FERMILAB-THESIS-2006-97

UNIVERSITÀ DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Fisica "G. Galilei"

DOTTORATO DI RICERCA IN FISICA

CICLO XVIII

First measurement of the top quark pair production cross section in the $\not\!\!E_T$ + jets channel at the Collider Detector Fermilab

Coordinatore: Prof. Attilio Stella

Relatore: Prof. Giovanni Busetto

Co-relatore: Dr. Tommaso Dorigo

Dottorando: Dr. Giorgio Cortiana

DATA CONSEGNA TESI 2 Gennaio 2006

Contents

First measurement of the top quark pair production cross section in the $\not\!\!E_T$ + jets channel at the Collider Detector Fermilab

Abstract

We present a measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} =$ 1.96 TeV which uses for the first time events with a signature of significant missing transverse energy and jets. The measurement is thus sensitive to leptonic W decays regardless of lepton type. Heavy flavor jets from top quark decay are identified with a secondary vertex tagging algorithm. From 311 pb^{-1} of data collected by the Collider Detector at Fermilab we measure a production cross section of $5.9 \pm 1.2^{+1.4}_{-1.0}$ pb for a top quark mass of 178 GeV/c^2, in agreement with previous determinations and standard model predictions.

Prima determinazione della sezione d'urto di produzione di coppie $t\bar{t}$ nello stato finale $\not\!\!E_T$ + jets all'esperimento CDF

Prefazione

Presentiamo una misura della sezione d'urto di produzione di coppie $t\bar{t}$ in collisioni $p\bar{p}$ all'energia del centro di mass di 1.96 TeV, per la quale per la prima volta vengono utilizzati eventi con alta energia trasversa mancante e jet multipli. La misura è in questo modo sensibile ai decadimenti leptonici dei bosoni W indipendentemente dall'identificazione di leptoni. I quark pesanti provenienti dal decadimento del quark top sono identificati attraverso un algoritmo di identificazione di vertici secondari di decadimento. Dall'analisi di 311 pb^{-1} di dati raccolti dal rivelatore CDF, assumendo una massa del quark top di 178 GeV/c^2, la misura di sezione d'urto di produzione$t\bar{t}$ è misurata essere pari a $5.9 \pm 1.2^{+1.4}_{-1.0}$ pb, in perfetto accordo con le stime teoriche nel contesto del Modello Standard e precedenti determinazioni.

Introduction

The work described in this thesis is focused on the inclusive search of the $t\bar{t} \to \not\!\!E_T$+jets process 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 (SM) the calculated cross section for pair production is $6.1^{+0.6}_{-0.8}$ pb for a top mass of 178 GeV/c^2, and varies by ± 0.2 pb for every $\mp 1\ GeV/c^2$ variation of the top mass in the range $170\ GeV/c^2 < m_t < 190\ GeV/c^2$. Because the Cabibbo-Kobayashi-Maskawa matrix element V_{tb} is close to unity and the top width is large, the SM top quark decays to a W boson and a b quark almost 100% of the time. The final state of top quark pair production thus includes two W bosons and two b-quark jets. When only one W decays leptonically, the $t\bar{t}$ event typically contains a charged lepton, missing transverse energy from the undetected neutrino, and four high transverse momentum jets, two of which originate from b quarks. Previous cross section analyses select this golden $t\bar{t}$ signature by requiring well identified leptons ($e,\ \mu$) with high transverse momentum. Since top pair final states are directly related to W branching ratios, the measurement of the $t\bar{t}$ rate into a particular final state measures both the production and decay properties of the top quark. Unexpected results could indicate either a non-SM source of top-like events or a modification of the top quark decay branching ratios.

In this thesis we describe a $t\bar{t}$ production cross section measurement which is sensitive to leptonic W decays regardless of the lepton type, and has a sizable acceptance to τ-lepton decays of the W boson. The measurement uses data collected by a multijet trigger, and selects top decays by requiring a high-P_T neutrino signature rather than charged lepton identification.

In Chapter 1, a brief review of our knowledge of elementary particles, summarized in the theory called Standard Model will be presented, with particular attention on the Standard Model top quark properties and its experimental signatures.

