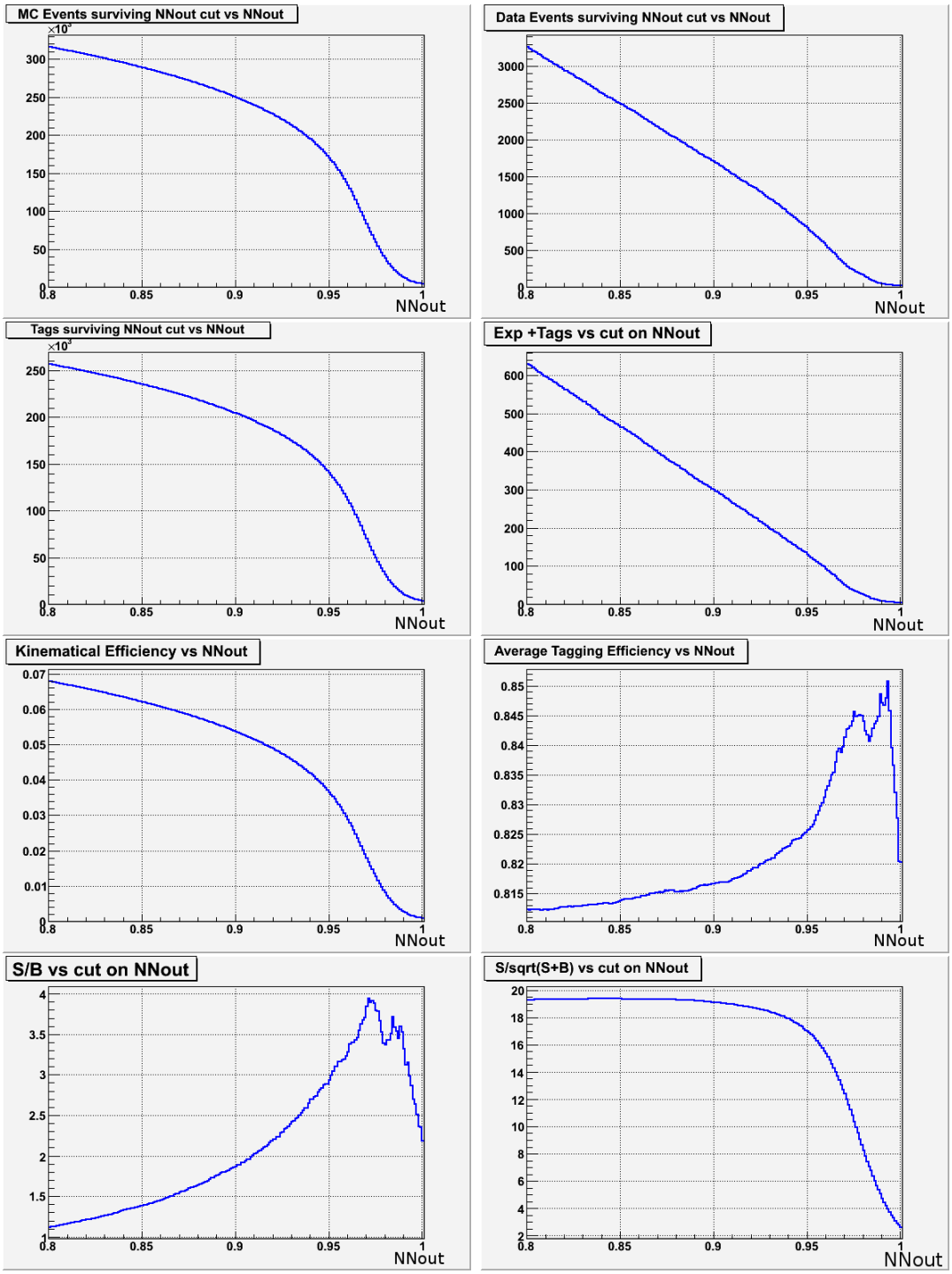


Figure 6.20: From top to bottom, from left to right: MC^{evt} and $Data^{evt}$, N_{MC} and N_{exp}, ϵ_{kin} and ϵ_{tag}^{ave}, S/B ratio and signal statistical significance $S/\sqrt{S+B}$ as a function of the cut on the Neural Network output. MC^{evt} and N_{MC} have not been rescaled to their expectation value in 1.9 fb^{-1}. S/B and ϵ_{tag}^{ave} plots show effects due to low statistics for cuts on neural network output in the region close to 1.

after the application of the given neural network cut, respectively. Using these values, the statistical uncertainty affecting the measurement can be computed before

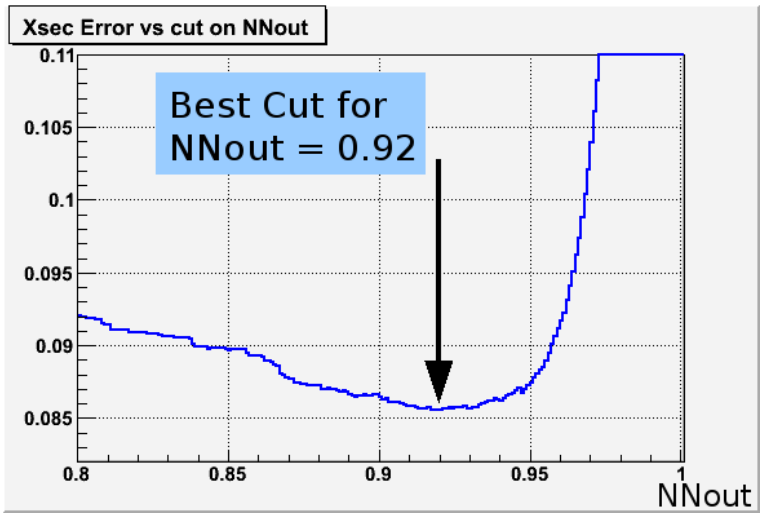


Figure 6.21: Expected relative statistical error on the cross section measurement $\sigma_{xsec}/xsec$ versus neural network cut. Best cut is set as $NNout \geq 0.92$.

looking at the "post-tagging" data sample, allowing in this way to choose the cut minimizing the relative error on the cross section measurement.

For each neural network output cut the following quantities are calculated:

- MC^{evt}, and $Data^{evt}$: number of inclusive Monte Carlo $t\bar{t}$ and data events in the selected sample, before any b-jet identification requirement.

- N_{MC} and N_{exp}: number of positive tags expected from Monte Carlo inclusive events and from tagging rate parameterization after the kinematical selection defined by the cut on the neural network output. Since we want to derive a "blind" minimization procedure, we don't want to look at the post-tagging sample, meaning we won't use any information on the number of observed b-tags N_{obs} in the sample obtained after the neural network cut. Since N_{obs} is necessary for our iterative correction procedure, we will then rely on the uncorrected matrix predictions only.

- ϵ_{kin} and ϵ_{tag}^{ave} are derived from the application of the cut to the Monte Carlo sample.

- the signal statistical significance obtained as $S/\sqrt{S+B}$: ratio of the number of tags expected for $t\bar{t}$ events and the square root of the number of tags expected from background processes plus the number of tags expected from signal.

- $\sigma_{xsec}/xsec$: relative error on the cross section measurement.

Results are reported in Fig. 6.20, while Fig. 6.21 shows the cross section uncertainty versus the neural network output cut, calculated using only the statistical errors of the involved quantities.

The final result of this procedure sets as the best event selection cut $NNout \geq 0.92$, promising a relative statistical cross section uncertainty of 8.6% and a S/B

ratio in terms of positive tags due to signal versus tags coming from background processes of 2.18. The pre-tagging combined kinematical efficiency of trigger, event clean-up, and neural network cut on $t\bar{t}$ inclusive events is measured to be $\epsilon_{kin} = 4.907 \pm 0.001\%$, where the uncertainty is statistical only. The average number of tags per $t\bar{t}$ event under these selections is found to be $\epsilon_{tag}^{ave} = 0.8188 \pm 0.0008$, and is determined by dividing the number of b-tagged jets in the kinetically selected sample ($N_{tag} = 187,201$) by the number of inclusive $t\bar{t}$ events surviving the selection ($N_{evt} = 228,614$). The ϵ_{kin} and ϵ_{tag}^{ave} values will be used for the cross section measurement as it will be described in Chapter 7.

Previous kinematical selection and new trigger effect

MC inc. $t\bar{t}$	ttopel	$\epsilon_{cut}(\%)$	ttop75 ot	$\epsilon_{cut}(\%)$	ttop75 nt	$\epsilon_{cut}(\%)$
Tot. evts			4719385	–	4719385	–
Good Run	1021924	–	3979960	–	3979960	–
Trigger			2579315	64.81	2311270	58.07
Vertex Req.			2466709	95.63	2213324	95.76
Lepton Veto	558528	–	2196636	89.05	1982264	89.56
$N_{Jet} \geq 4$	549138	98,32	2156643	98.18	1958949	98.82
$\not\!\!E_T{}^{sig} \geq 4\ GeV^{1/2}$	78145	14.23	309425	14.35	238426	12.17
$DPhiMin \geq 0.4$	49848	63.79	197264	63.75	145197	60.90
Tot. Eff.		4.88		4.96		3.65

Table 6.5: Effect of the introduction of the new L2 TOP_MULTI_JET trigger on inclusive $t\bar{t}$ Monte Carlo events. The column for *ttopel* shows results taken from [3]; the one for *ttop75 ot* shows the results obtained on ttop75, the Monte Carlo dataset used in our neural network analysis, after a full simulation of the old trigger. Column for *ttop75 nt* show the results of the selection on ttop75 with a full simulation of the new trigger path. For each cut, the efficiency with respect to the previous selection is provided. Last line shows the overall efficiency of the selection. The expected loss in the efficiency of the selection on inclusive $t\bar{t}$ signal due to the introduction of the new trigger is $\sim 26.4\%$.