In Chapter 2, an overview of the Tevatron accelerator chain that provides $p\bar{p}$ collisions at the center-of-mass energy of $\sqrt{s} = 1.96\ TeV$ will be provided. This Chapter will describe the final proton and antiproton beams production procedure, through a complex acceleration apparatus which involves different stages, spanning from proton and antiproton production, their acceleration and transfer towards different sub-systems, to their actual collision in specified interaction points where particle detectors are installed.

The CDF detector, used for the study of $p\bar{p}$ collisions provided by the Tevatron, and the physics objects reconstruction tools will be described in Chapter 3.

Chapter 4 will be devoted to the description of the main characteristics of the $t\bar{t} \to$

$\not\!\!E_T$+jets decay channel by which to distinguish the top pair production from background processes. The optimization of the event selection and the strategy adopted for the background prediction, relying on b-jets identification rate parameterizations will be highlighted in Chapter 5; on the other hand, Chapter 6 will provide a characterization of the final data sample, which will be used for our measurement, by means of binned likelihood fits to kinematical variables.

The discussion of the systematic uncertainties, and of the methodology applied for the top pair production cross section measurement will be presented in Chapter 7, where in addition the latest CDF cross section determinations combination procedure and results will be provided.

Chapter 1

Theoretical overview

During the last decades high energy physics has played a crucial role in changing our understanding of the fundamental interactions.

In the first part of this Chapter a brief review of our knowledge of elementary particles, summarized in the theory called *Standard Model*, will be presented, while in the second part particular attention will be given on the Standard Model top quark properties and its experimental signatures.

1.1 The Standard Model of particle physics

The extraordinary predictive capability and the formal elegance that determined in the last 50 years the success of *Quantum Electro-Dynamics* (QED), on which our present understanding of electromagnetic interactions is based, induced theoretical physicists to try to extend its formulation to an analogous gauge theory able to describe also the weak and strong nuclear interactions. These models are based on the simple idea that, by requiring the physical laws to be invariant with respect to opportune local gauge transformations, fermionic fields describing ordinary matter can be *naturally* associated with gauge fields responsible of their interactions. The computation of physical quantities leads to divergences that can be reabsorbed in a finite number of measurable parameters (mass, charge, etc.). The success of this procedure, named *renormalization*, is ensured by the invariance under gauge transformations of the theory itself.

The Standard Model[1, 2] 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-Dynamic*, 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.

Within the Standard Model framework matter is composed of elementary particles that are subject to interactions mediated by what we call *gauge particles*. The fundamental constituents of matter are classified into leptons and quarks, which are further organized into three families, called *generations*. To each generation are associated particles and their corresponding anti-particles, the latter having the same properties as the partner

	Generation		
Quarks	1^{st}	2^{nd}	3^{rd}
$Q = +\frac{2}{3}$	up (u) $M_u \approx 1.5 \div 4.5\ MeV/c^2$	charm (c) $M_c \approx 1.15 \div 1.35\ GeV/c^2$	top (t) $M_t = 172.7 \pm 2.9\ GeV/c^2$
$Q = -\frac{1}{3}$	down (d) $M_d \approx 4.0 \div 8.0\ MeV/c^2$	strange (s) $M_s \approx 80 \div 130\ MeV/c^2$	bottom (b) $M_b \approx 4.1 \div 4.9\ GeV/c^2$
Leptons	1^{st}	2^{nd}	3^{rd}
$Q = -1$	electron (e) $M_e = 0.511\ MeV/c^2$	muon (μ) $M_\mu = 106\ MeV/c^2$	tau (τ) $M_\tau = 1.78\ GeV/c^2$
$Q = 0$	electron neutrino (ν_e) $0 < M_{\nu_e} < 3\ eV/c^2$	muon neutrino (ν_μ) $0 < M_{\nu_\mu} < 0.19\ MeV/c^2$	tau neutrino (ν_τ) $0 < M_{\nu_\tau} < 18.2\ MeV/c^2$

Table 1.1: Fermions in the Standard Model. Mass values are from [18] and [22].

particles (especially, the same mass), but opposite charges (the *charge* of the particle is the quantum number that defines the coupling of the particle to the force carriers).

Ordinary matter is composed by particles belonging to the first generation: the electron, e^-, with electric charge $Q = -1$, its corresponding neutrino ν_e with $Q = 0$, the *up* and *down* quarks and their corresponding antiparticles (e^+, $\bar{\nu}_e$, $\bar{u}$ and $\bar{d}$). The up and down quarks, denoted by u and d, carry a fractionary electric charge $Q_u = \frac{2}{3}$ and $Q_d = -\frac{1}{3}$ respectively. Aside the electric charge, an additional quantum number is associated to quarks: *color*. The color of particles can be of three types: red, green, and blue, denoted as R, G, and B respectively, while for antiparticles the corresponding anti-colors are defined, $\bar{R}$, $\bar{G}$, $\bar{B}$. "Colored" particles are not observed in nature, therefore quarks must be confined into colorless composite particles, called *hadrons*. Hadrons are categorized as baryons and mesons, depending on their quark composition. In the quark model baryons, like the proton and the neutron, are basically constituted by three valence quarks: $p \sim uud$ and $n \sim udd$. On the other hand, mesons are composed by a quark-antiquark pair. This is the case for instance of the pions $\pi^+ \sim u\bar{d}$ and $\pi^- \sim d\bar{u}$. Second and third generation particles have identical properties to first generation ones, except the mass (Tab.1.1).

Gauge particles include the massless and chargeless photon, γ, which is exchanged in electromagnetic interactions; the massive $W^\pm$ and Z^0, with $Q = \pm 1$ and 0 respectively, mediating weak interactions; and the eight gluons, g_α, carrying strong interactions among quarks. Each gluon is massless and chargeless but carries two color charges (color + anti-color), determining the possibility for gluons to interact between themselves as well as with quarks. Gravitational interactions are not incorporated in the Standard Model framework.

1.1.1 Electromagnetic interactions

Fundamental particles are spin-$\frac{1}{2}$ fermions for which in absence of gauge fields the dynamics is governed 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 global $U(1)$ transformations, that act on the fields and their derivatives as follows:

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

Promoting the invariance from global to local, *i.e.* by allowing the parameter θ in eq. 1.2 to depend on the space-time point x, destroys the Lagrangian (eq. 1.1) invariance.

The requirement of invariance under local $U(1)$ transformations of the type $\Psi \to \Psi e^{iQ\theta(x)}$ asks the introduction of one gauge vector boson field $A_\mu(x)$, associated to the photon, which interacts with the field Ψ and whose transformations compensate non-invariant terms in the Lagrangian. In this way, the $U(1)$ gauge invariant Lagrangian of Quantum Electrodynamic (QED) will read 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 the covariant derivative, D_μ, is defined by the following:

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

and contains the interaction terms between the photon and fermions, while the propagation of the photon field is specified in terms of the field strength tensor $F_{\mu\nu}$, which is defined by:

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

1.1.2 Strong interactions

The gauge theory for strong interactions is based on $SU(3)_C$, which is a non-abelian Lie group generated by color transformations. The label C stands indeed for color and the 3 refers to the three possible color quark states.

The QCD (quantum chromodynamics) invariant Lagrangian is built similarly to the QED one. The gauge symmetry $SU(3)$ is promoted by the QCD invariant derivative:

$$D_\mu q = \left(\partial_\mu - ig_s\frac{\lambda_\alpha}{2}A_\mu^\alpha\right)q, \tag{1.6}$$

where

$$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \tag{1.7}$$

represents 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, $\alpha = 1, ..., 8$. The QCD Lagrangian is consequently defined as:

$$\mathcal{L}_{QCD} = \sum_q \bar{q}(x)(iD_\mu\gamma^\mu - m_q)q(x) - \frac{1}{4}F^\alpha_{\mu\nu}F^{\mu\nu}_\alpha; \tag{1.8}$$

in which the gluon field strength tensor is defined by the following relation:

$$F^\alpha_{\mu\nu}(x) = \partial_\mu A^\alpha_\nu(x) - \partial_\nu A^\alpha_\mu(x) + g_s f^{\alpha\beta\gamma} A_{\mu\beta} A_{\nu\gamma} \tag{1.9}$$

and contains a bilinear term which provides gluon-gluon interactions. In expression 1.9, 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:

$$\alpha_S(q^2) = \frac{4\pi}{\left(11 - \frac{2}{3}N_f(q^2)\right)\ln\left(\frac{-q^2}{\Lambda^2_{QCD}}\right)}. \tag{1.10}$$

In eq. 1.10, Λ_{QCD} is the QCD energy scale, $N_f(q^2)$ is the number of quark flavors that can be pair-produced at a given energy (*i.e.* the number of quark flavors 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, giving rise to the so-called "asymptotic freedom", but to be strong and divergent[1] at low energy. The behavior of the strong coupling constant at low energy is responsible of quark confinement into hadrons.