The kinematical selection studied in [3] for the cross section measurement in the $t\bar{t} \rightarrow \not\!\!E_T + jets$ channel with the 311 pb^{-1} data sample was optimized for the old TOP_MULTI_JET trigger path and reached a kinematical efficiency on inclusive $t\bar{t}$ events of $\epsilon_{kin} = 4.878 \pm 0.021\%$.

Tab. 6.5 shows the effect caused on this kinematical selection by the introduction of the higher energy treshold in the new Level 2 requirements of the TOP_MULTI_JET trigger, starting from Run 194328 (see 6.2 for details). Results in the table are obtained with a Good Run List older than the one used in our analysis.

First two columns, referring to the dataset *ttopel*, show results of the selection on inclusive $t\bar{t}$ Monte Carlo events, taken from the published article [3]. For comparison, the results of the old selection on the inclusive Monte Carlo dataset used in

this analysis, *ttop75*, treated with a full simulation of the old TOP_MULTI_JET trigger path are reported in the two central columns ("ttop75 ot"). Relative efficiencies of the different selections agree very well, and thus give us the possibility of estimating the effect of the new trigger introduction by simulating its requirements on *ttop75*. The effects of the new trigger and the kinematical selection are shown in the last two columns of Tab. 6.5 under the label "ttop75 nt". Relative efficiencies of the analysis kinematical cuts are almost left unchanged by the introduction of the new trigger, but we notice an overall reduction of the kinematical efficiency on the sample with the new trigger simulation from $4.96 \pm 0.01\%$ to $3.65 \pm 0.01\%$, causing an expected $t\bar{t}$ signal loss around the 26.4% with respect to the previous trigger.

It is interesting to note that the neural network selection described in this work achieves an efficiency $\epsilon_{kin} = 4.907 \pm 0.001\%$ on $t\bar{t}$ inclusive events taken with the new trigger path, comparable with the one obtained in [3] for the old trigger. In conclusion, with the introduction of the neural network selection we were then able to mitigate the effect of the new trigger on the signal acceptance of the old kinematical selection in this decay channel.

Event selection acceptances

The impact of the trigger, prerequisites and optimized neural network selection on exclusive $e + jets$, $\mu + jets$, $\tau + jets$, all-hadronic and di-lepton $t\bar{t}$ Monte Carlo events is shown in Tab. 6.6. We note that, as expected, the di-lepton decay channel is highly suppressed by our choice of TOP_MULTI_JET trigger, that is not designed for this kind of analysis. Moreover, the all-hadronic $t\bar{t}$ decay channel is highly suppressed by the requirement of large missing E_T significance $\not{E}_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$: as we already pointed out, this analysis prerequisite rejects those events whose missing E_T is due mainly to residual energy mis-measurement effects, such as all-hadronic events, while it focuses on events containing physics-induced $\not{E}_T$.

Considering the leptonic decay channels, we note that the selection we described allows $t\bar{t}$ isolation by requiring the presence of large $\not{E}_T^{sig}$ in the event, thus searching for high-P_T neutrino signature produced in the leptonic decay of the W boson. This signature is produced in a similar way for all the top pair $e + jets$, $\mu + jets$ and $\tau + jets$ decay channels.

Even if we didn't use any lepton identification procedure and despite the well-identified high-P_T lepton veto imposed for e and μ, the final acceptance provided by the kinematical selection is comparable for all the lepton+jets decays. The efficiencies calculated with respect to the number of events for each decay mode of all selection requirements are found to be $9.44 \pm 0.03\%$, $7.38 \pm 0.03\%$ and $12.66 \pm 0.04\%$ for $e + jets$, $\mu + jets$ and $\tau + jets$ channels respectively.

This is due to the fact that the trigger requirement has a large impact on μ+$jets$ $t\bar{t}$ decays with respect to the other leptonic ones (the muons does not make any jet), while the tight lepton veto prerequisite decreases the selection efficiency for both e and μ plus jets events. Additionally, the relative efficiency for the $\not{E}_T^{sig}$ selection with respect to the N_{jet} requirement is more or less the same for τ and electron plus jets events, but it is higher for muon: we correct the $\not{E}_T$ for muon transverse momentum but we do not include the muon P_T in the calculation of the

N evt $t\bar{t}$	$e+jets$	$\epsilon_{cut}(\%)$	$\mu+jets$	$\epsilon_{cut}(\%)$	$\tau+jets$	$\epsilon_{cut}(\%)$
Branching Ratio	689,083	–	688,787	–	689,743	–
Good Run	680,317	–	679,962	–	680,758	–
Trigger	411,734	60.52	201,436	29.62	315,536	46.35
Vertex Req.	393,562	95,59	193,513	96.07	302,867	95.98
Lepton Veto	220,834	56.11	128,906	66.61	285,637	94.31
$N_{Jet} \geq 4$	215,022	97.37	124,367	96.48	280,032	98.04
$\not\!\!E_T{}^{sig} \geq 3\ GeV^{1/2}$	117,360	54.58	82,472	66.31	160,683	57.38
$NNout \geq 0.92$	65,072	55.45	50,819	61.62	87,352	54.36
Tot. Eff wrt BR		9.44		7.38		12.66

N evt $t\bar{t}$	all-hadronic	$\epsilon_{cut}(\%)$	di-lepton	$\epsilon_{cut}(\%)$
Branching Ratio	2,154,096	–	221,659	–
Good Run	2,126,366	–	218,737	–
Trigger	1,744,325	82.03	18,555	8.48
Vertex Req.	1,673,326	95.93	17,172	92.55
Lepton Veto	1,672,094	99.93	7,853	45.73
$N_{Jet} \geq 4$	1,662,229	99.41	7,089	90.27
$\not\!\!E_T{}^{sig} \geq 3\ GeV^{1/2}$	74,833	4.50	4,890	68.98
$NNout \geq 0.92$	10,432	13.94	2,851	58.30
Tot. Eff wrt BR		0.48		1.29

Table 6.6: Effect of the trigger, prerequisites and neural network selection cuts for $e/\mu/\tau+jets$ (top) and all-hadronic and di-lepton (bottom) exclusive $t\bar{t}$ Monte Carlo events. For each cut, the efficiency with respect to the previous selection is provided for each $t\bar{t}$ decay channel. Last line shows the overall efficiency of the selection with respect to the branching ratio of the channel.

event $\sum E_T$. This increases the possibility for a $\mu+jets$ event to pass this cut, given that $\sqrt{\sum E_T}$ is the denominator of the missing E_T significance. The same effect happens for the neural network selection cut, that is sensitive to the event $\sum E_T$ and missing E_T significance since it uses these two variables as inputs: the neural network cut shows a higher efficiency on $\mu+jets$ $t\bar{t}$ decays.

Overall, the described event selection provides comparable efficiencies in the pre-tagging sample and thus selects comparable $t\bar{t}$ signal contributions from each lepton+jets decay mode. Rescaling the number of Monte Carlo events surviving the selection according to the 1.9 fb^{-1} data luminosity, we expect about 178/139/239 events from $e/\mu/\tau+jets$ decays, respectively.

Background prediction systematic uncertainty

The optimized neural network cut definition found in previous sections allows us to define control data samples in which to further check the b-tagging rate parameterization for the background. In fact, once the selection is defined, we can reverse its cut to construct control data samples close to the signal region but depleted as much as possible of signal contribution, where to compare the number

N jet	3	4	5	6	7	8
Obs + tags	2230	4977	3361	1329	405	76
Exp + tags	2231.18	4974.27	3231.98	1259.23	359.62	80.26
error	±51.34	±130.05	±93.99	±40.52	±14.22	±4.96
Exp/Obs ratio	≈ 1	≈ 1	0.96	0.95	0.89	1.06

Table 6.7: Tagging matrix check in the data sample with $NNout <= 0.9$. For each jet multiplicity bin, the number of observed and predicted positive b-tags is shown. Uncertainties are statistical only.

Exp + tags, N Jets:	3	4	5	6	7	8
Complete mtx	2238	5049.75	3336.98	1335.07	389	90.77
error	±51.53	±134.08	±99.18	±43.66	±15.52	±5.49
Even mtx	2261.82	5050.97	3369.02	1342.38	392.39	91.09
error	±84.09	±208.51	±156.68	±72.64	±24.09	±7.87
Odd mtx	2184.92	4851.31	3153.2	1258.63	357.03	80.67
error	±86.67	±210.09	±158.72	±69.22	±24.76	±8.69
$\frac{Even-Odd}{Complete}$ ratio (%)	3.44	3.95	6.47	6.27	9.09	11.48

Table 6.8: Even-odd tagging matrix check in the complete data sample after prerequisites. For each jet multiplicity bin, the number of predicted positive b-tags with the complete, even and odd matrix is shown. Uncertainties are statistical only.

of observed positive tags to the number of predicted tags derived from the tagging rate parameterization applied to data. This will allow us to verify the predictions of the matrix and to account for possible deviations from the desired behaviour deriving a systematic uncertainty on the background prediction.