1.1.3 Electroweak interactions

The electroweak sector of the Standard Model is based on the gauge symmetry group $SU(2)_L \times U(1)_Y$. $SU(2)_L$ is the weak isospin group which acts on left-handed fermions; $U(1)_Y$ is the weak hypercharge group. $SU(2)_L \times U(1)_Y$ has four generators, three of which from $SU(2)_L$: $T_i = \frac{\sigma_i}{2}$ with $i = 1, 2, 3$; and the fourth one from $U(1)_Y$, $\frac{Y}{2}$. The commutation relations of the group read:

$$[T_i, T_j] = i\epsilon_{ijk}T_k; \quad [T_i, Y] = 0; \quad i, j, k = 1, 2, 3. \tag{1.11}$$

The left-handed fermions transform as doublets under $SU(2)_L$,

$$f_L \to e^{i\vec{T}\vec{\theta}} f_L; \quad f_L = \begin{pmatrix}\nu_L \\ e_L\end{pmatrix}, \quad \begin{pmatrix}u_L \\ d_L\end{pmatrix}, ... \tag{1.12}$$

whereas the right-handed fermions are singlets for $SU(2)_L$ transformations:

$$f_R \to f_R; \quad f_R = e_R, u_R, d_R, ... \tag{1.13}$$

First-generation fermion quantum numbers are provided in Tab. 1.2 and are related to each other by the following equation:

$$Q = T_3 + \frac{Y}{2}. \tag{1.14}$$

[1] When the strong coupling is large, perturbative calculations are no longer valid.

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 1.2: First-generation fermion quantum numbers within the Standard Model.

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

The Standard Model Lagrangian construction for the electroweak sector follows the same rules as for any gauge theory. In particular, $SU(2)_L \times U(1)_Y$ symmetry is promoted from global to local by replacing the field derivatives by their corresponding covariant derivatives. For a generic fermion field f, the covariant derivative reads:

$$D_\mu f = \left(\partial_\mu - ig\vec{T}\cdot\vec{W}_\mu - g'\frac{Y}{2}B_\mu\right) f, \tag{1.15}$$

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

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

$$\mathcal{L}_G = -\frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.16}$$

where the field strength tensors are defined as follows:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk}W_{\mu j}W_{\nu k} \tag{1.17}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.18}$$

The physical gauge bosons $W^\pm_\mu$, Z_μ and A_μ are obtained from the electroweak interaction eigenstates by the following expressions, in which θ is the weak mixing angle:

$$W^\pm_\mu = \frac{W_{1\mu} \mp iW_{2\mu}}{\sqrt{2}} \tag{1.19}$$

$$Z_\mu = W_{3\mu}\cos\theta - B_\mu\sin\theta \tag{1.20}$$

$$A_\mu = W_{3\mu}\sin\theta + B_\mu\cos\theta \tag{1.21}$$

The gauge invariant interactions and the fermion kinematics are generated by $\bar{f}iD_\mu\gamma^\mu f$ terms in the Lagrangian. The Standard Model Lagrangian of the electroweak sector will thus be:

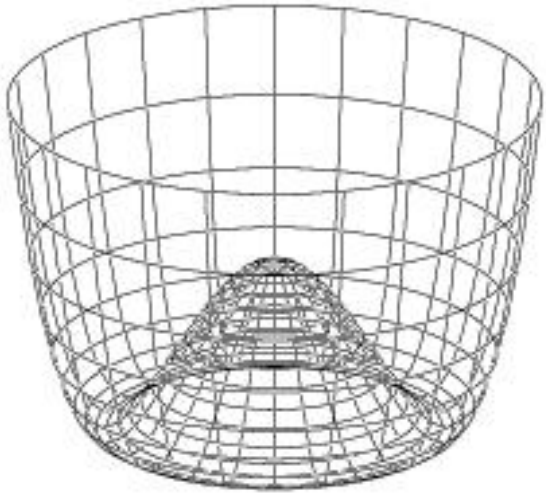


Figure 1.1: The Higgs potential

$$\mathcal{L}_{SM} = \mathcal{L}_f + \mathcal{L}_G + \mathcal{L}_{SSB} + \mathcal{L}_{YW}, \tag{1.22}$$

where the last two terms are the spontaneous symmetry breaking and Yukawa Lagrangians respectively, and are needed in order to properly introduce gauge bosons and fermion mass terms in a gauge invariant way. For $SU(2)_L \times U(1)_Y$ gauge invariance terms like $M_W^2 W_\mu W^\mu$, $M_Z^2 Z_\mu Z^\mu$ and $m_f^2 \bar{f} f$ are indeed forbidden. It will be the spontaneous symmetry breaking and the Higgs mechanism to provide the right mass generation as will be described in the following.

Spontaneous symmetry breaking

To resolve the mass generation problem, the spontaneous symmetry breaking (SSB) phenomenon is introduced. The SSB happens when the Lagrangian describing the dynamics of a physical system has a symmetry that is not preserved by the system ground states.

With the SSB, for a given gauge theory based on a local invariance with respect to a symmetry group G, given $H \subset G$ the symmetry group of the vacuum state, and being $dim(G) = N$ and $dim(H) = M$, $N - M$ massless Goldstone 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, as the simplest choice, the introduction of a doublet of complex fields, the Higgs field, of which three of the four degrees of freedom will be spent for the longitudinal polarization state of the massive bosons. The remaining degree of freedom is associated to the presence of the undetected Higgs particle, H_0.

The results of this theoretical environment is that the SSB 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.

The simplest SSB Lagrangian for the $SU(2)_L \times U(1)_Y$ can be written as:

$$\mathcal{L}_{SSB} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi); \tag{1.23}$$

$$V(\Phi) = -\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2; \quad \lambda > 0; \tag{1.24}$$

where $\Phi = \binom{\phi^+}{\phi^0}$ is a complex doublet with hypercharge $Y(\Phi) = 1$, and $V(\Phi)$ is the simplest renormalizable potential. For $(-\mu^2) < 0$, the minimum of the potential is reached on a circle of radius $v = \sqrt{\mu^2/\lambda}$ (see Fig. 1.1), and

$$|<0|\Phi|0>| = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}. \tag{1.25}$$

Consequently, in the lowest energy state the system spontaneously chooses one vacuum expectation value which no longer reflects the symmetry of the potential $V(\Phi)$. The physical spectrum is then realized by performing "small oscillations" around the vacuum state. By parameterizing $\Phi(x)$ as

$$\Phi(x) = \exp\left(i\frac{\vec{\xi}(x)\vec{\sigma}}{v}\right)\begin{pmatrix} 0 \\ (v + H(x))/\sqrt{2} \end{pmatrix}, \tag{1.26}$$

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

$$(D_\mu\Phi')^\dagger(D^\mu\Phi') = \frac{g^2v^2}{4}W_\mu^+W^{-\mu} + \frac{1}{2}\frac{(g^2+g'^2)v^2}{4}Z_\mu Z^\mu + \ldots \tag{1.27}$$

$$V(\Phi') = \frac{1}{2}2\mu^2H^2 + \ldots \tag{1.28}$$

$$\mathcal{L}_{YW} = \lambda_e\frac{v}{\sqrt{2}}\bar{e'_L}e'_R + \lambda_u\frac{v}{\sqrt{2}}\bar{u'_L}u'_R + \lambda_d\frac{v}{\sqrt{2}}\bar{d'_L}d'_R + \ldots \tag{1.29}$$

Explicitly, the tree level mass predictions for gauge and Higgs bosons read:

$$\begin{aligned} M_{W_\mu^\pm} &= \frac{gv}{2} \\ M_{Z_\mu} &= \frac{\sqrt{g^2+g'^2}v}{2} \\ M_{A_\mu} &= 0 \\ M_{Higgs} &= \sqrt{2\lambda}v, \end{aligned} \tag{1.30}$$

where

$$v = \sqrt{\frac{\mu^2}{\lambda}} \tag{1.31}$$

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

The mechanism here described, called the Higgs mechanism, gives mass terms for $W^{\pm}$, Z, as well as for quarks and leptons preserving the gauge invariance of the theory. On the other hand, it introduces a new scalar particle, not yet experimentally observed, the Higgs boson, whose mass and self-interaction are not theoretically determined.