The first sample we want to analyze is made of events with $NNout <= 0.9$. The performances of the tagging matrix as a function of the jet multiplicity are shown in Tab. 6.7. The agreement is quite good and any discrepancy can be limited at the level of few percent.

Additionally, we can try to give an estimate of how our background predictions are affected by statistical fluctuations in the sample we used to construct the matrix parameterization. To do so, we split our 3-jets data sample after prerequisites in *even* and *odd* events and use this two orthogonal samples to build two tag matrixes with the same characteristics of the one we used in the analysis.

The comparison of the tag rates obtained with these two samples and the ones used in the analysis is shown in Fig. 6.22. We now use the matrixes to derive the number of expected positive b-tags for each jet multiplicity in the data sample obtained after prerequisites application, and account any discrepancy among the two and the complete matrix as a systematic error on our background prediction. Tab. 6.8 shows the results of this additional check.

From the studies performed, a systematic uncertainty on the background prediction can be derived.

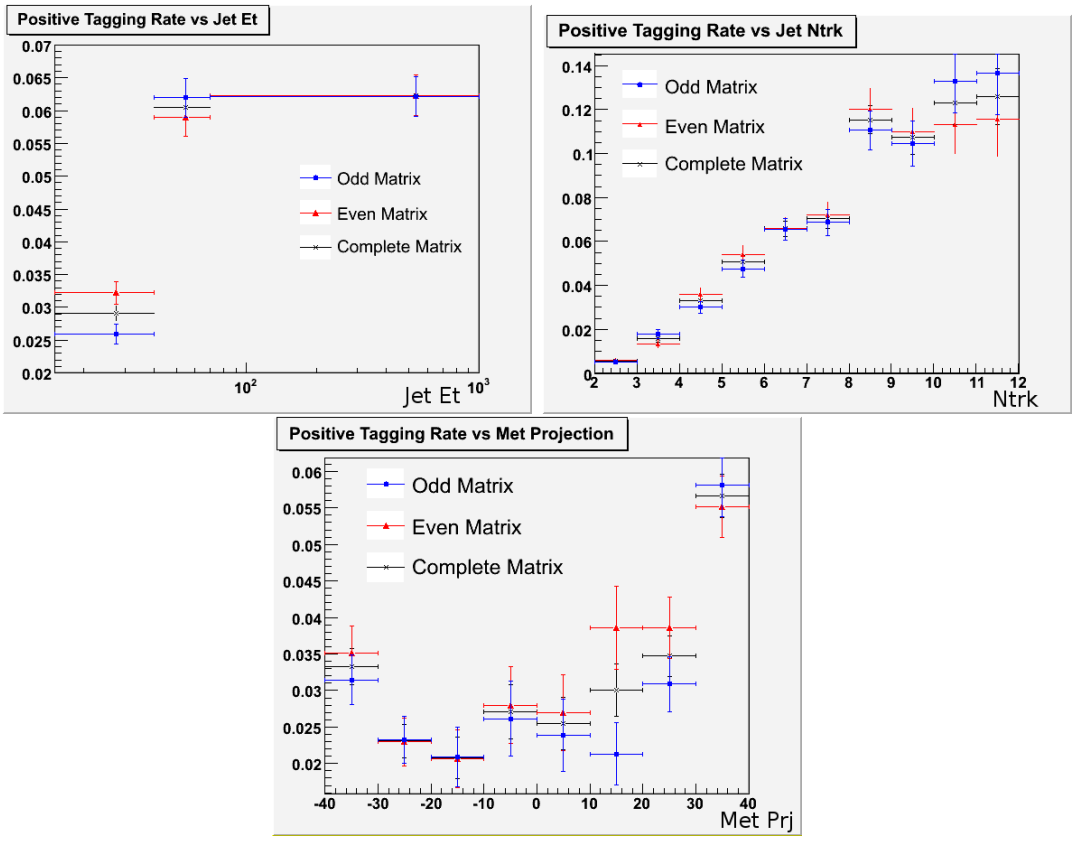


Figure 6.22: Tag Rates for the variables jet E_T, jet N_{trk} and $\not\!\!E_T{}^{prj}$ used in the matrix parameterization are compared for the matrix built using all 3-jets data and for the odd and even matrixes.

Considering the *obs*/*exp* b-tag ratio as a function of the jet multiplicity , the overall discrepancy between observed and matrix predicted number of b-tags due to intrinsic limits of the matrix and to the dependance from the sample in which the matrix has been built can be quoted conservatively at 15%. This value will be assumed as the systematics uncertainty to be associated to our background prediction, and will be used in Chapter 7 for the cross section measurement.

Bibliography

[1] F. Abe *et al.*, *The mu tau and e tau Decays of Top Quark Pairs Produced in p anti-p Collisions at* $\sqrt{s} = 1.8\ TeV$, Phys. Rev. Lett. **79**, 3585 (1997).

[2] G. Cortiana *et al.*, *Measurement of the* $t\bar{t}$ *production cross section in the* τ*+jets channel with SecVtX tags using* $311pb^{-1}$ *of* $p\bar{p}$ *data*, CDF Internal Note 7689.

[3] A. Abulencia *et al.* [CDF Collaboration], *Measurement of the* $t\bar{t}$ *production cross section in* $p\bar{p}$ *collisions at* $\sqrt{s} = 1.96\ TeV$ *using* $\not\!\!E_T$ *+ jets events with secondary vertex b-tagging*, Phys. Rev. Lett. **96**, 202002 (2006).

[4] T. Sjostrand *et al.*, *High-energy-physics event generation with PYTHIA 6.1*, Comput. Phys. Commun. **135** 238 (2001), [arXiv:hep-ph/0010017].

[5] G. Corcella *et al.*, *HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** 010 (2001), [arXiv:hep-ph/0011363].

[6] R. Brun *et al.*, *Geant: Simulation Program For Particle Physics Experiments. User Guide And Reference Manual*, CERN-DD-78-2-REV.

[7] S. Carron *et al.*, *Parametric Model for the Charge Deposition of the CDF RunII Silicon Detectors*, CDF Internal Note 7598.

[8] W. Riegler *et al.*, *COT detector physics simulations*, CDF Internal Note 5050.

[9] S.Y. Jun *et al.*, *GFLASH Tuning in the CDF Calorimeter for the 2003 Winter Release*, CDF Internal Note 7060.

[10] TRIGSIM++ Web Page:
http://www-cdf.fnal.gov/internal/upgrades/daq_trig/twg/trgsim/trgsim.html.

[11] D. Tsybychev *et al.*, *A study of missing* E_T *in Run II minimum bias data*, CDF Internal Note 6112.

[12] G. Cortiana *et al.*, *t tbar -> tau + jets nt5 analysis update*, CDF Internal Note 7553.

[13] S. Grinstein *et al.*, *Electron-Method SecVtx Scale Factor for Winter 2007*, CDF Internal Note 8625.

[14] F. Garberson *et al.*, *SecVtx B-Tag Efficiency Measurement Using Muon Transverse Momentum for 1/fb Analyses* CDF Internal Note 8434.

[15] S. Grinstein and D. Sherman, *SecVtx Scale Factors and Mistag Matrices for the 2007 Summer Conferences*, CDF Internal Note 8910.

[16] F. Garberson *et al.*, *Combination of the SECVTX b-Tagging Scale Factors for 1 fb-1 Analyses*, CDF Internal Note 8435.

[17] V.D. Barger, R.N. Phillips, **Collider Physics**, *Addison-Wesley (1987).*

[18] T. Aaltonen *et al.* [The CDF Collaboration], *Measurement of the Top Pair Production Cross Section in the Lepton+Jets Decay Channel*, CDF Public Note 8795.

[19] T. Aaltonen *et al.* [The CDF Collaboration], *Measurement of the ttbar Cross Section and Top Quark Mass in the All Hadronic Channel*, Phys. Rev. D **76**, 072009 (2007).

[20] Data Quality Web Page:
http://www-cdf.fnal.gov/internal/dqm/goodrun/v17/goodv17.html.

[21] K. Copic and M. Tecchio, *Event vertex studies for dilepton events*, CDF Internal Note 6933.

[22] A. Castro *et al.*, *Top Production Cross Section in the all-hadronic channel*, CDF Internal Note 3464.

[23] G. Cortiana *et al.*, *Background studies for the* $t\bar{t} \rightarrow \tau + jets$ *search*, CDF Internal Note 7292.

Chapter 7

Cross section measurement and systematic uncertainties

In this Chapter we will finalize our $t\bar{t} \rightarrow \not\!\!E_T + jets$ cross section measurement. After describing the final data sample obtained with our neural network selection, we will analyze and discuss the sources of systematic uncertainties and finally we will determine the cross section measurement by means of a likelihood maximization.

7.1 The final sample

The optimized neural network selection described in previous chapter allows to isolate a tagged data sample in which the $t\bar{t} \rightarrow \not\!\!E_T + jets$ signal is estimated to contribute with a signal to background ratio of at least $S/B \sim 2$ after the requirement of selecting events containing at least one positive SECVTX tagged jet. Indeed the tagging probability for a b-jet produced by top quark decay is expected to be higher than the probability to identify b-quark jets yielded by background processes. Additionally, the selection is expected to provide a statistical uncertainty of 8.6% on the cross section measurement. Moreover, we will see at the end of this section that after the correction of the expected background for the presence of signal in the pre-tagging sample used for b-tagging rate parameterization application, the S/B ratio will increase to ~ 4.5, since almost 50% of the expected background b-tags will be found to be due to $t\bar{t}$ contamination.

Fig. 7.1 displays the sample composition after the cut on the neural network: the predicted amount of background b-tags after the selection is shown together with the expected inclusive $t\bar{t}$ contribution; the observed positive tags in the data are shown by dots. After selection we are left with a sample containing 1415 events, and 627 positive b-tags.

As already discussed, since the data events selected before the tagging requirement are expected to contain a non-negligible $t\bar{t}$ component, the tagging rate parameterization procedure overestimates the background. We correct for this effect each jet multiplicity bin of Fig. 7.1 using the iterative procedure discussed in previous chapter.

A good agreement between observed and predicted background tags is noted

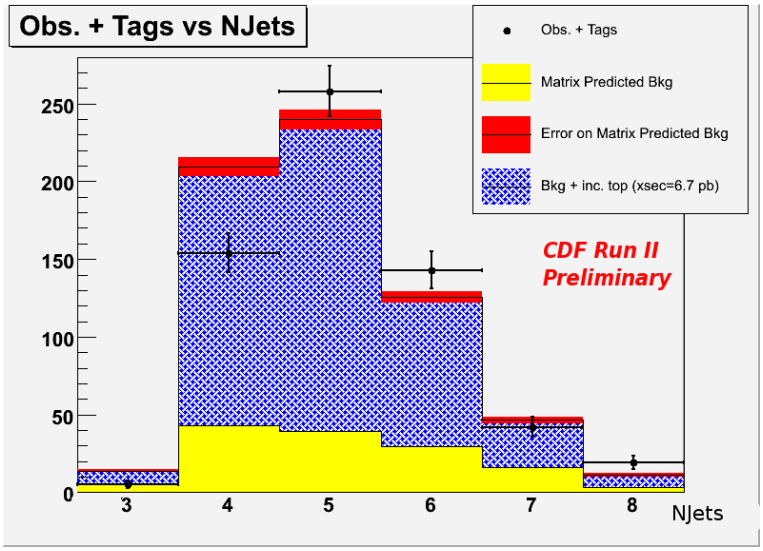


Figure 7.1: Number of tagged jets versus jet multiplicity. Data (points), iteratively corrected background (yellow histogram) and $t\bar{t}$ expectation (blue histogram) for $\sigma_{t\bar{t}} = 6.7\ pb$ are shown after neural network selection.

in the 3-jet bin, where the tagging matrix is computed before the kinematical selection, while on the contrary for 4 to 8 jet bins the addition of Monte Carlo inclusive contribution is required in order to explain the data behavior. However, the 4-jets bin shows a noticeable mismatch between the number of observed and predicted b-tags: this effect is still under investigation.

Now that we estimated the $t\bar{t}$ component in the selected sample and the overall background to the signal signature by means of the tagging rate parameterization applied to data, we can finally provide the top pairs production cross section measurement in the selected data sample.

We have all the ingredients to proceed directly for a measurement, we are only missing the systematic uncertainties determination. The measurement we will describe uses the excess of b-tagged jets over the background prediction to estimate the top pairs production cross section. In order to properly account for each systematic source affecting the measurement, a likelihood function will be used to determine the cross section value.

The cross section measurement will be obtained by maximizing $\log \mathcal{L}$, where the likelihood function is defined as follows:

$$\mathcal{L} = e^{-\frac{(L-\bar{L})^2}{2\sigma_L^2}} \cdot e^{-\frac{(\epsilon_{kin}-\bar{\epsilon}_{kin})^2}{2\sigma^2_{\bar{\epsilon}_{kin}}}} \cdot e^{-\frac{(\epsilon^{ave}_{tag}-\bar{\epsilon}^{ave}_{tag})^2}{2\sigma^2_{\epsilon^{ave}_{tag}}}} \cdot e^{-\frac{(N_{exp}-\bar{N}_{exp})^2}{2\sigma^2_{N_{exp}}}} \cdot \frac{(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N_{exp})^{N_{obs}}}{N_{obs}!} \cdot e^{-(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N_{exp})} \quad (7.1)$$

where L is the integrated luminosity of the data sample we used, ϵ_{kin} is the combined trigger, prerequisites and neural network selection efficiency on inclusive Monte Carlo $t\bar{t}$ events, and ϵ^{ave}_{tag} is the average number of b-tags per $t\bar{t}$ event. N_{exp} is the number of background b-tags returned by the tagging matrix application to the selected data sample; N_{obs} is the number of observed b tags in the data.

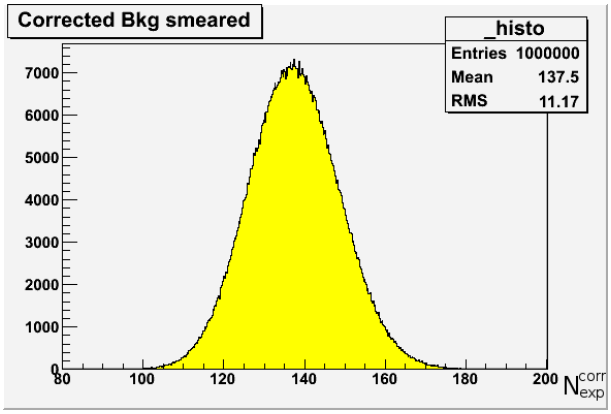


Figure 7.2: Results of background correction, see text for details.

The cross section central value is then given by the likelihood maximization as:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{exp}}{\epsilon_{kin} \cdot \epsilon_{tag}^{ave} \cdot L} \tag{7.2}$$

In the following, we review the input values we need to perform the likelihood maximization.

In previous Chapter we determined the overall kinematical efficiency and the average number of b-tagged jets per $t\bar{t}$ event to be $\epsilon_{kin} = 4.907 \pm 0.001\%$ and $\epsilon_{tag}^{ave} = 0.8188 \pm 0.0008$ respectively. Using the tagging rate parameterization applied to the 1415 events passing the selection, the background amount in terms of b-tagged jets is calculated to be $237.64 \pm 15.28(stat) \pm 35.64(syst) = 237.64 \pm 38.77$, where the first uncertainty is statistical only, while the latter is systematic and is calculated by comparing observed to matrix-predicted b-tags in data control samples and quoting a 15% systematic uncertainty. This value needs to be corrected for the signal presence in the pre-tagging sample, as seen in the previous chapter: the application of our iterative correction procedure yelds a top-free background determination of $N_{exp}^{corr} = 137.5$. The uncertainty on the background correction depends both on the uncertainty on N_{exp} and the uncertainty on ϵ_{tag}^{ave}. In order to evaluate both contributions we follow the tecnique adopted in [1]: we generate $1,000,000$ random samples of N_{exp} events smeared with its ± 15.28 statistical uncertainty and apply the iterative correction using ϵ_{tag}^{ave} smeared with its statistical uncertainty. The resulting N_{exp} distribution is shown in Fig. 7.2 and gives $N_{exp}^{corr} = 137.5 \pm 11.2$, so the relative statistical uncertainty on the expected background becomes 8.1%.

On the other hand, the number of observed b-tagged jets in the data sample selected with the neural network selection is found to be 627.

Finally, the integrated luminosity of the considered data sample is $L = 1906.8 \pm 110.6\ pb^{-1}$.

For a proper determination of the cross section, we need to assign to each of the input values its corresponding uncertainty, accounting for both the statistical and systematic effects.

In the following the sources of systematic uncertainty will be described and quantified.

N evt	PYTHIA		HERWIG	
Tot. Events	4,719,385	–	979,066	–
Good Run	4,658,603	–	975,615	–
L2 Trigger	2,786,636	59.82 %	608,639	62.39 %
L3 Trigger	2,719,975	97.61 %	595,200	97.79 %
Good Vertex	2,607,087	95.85 %	569,944	95.76 %
Nlep = 0	2,333,998	89.53 %	505,885	88.76 %
$N_{jet} \geq 3$	2,333,351	99.97 %	505,714	99.97 %
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$	464,067	19.89 %	109,140	21.58 %
$N_{jet} \geq 4$	452,009	97.40 %	106,141	97.25 %

Table 7.1: Effect of the trigger and prerequisites selection on PYTHIA and HERWIG inclusive $t\bar{t}$ generated events. Last requirement before neural network application is is $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$.

7.2 Systematics

7.2.1 Background prediction systematic

The systematic uncertainty on the background prediction is calculated, as already explained in previous Chapter, by comparing the number of b-tags yelded by the tagging matrix application to the actual number of positive SECVTX tags in a control sample depleted of signal contamination (we chose the one with $NNout \leq 0.9$), obtained from the TOP_MULTI_JET triggered dataset; additionally, another source of systematic uncertainties has been considered, depending on the statistical fluctuations in the sample used for the tagging matrix parameterization. As a result of these checks a 15% systematic uncertainty to the number of background b-tags returned by the tagging matrix application to data is assigned.

7.2.2 Luminosity systematic

The integrated luminosity calculation is based on the instantaneous luminosity measurement provided by the CLC detector described in Sec. 3.5. Two components of uncertainty play a role in the luminosity measurement determination: the acceptance and operation of the luminosity monitor (the CLC detector) and the theoretical uncertainty of the total inelastic $p\bar{p}$ cross section ($60.7 \pm 2.4\ mb$). The uncertainties on these quantities are 4.2% and 4.0% respectively, giving a total uncertainty of 5.8% on the integrated luminosity calculated for any given CDF dataset [2].

7.2.3 Monte Carlo generator dependent systematics

The base Monte Carlo sample adopted for this work consists of almost 4 millions inclusive $t\bar{t}$ events (exactly $3,979,960$ after the Good Run requirement) generated using PYTHIA with $M_{top} = 175\ GeV/c^2$, and with corresponding integrated luminosity of 594 fb^{-1} assuming $\sigma_{t\bar{t}} = 6.7\ pb$.

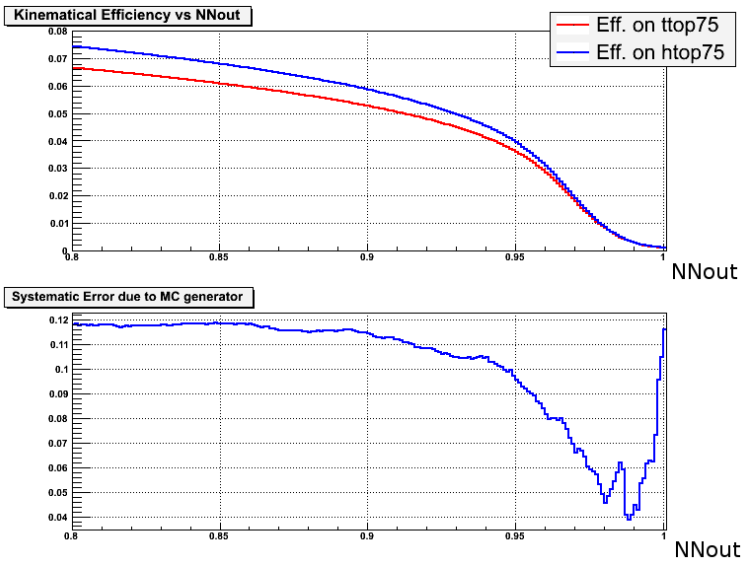


Figure 7.3: Top: kinematical efficiency of trigger, prerequisites and neural network selection versus cut applied on neural network output for both PYTHIA Monte Carlo sample (ttop75) and HERWIG sample (htop75) generated at $M_{top} = 175\ GeV/c^2$. Bottom: Monte Carlo generator dependent systematic versus cut on neural network output. The error peak for neural network output cuts close to 1 is due to low statistics effects.

The neural network selection optimization was derived using these PYTHIA inclusive $t\bar{t}$ Monte Carlo events. In order to evaluate the generator dependence of the kinematical efficiency computed for signal events we used a sample of almost 1 million (exactly 975,615 after the Good Run requirement) $t\bar{t}$ events generated with HERWIG, corresponding to an integrated luminosity of 145.6 fb^{-1}.

All these samples are processed through the CDF detector and trigger simulation, as described in Sec. 6.1.

Tab. 7.1 shows the effect of each cut of trigger and prerequisites selection, before neural network application, for inclusive $t\bar{t}$ events generated with PYTHIA and HERWIG. The efficiency of each cut with respect to the previous one is also reported.

The overall systematic uncertainty to be assigned to generator effects can then be computed for each neural network output cut as:

$$syst_{gen}(cut) = \frac{\Delta\epsilon(cut)}{\epsilon(cut)} = \frac{\epsilon_{HERWIG}(cut) - \epsilon_{PYTHIA}(cut)}{\epsilon_{PYTHIA}(cut)} \tag{7.3}$$

where $\epsilon_{PYTHIA}(cut)$ and $\epsilon_{HERWIG}(cut)$ are the kinematical efficiency for the chosen cut on $t\bar{t}$ inclusive Monte Carlo events generated with PYTHIA and HERWIG, respectively. Fig. 7.3 shows the results of this calculation in the $0.8 - 1.0$ neural network output cut range.

For the optimized cut we chose in previous chapter $NNout \geq 0.92$ the corresponding systematic uncertainty to be assigned to generator dependence effects is $syst_{gen} = 10.84\%$.

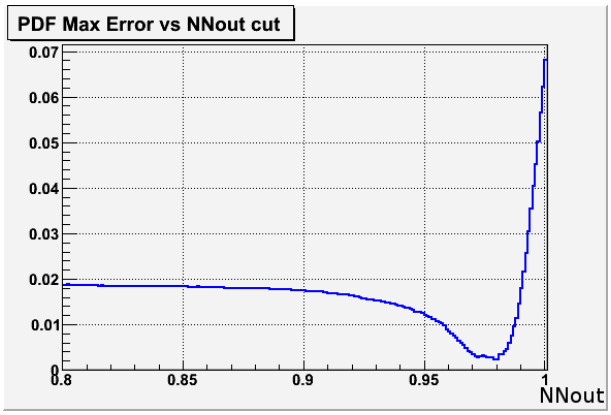


Figure 7.4: PDF dependent systematic, obtained with Monte Carlo reweighting tecnique, versus cut on neural network output. The error increases for neural network output cuts close to 1 because of low statistics effects.

7.2.4 PDF-related systematics

The parton distribution functions (PDFs) chosen for the standard CDF Monte Carlo generation correspond to the CTEQ parameterization outlined in [3]. There are uncertainties associated with this parameterization, since the usage of different parameterizations of the PDFs could slightly change the kinematics and thus the acceptance for signal events.

In order to account for these effects, we used a standard Monte Carlo reweighting tecnique. Instead of generating new samples for each different PDF, we reweighted the events already generated with PYTHIA according to different PDF eigenvectors. The weight for each event is calculated as the ratio of the new PDFs with respect to the standard one. We then sum the weights in order to determine the effect on the total kinematic efficiency [4].

The results of the calculation for neural network output cuts in the range $0.8-1.0$ are shown in Fig. 7.4. For the optimized cut we chose in previous chapter $NNout \geq 0.92$, as a result of this approach we set a $syst_{PDF} = 1.64\%$ systematic uncertainty associated with our choice of PDFs.

7.2.5 ISR/FSR-related systematics

In general it is very difficult for Monte Carlo generators to model accurately initial and final state radiation processes. If more or less extra radiation is present in the event with respect to the default values set in the base Monte Carlo sample, the event kinematics could change affecting the kinematic efficiency determination. Indeed the presence of less or more radiation associated to the $t\bar{t}$ production can alter the acceptance of the N_{jet} and $\not\!\!E_T/\sqrt{\Sigma E_T}$ requirements.

We evaluated this effect using different inclusive Monte Carlo $t\bar{t}$ samples generated with different tunings for initial (ISR) and final state (FSR) radiation: less/more ISR, and less/more FSR.

The impact of trigger and prerequisites selection for different ISR/FSR radiation settings is presented in Tab. 7.2.

N evt MC_{incl}	less ISR		more ISR	
Tot. Events	1,180,947	–	1,172,571	–
Good Run	1,165,597	–	1,157,221	–
L2 Trigger	701,208	60.16 %	714,682	61.76 %
L3 Trigger	684,004	97.55 %	698,426	97.73 %
Good Vertex	656,016	95.91 %	669,708	95.89 %
Nlep = 0	588,301	89.68 %	598,031	89.30 %
$N_{jet} \geq 3$	588,128	99.97 %	597,845	99.97 %
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$	118,009	20.07 %	122,629	20.51 %
$N_{jet} \geq 4$	114,694	97.19 %	119,349	97.33 %
	less FSR		more FSR	
Tot. Events	1,150,650	–	1,142,486	–
Good Run	1,135,300	–	1,127,136	–
L2 Trigger	693,212	61.06 %	684,307	60.71 %
L3 Trigger	676,968	97.66 %	668,192	97.65 %
Good Vertex	648,582	95.81 %	640,481	95.85 %
Nlep = 0	579,746	89.39 %	573,900	89.60 %
$N_{jet} \geq 3$	579,571	99.97 %	573,729	99.97 %
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$	118,308	20.41 %	115,643	20.16 %
$N_{jet} \geq 4$	115,022	97.22 %	112,548	97.32 %

Table 7.2: Effect of ISR/FSR radiation variation on trigger and prerequisites selection, before neural network application.

Here and in the next sections, we will calculate systematic effects for each cut on neural network output with the following approach: taking as an example the systematic uncertainty to be related to initial state radiation effect we will compute it as:

$$syst_{ISR}(cut) = \frac{|\epsilon_{+ISR}(cut) - \epsilon_{-ISR}(cut)|}{2\epsilon_{PYTHIA}(cut)}$$

when our nominal value for the kinematical efficiency $\epsilon_{PYTHIA}(cut)$ is in between the values $\epsilon_{+ISR}(cut)$ and $\epsilon_{-ISR}(cut)$ we use for comparison; when it is not, we will use half the maximum difference:

$$syst_{ISR}(cut) = \frac{max\left(|\epsilon_{+ISR}(cut) - \epsilon_{PYTHIA}(cut)|, |\epsilon_{PYTHIA}(cut)\epsilon_{-ISR}(cut)|\right)}{2\epsilon_{PYTHIA}(cut)} \tag{7.4}$$

The same will hold on the other hand for final state radiation effect, wich we will compute for each cut as:

$$syst_{FSR}(cut) = \frac{|\epsilon_{+FSR}(cut) - \epsilon_{-FSR}(cut)|}{2\epsilon_{PYTHIA}(cut)}$$

or

$$syst_{FSR}(cut) = \frac{max\left(|\epsilon_{+FSR}(cut) - \epsilon_{PYTHIA}(cut)|, |\epsilon_{PYTHIA}(cut) - \epsilon_{-FSR}(cut)|\right)}{2\epsilon_{PYTHIA}(cut)}$$

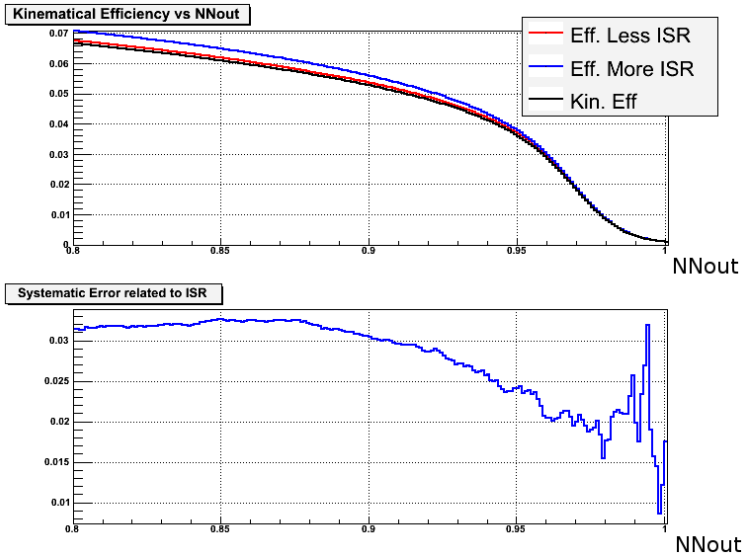


Figure 7.5: Top: kinematical efficiency of trigger, prerequisites and neural network selection versus cut applied on neural network output for PYTHIA Monte Carlo sample (ttop75) and samples generated with more and less initial state radiation. Bottom: Initial state radiation systematic uncertainty versus cut on neural network output. The behaviour of the error function for neural network output cuts in the region close to 1 is due to low statistics effects.

according to the criteria described above.

Fig. 7.5 and Fig. 7.6 show the results of this calculation in the $0.8 - 1.0$ neural network output cut range for both the ISR and FSR contributions respectively.

For the optimized cut $NNout \geq 0.92$, we can estimate a systematics to be assigned to initial state radiation effects of $syst_{ISR} = 2.86\%$, and a systematic for final state radiation effects of $syst_{FSR} = 1.71\%$. Summing in quadrature the two effects we can estimate a total systematic uncertainty to be assigned to initial and final state radiation effects of $syst_{ISR/FSR} = 3.33\%$.

7.2.6 Systematics due to the jet energy response

In this section we discuss the systematic uncertainty related to the jet energy response. In Sec. 4.3.1 the total systematic uncertainty on the corrected jet E_T was found to vary in the range [3,10]%, where the extreme values are reached for high and low jet E_T, respectively. Moreover, the uncertainty associated to the jet energy response was found to be largely independent of the level of correction applied but to be mostly arising from the jet description provided by the Monte Carlo simulation.

In order to account for the jet response systematic in the cross section measurement, we varied the corrected jet energies within $\pm 1\sigma$ of their corresponding systematic uncertainty. Therefore, signal trigger and prerequisites efficiencies are recalculated after these variations. The results are provided in Tab.7.3.

As described in previous section, we can assign a systematic uncertainty de-

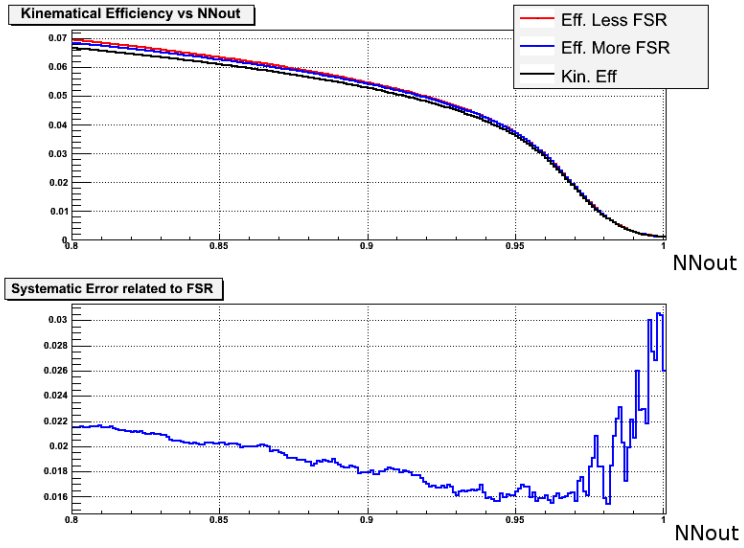


Figure 7.6: Top: kinematical efficiency of trigger, prerequisites and neural network selection versus cut applied on neural network output for PYTHIA Monte Carlo sample (ttop75) and samples generated with more and less final state radiation. Bottom: Final state radiation systematic uncertainty versus cut on neural network output. The behaviour of the error for neural network output cuts close to 1 is due to low statistics effects.

N evt MC_{incl}	standard jet corrs	$+1\sigma$ jet systs	-1σ jet systs
Total	4,719,385	4,719,385	4,719,385
Prereq	2,333,998	2,333,998	2,333,998
$N_{jet} \geq 4$	2,306,282	2,310,830	2,300,028
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$	452,009	469,481	449,568

Table 7.3: Effect of the jet energy correction within their uncertainty on the trigger and prerequisites selection on $t\bar{t}$ inclusive events, before neural network application.

pending on the cut we apply on the neural network output as follows:

$$syst_{jetcorr}(cut) = \frac{|\epsilon_{jetcorr,+1\sigma}(cut) - \epsilon_{jetcorr,-1\sigma}(cut)|}{2\epsilon_{kin}(cut)}$$

when our nominal value for the kinematical efficiency $\epsilon_{kin}(cut)$ is in between the values $\epsilon_{jetcorr,+1\sigma}(cut)$ and $\epsilon_{jetcorr,-1\sigma}(cut)$, while in the other case we will use half the maximum difference defined according to Eq. 7.4.

Fig. 7.7 show the results of this calculation in the $0.8 - 1.0$ neural network output cut range.

For the optimized cut $NNout \geq 0.92$, we can estimate a systematic uncertainty to be assigned to jet energy response of $syst_{jetcorr} = 4.73\%$.

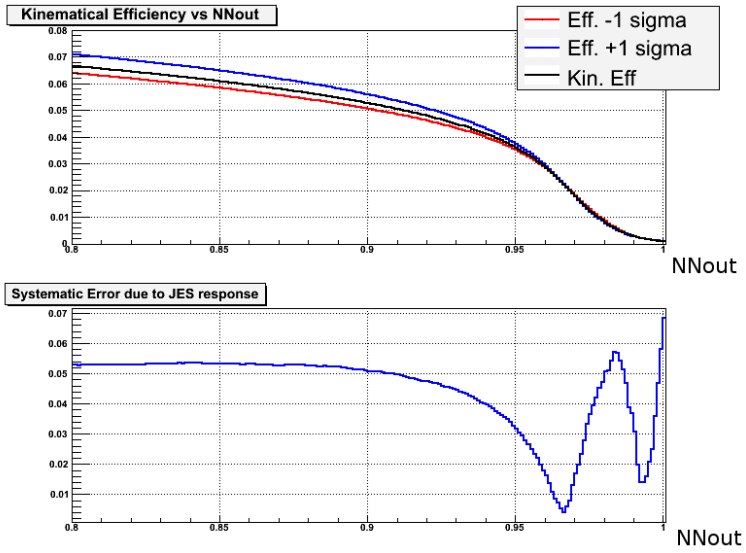


Figure 7.7: Top: kinematical efficiency of trigger, prerequisites and neural network selection versus cut applied on neural network output for the Monte Carlo sample (ttop75) with standard jet corrections and with jet energy corrections shifted by $\pm\sigma$ of their systematic error. Bottom: Systematic uncertainty due to jet energy repsonse versus cut applied on neural network output. In the neural network output cut region close to 1 low statistics effects arise, causing the error to increase.

7.2.7 b-tagging scale factor systematics

As described in Sec. 6.4, the SECVTX efficiency scale factor we used in this analysis, to count the number of b-tags on Monte Carlo events is $SF = 0.95 \pm 0.050$. Since the average number of b-tags per $t\bar{t}$ event, ϵ_{tag}^{ave}, enters directly in the cross section measurement we have to compute the systematics effect related to its determination.

In particular, to account for the scale factor uncertainty we varied it from its central value of 0.95 within the $\pm 1\sigma$ range and we determined the difference in terms of average number of b-tags per event on the Monte Carlo sample with respect to the standard value, taking into account that the SECVTX scale factor has the same central value for both b- and c-quarks, but for the latter has a doubled uncertainty: $SF_b = 0.95 \pm 0.050$, $SF_c = 0.95 \pm 0.100$.

For each cut on neural network output we can assign the following systematic uncertainty:

$$syst_{\epsilon_{tag}}(cut) = \frac{|\epsilon_{tag,+1\sigma}(cut) - \epsilon_{tag,-1\sigma}(cut)|}{2\epsilon_{tag}^{ave}(cut)} \tag{7.5}$$

The results are shown in Fig. 7.8. As expected, the systematic uncertainty due to the scale factor application does not depend much on the choice of the cut on the network output, since it only rescales the number of positive tags in a given sample.

For the cut $NNout \geq 0.92$, we can estimate a systematic uncertainty to be assigned to b-tagging scale factor application of $syst_{\epsilon_{tag}} = 3.98\%$.

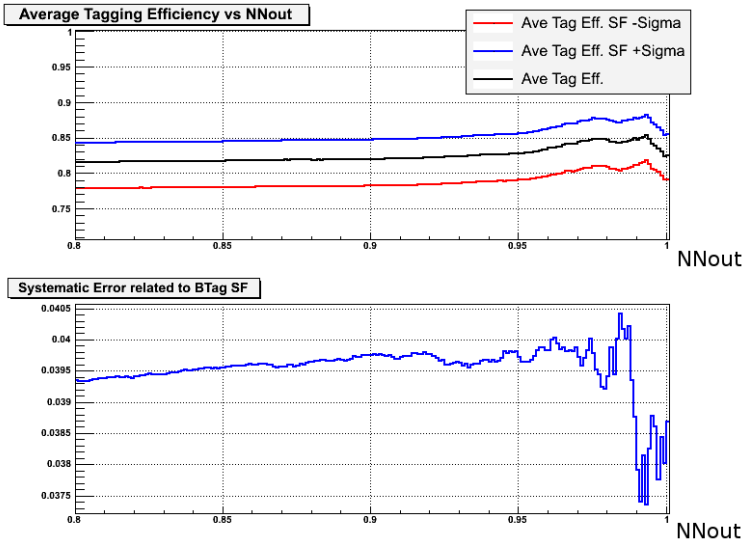


Figure 7.8: Top: Average tagging efficiency in the sample obtained after trigger, prerequisites and neural network selection versus cut applied on neural network output for the Monte Carlo sample (ttop75) with standard b-tagging Scale Factor and with Scale Factor shifted by $\pm\sigma$ of its systematic error. Bottom: Systematic uncertainty due to b-tagging scale factor application versus cut applied on neural network output. The behaviour of the error function in the output region close to 1 is due to low statistic effects.

7.2.8 Trigger systematics

Trigger systematics have already been studied and characterized for the old TOP_MULTI_JET trigger (used up to Run 194328) [6] using the following tecnique. A sample of collider data called "Single Tower-10" is used, triggered with the following requirements

- at Level 1: at least one calorimetric tower with $E_T \geq 10\ GeV$.
- at Level 2: a static prescaling factor of 1K.
- at Level 3: auto-accept.

with a corresponding integrated luminosity of $196 \pm 12\ pb^{-1}$. This dataset is then used to extract the efficiency of the TOP_MULTI_JET trigger on a data-driven basis, evaluating its efficiency by applying its L2 requirements directly on "Single Tower-10" triggered data.

Then the systematic uncertainty affecting the trigger efficiency measurement on Monte Carlo $t\bar{t}$ events is evaluated by comparing trigger turn-on curves for Tower-10 data and $b\bar{b}$ and $b\bar{b}+6$ partons Monte Carlo samples. The trigger turn-on curves are derived as functions of of the 4^{th} jet L5-corrected E_T in the event.

From a study of the mismatch of turn-on curves between Monte Carlo samples and data, a trigger efficiency systematic of a few percent (2.0%) is derived.

Since we didn't have enogh time to reproduce such a detailed study of the trigger turn-on curves for the new TOP_MULTI_JET trigger, we will rely on this

Source	Method	Uncertainty
ϵ_{kin} systematics		
Generator dependence	$\frac{\lvert\epsilon_{PYTHIA}-\epsilon_{HERWIG}\rvert}{\epsilon_{PYTHIA}}$	10.84 %
PDFs	MC reweighting	1.64 %
ISR/FSR	samples comparison	3.33 %
Jet Energy Scale	$\frac{\lvert\epsilon_{jetcorr,+1\sigma}-\epsilon_{jetcorr,-1\sigma}\rvert}{2\epsilon_{kin}}$	4.73 %
Trigger simulation	turn-on curves	2.0 %
ϵ_{tag} systematics		
SecVtX scale factor	$\frac{\lvert\epsilon_{tag,+1\sigma}-\epsilon_{tag,-1\sigma}\rvert}{2\epsilon_{tag}}$	3.98 %
Tagging matrix systematics		
Data control samples	N_{obs}/N_{exp}	15.0 %
Luminosity systematics		
Luminosity measurement	–	5.8 %

Table 7.4: Summary of the sources of systematic uncertainty. Trigger simulation systematic uncertainty is based on a determination on the old TOP_MULTI_JET trigger and is to be considered a preliminary result.

Variable	Symbol	Input Value	Output Value
Integrated Luminosity (pb^{-1})	$\mathcal{L}$	1906.8 ± 110.6	1907.1 ± 111.7
Observed Tags	N_{obs}	627	–
Expected Tags	N_{exp}^{corr}	137.5 ± 23.4	137.1 ± 23.3
Kin. efficiency (%)	ϵ_{kin}	4.907 ± 0.613	4.889 ± 0.668
Ave. b-tagging efficiency	ϵ_{tag}^{ave}	0.8188 ± 0.0325	0.8187 ± 0.0327

Table 7.5: Input and output values of the likelihood maximization.

previous determination to give a *preliminary* cross section measurement in this channel.

7.3 Cross section measurement

The summary of all the sources of systematic uncertainty to the cross section evaluation is listed in Tab. 7.4.

Now that we have evaluated all the sources of systematic uncertainty affecting the kinematical selection efficiency as well as the determination of the average number of b-tags per $t\bar{t}$ event and the background prediction, we are ready to use all the ingredients described in 7.1 to perform the cross section measurement.

We remind that we will interpret the excess in the number of tags defined as $N_{obs} - N_{exp}^{corr}$ as a sign of $t\bar{t}$ production and it will be used for the cross section measurement.

As already mentioned, the cross section is measured by maximizing $\log \mathcal{L}$, where

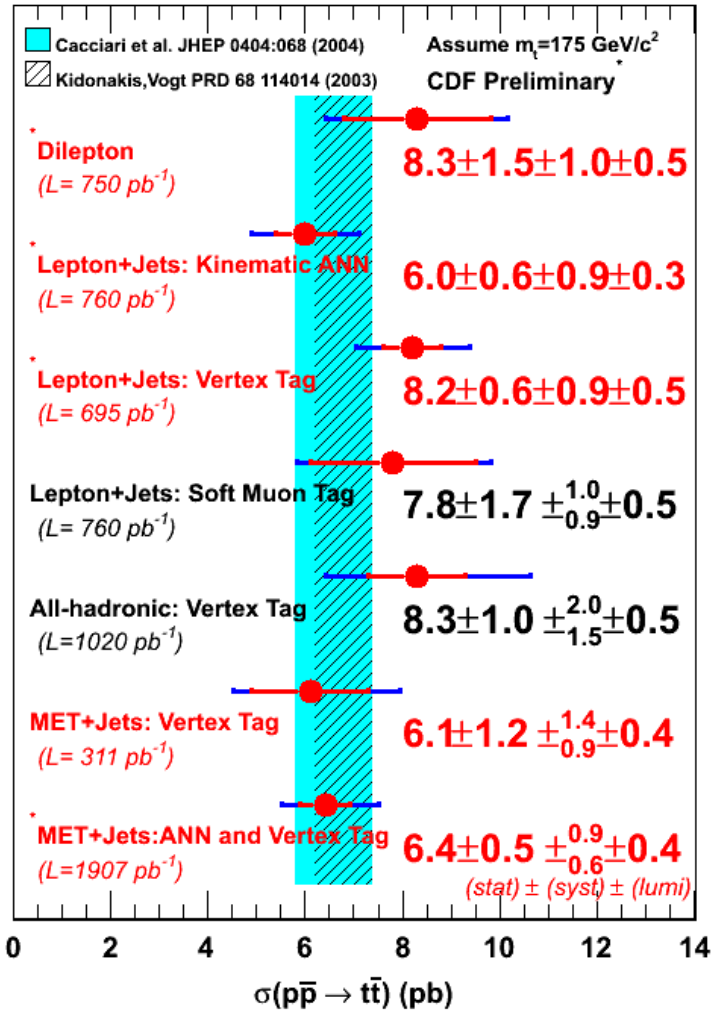


Figure 7.9: Latest CDF cross section results. Last two rows show the measurement obtained with kinematical selection and secondary vertex tag on 311 pb^{-1} and the determination presented in this work.

the likelihood function is defined as follows:

$$\mathcal{L} = e^{-\frac{(L-\bar{L})^2}{2\sigma_L^2}} \cdot e^{-\frac{(\epsilon_{kin}-\bar{\epsilon}_{kin})^2}{2\sigma^2_{\bar{\epsilon}_{kin}}}} \cdot e^{-\frac{(\epsilon^{ave}_{tag}-\bar{\epsilon}^{ave}_{tag})^2}{2\sigma^2_{\epsilon^{ave}_{tag}}}} \cdot e^{-\frac{(N^{corr}_{exp}-\bar{N}^{corr}_{exp})^2}{2\sigma^2_{N^{corr}_{exp}}}} \cdot \frac{(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N^{corr}_{exp})^{N_{obs}}}{N_{obs}!} \cdot e^{-(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N^{corr}_{exp})} \tag{7.6}$$

The central value is given by the likelihood maximization, that is:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N^{corr}_{exp}}{\epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L} \tag{7.7}$$

The input and output parameters of the likelihood maximization are quoted in Tab.7.5.

The measured cross section value is:

$$\begin{aligned} \sigma_{t\bar{t}} &= 6.42 \pm 0.51\ (stat)\ ^{+0.98}_{-0.74}(syst)\ pb \\ &= 6.42\ ^{+1.1}_{-0.9}\ pb \end{aligned}$$

Separating the contribution due to the uncertainty coming from the luminosity measurement, we can rewrite the result as follows:

$$\sigma_{t\bar{t}} = 6.42 \pm 0.51\ (stat)\ ^{+0.90}_{-0.62}(syst)\ ^{+0.40}_{-0.37}(lum)\ pb$$

To have an idea on how this preliminary, not yet officially approved by the collaboration, cross section measurement compares with the other CDF determinations, Fig. 7.9 shows a summary of latest results from the experiment together with the previous $t\bar{t} \to \not\!\!E_T + jets$ cross section determination obtained on 311 pb^{-1} in [5]. We note that all measurements reported assume $M_{top} = 175\ GeV/c^2$ and that our preliminary determination is extracted from the highest luminosity sample.

As Fig. 7.9 summarizes, CDF has produced several measurements of the top pair production cross section in the di-lepton, lepton+jets and all-hadronic channels. Since several of the measurements are based on totally or partially uncorrelated data samples and have different sources of systematic uncertainty, the combination of the results reduces the experimental uncertainty.

The combination technique uses the BLUE algorithm [7], which stands for Best Linear Unbiased Estimate, and needs as inputs the statistical, systematic uncertainties as well as the correlation between different analyses. These are used to construct a covariance matrix, which is inverted to obtain weights for each analysis. Thanks to the fact that it was derived on a data sample completely uncorrelated to the ones used in the remaining analyses, the previous cross section measurement in the $\not\!\!E_T + jets$ channel in 311 pb^{-1} was found in [8] to carry a relative weight of 17% on the final combination, thus giving a very important contribution to the combined cross section determination. We finally note that the new cross section measurement we have obtained in this work has a lower statistic uncertainty than the previous one, thanks to the luminosity increase of the dataset and to the fact that we could compensate the signal loss caused by the introduction of a higher treshold in the Level 2 TOP_MULTI_JET trigger by means of a neural network selection. Indeed, also this measurement is derived from a data sample that was chosen by prerequisites to be orthogonal to the ones used for the other cross section determinations at CDF. We thus think that once approved officially by the collaboration, this cross section measurement could have an important impact in the combination of the results obtained by the CDF experiment.

Bibliography

[1] A. Castro *et al.*, *Top Production Cross Section in the all-hadronic channel*, CDF Internal Note 3464.

[2] S. Jindariani *et al.*, *Luminosity Uncertainty for Run 2 up until August 2004*, CDF Internal Note 7446.

[3] J. Pumplin *et al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** 012 (2002).

[4] O. Gonzalez *et al.*, *Uncertainties due to PDFs for the gluino-sbottom search*, CDF Internal Note 7051.

[5] A. Abulencia *et al.*, *Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\ TeV$ using $\not\!\!E_T + jets$ events with secondary vertex b-tagging*, Phys. Rev. Lett. **96**, 202002 (2006).

[6] G. Cortiana *et al.*, *Report on trigger systematic uncertainty evaluation for the ttbar->met+jets cross section*, CDF Internal Note 7964.

[7] L. Lions and D. Gibaut, *How to combine correlated estimates of a single physical quantity*, NIMA **A**270, 110 (1988).

[8] CDF Collaboration, *Combination of CDF top pair production cross section measurements*, CDF Internal Note 7794.

Conclusions

We presented a research aimed at the isolation of the $t\bar{t} \rightarrow \not\!\!E_T + jets$ signal by means of neural network tools from a dataset containing "multijet" triggered events with a total integrated luminosity amounting to 1.9 fb^{-1}.

The decay channel has been extracted using neutrino signatures such as presence of high $\not\!\!E_T$ in the event and explicitly vetoing well identified high-P_T electrons or muons from W boson decay.

A 2-hidden layers neural network trained with input variables related to jet characteristics and energy and event topology and energy has been used to classify and discriminate between top-like events obtained from a Monte Carlo sample generated at $M_{top} = 175\ GeV/c^2$ and background processes contained in the data sample after prerequisites requirements.

Secondary vertex b-tagging algorithm has been exploited to indentify heavy flavour jets due to top quark decay, while the amount of tags coming from background processes has been evaluated by means of a parameterization of the b-tagging rate as a function of the jet transverse energy, jet number of tracks and projection of the $\not\!\!E_T$ of the event along the jet direction, in a data sample with negligible signal contamination containing exactly 3 tight jets.

Once checked the performance of the tagging parameterization and the correctness of its predictions, the optimized cut on the neural network output $NNout \geq 0.92$ has been computed by minimizing the relative statistical error on the cross section measurement.

With the resulting selection we obtained a pre-tagging sample of 1415 events: in order to derive our final cross section measurement, we added the requirement of the presence of at least one jet identified as originating from a b-quark, observing 627 b-tags. Thanks to our b-tagging rate parameterization we accounted for 490 tags coming from $t\bar{t}$ events.

A likelihood function in which the input parameters are subject to Gaussian constraints was finally used for a proper determination of the top pair production cross section, after having taken into account the possible sources of systematic uncertainties. Assuming a top quark mass of 175 GeV/c^2, our final measurement was:

$$\begin{aligned} \sigma_{t\bar{t}} &= 6.42 \pm 0.51\ (stat)\ ^{+0.98}_{-0.74}(syst)\ pb \\ &= 6.42\ ^{+1.1}_{-0.9}\ pb \end{aligned}$$

in agreement with Standard Model predictions and with previous determinations. Moreover, being derived from a data sample that was chosen by prerequisites to

be orthogonal to the ones used for the other cross section determinations at CDF, this measurement promises to be particularly important in the combination of the results obtained by the experiment.

Issues and Future plans

As discussed during the systematic uncertainties analysis, this result lacks a correct and detailed description of the trigger systematic error related to the impact of the new Level 2 TOP_MULTI_JET trigger to our selected sample; our estimation relies instead on a previous determination made for the old trigger path. Even if we expect the new trigger systematic error to be of the same order of magnitude of the old one (i.e. $\leq 10\%$), we should nevertheless consider this cross section measurement as a preliminary result still subject to changes.

Additionally, we believe there's still room for improvement in the b-tagging rate parameterization: many different choices of variables have been tested and the one used in this work has been selected for its best performances, but a finer tuning of the limits of the matrix bins could provide a better prediction of the background in the control samples and consequently a lower systematic uncertainty on the tagging rate parameterization predictions in the selected sample.

Acknowledgements

Among the many outstanding people I met during these last three years, several started by being just collegues and soon became really good friends; I'm sure that having the chance of getting in touch and sharing with them an important part of my life was the best thing ever about this whole Ph.D. experience. A countless number of them is responsible for what I am now (so shame on you!) and for the things I learned (still too few, I admit it!) in the field of experimental high energy physics, but some among this impossible to remember list played a special role and really deserve to be mentioned.

The first person I would like to thank is dr. *Silvia Amerio* (Towanda!): without her spreading passion for this field, her constant support and encouragement during the toughest periods of my work and her calm and serene attitude towards problem solving, my whole thesis would have never been possible. She guided and advised me constantly through the progress of the analysis, and also offered me the gift of her friendship: Silvia, I'll always be grateful with all my heart for everything you did for me!

I also wish to thank dr. *Giorgio Cortiana* for having "inspired" this work, for his continuous helpfulness and for sharing with me not only his great experience about the MET+jets channel, but also many useful pieces of code and technical informations that were fundamental for this thesis. I will never forget how much fun I had in the evenings out with Giorgio during our Summer of work spent at FNAL, as I won't be able to forget (for its consequences!) our wonderful idea of running the infamous "Tevatron lap" at 5 a.m. in the morning with tequila as fuel.

I'm also thankful to *Donatella Lucchesi* for her constant assistance, encouragement, support and patience throughout this experience; I owe the happy ending of this story to her.

I want to show all my appreciation to *Simone Pagan*, for his help in our "beloved" geeky computing stuff and for having saved for me all the fun about LcgCAF!

My sincere gratefulness goes to *Igor* and *Loredana* for welcoming me in their house and treating me as family: I owe to them much more than I can write.

I have no words to describe my gratitude towards *Michele Giunta* for his friendship throughout these years and for the wonderful time spent together; despite his tastes regarding music and the strange soundtrack of our nights out, my experience

in CDF would have never been so exciting without him!

I thank *Francesco Sborzacchi* for existing: he can turn the world's most boring, soporific and unexciting place in a fun fair, I feel blessed in being among his friends.

Moreover, thanks to (in random order): *Luca Brigliadori*, for letting me understand I could run a marathon worse than I do physics; *Renzo*, for the training in running and software development: sorry for the poor results! *Stefano Torre*, for being a good host (but please keep watermelon out of his touch!) and for suggesting me a different, wealthier career; *Fabrizio*, for his help in trying to fix the path of my thesis: sorry! *Paolo Mastrandrea*, for the unforgettable dinners and evenings at Roma house, not to mention the fun with him around the Windy City; the *CAF team*, for giving me the opportunity of learning something that might be useful in the real world.

I also want to thank all the wonderful people met in Trento, I regret being so lazy I didn't spend enough time with them: *David & Giacomo* for their "travel agency", *Mario*, *Giovanni*, *Sara* and all the others that made me feel like home.

Finally, I want to thank *Patrizia*, *Mario* and *Michele* for having guided me all this time, and *Marilena* for giving me the chance of sharing my life with her. All I've done so far would have never been possible without their presence, support and understanding.

March 6, 2008
Gabriele Compostella