1.2 Physics beyond the Standard Model

Despite all the efforts devoted to precision measurements aimed at probing the Standard Model of high energy physics, the question concerning what physics scenario to expect beyond the electroweak scale is still open. Several theoretical arguments support the idea that new physics is likely to be expected above the electroweak energy boundary and several models for extending the SM have been suggested.

In this context, the Standard Model of particle physics is incomplete and many open issues remain. The recently proved non zero masses of neutrinos call for an extension of the model. Another caveat, concerning the corrections to the Higgs mass, demands to be solved: 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. This issue is referred as to the hierarchy problem. Several ways of solving the hierarchy problem have been explored. For example, new strong dynamics could appear at the TeV scale (Technicolor theories). Another possibility allows the divergent corrections to the m_H to be cancelled by a new spectrum of particles at the electroweak scale: supersymmetric (SUSY) theories propose a supersymmetric partner, with different spin, to each SM particle. In this context, the hierarchy problem is solved considering radiative corrections from supersymmetric partners. To Standard Model fermions correspond bosonic superpartners (squarks and sleptons), while to Standard Model gauge bosons correspond fermionic superpartners (gluinos and gauginos). SUSY requires additional Higgs fields in order to provide mass to both up and down families. 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, a call for a Standard Model extension comes from the need to incorporate gravitational interactions in the theory.

1.3 The Top quark

1.3.1 Indirect evidence for the top quark

The discovery of the top quark[3, 4] in 1995 by the CDF and DØ collaborations at Fermilab marked the end of a long and difficult search. It was on the other hand another success of the Standard Model (SM) of particle physics which had strongly predicted its existence.

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

in its isospin multiplet.

In the following a short historical description of the experimental and theoretical hints for the top quark existence will be provided.

CP violation

In 1964 Christenson *et al.* at the Brookhaven National Laboratory observed rare decays of neutral kaons which violate CP symmetry [5].

In the light of this result, in 1973, Kobayashi and Maskawa added a phase factor, $e^{i\delta}$, into their quark mixing matrix in order to accommodate the observed CP violation into the model [6]. At that time, only three quarks (u, d, s), were known as mathematical entities. In their work, *in a framework of the renormalizable theory of weak interaction, problems of CP-violation [were] studied. It [was] concluded that no realistic models of CP-violation exist [...] without introducing any other new fields* (from [6]). The addition of the phase necessarily led to the proposal for three complete generations of quarks, since the smallest unitary matrix which can exhibit a non removable complex phase is 3×3 in size. At that time, even if CP violation had been experimentally proved, but not even the second family of fermions had been observed, and nothing guaranteed CP violation to be consequence of a phase in the mixing matrix, the concept of three fermion generations began to be kept in mind by scientists.

Observation of c-quark and of τ lepton

The mid-1970s marked two major discoveries in particle physics: in 1974, two experiments, at Brookhaven[7] and SLAC[8], independently observed a new resonance at 3.1 GeV/c^2, the J/ψ, which was immediately interpreted as a $c\bar{c}$ bound state. Furthermore, one year later, in 1975, M. L. Perl *et al.* at SLAC observed for the first time the τ lepton[9].

The discovery of the charm-quark completed the second generation of quarks, while the observation of the τ lepton strongly suggested the existence of a third lepton and quark generation.

The b-quark and its weak isospin measurement

It was in 1977 that the FNAL-E-0288 experiment collaboration at Fermilab discovered the b-quark ($\Upsilon = b\bar{b}$) [10]. The searches for a companion, the top quark, initiated immediately thereafter, based on the existence of the b and the empirically observed generation grouping of the quarks and leptons previously discovered.

Although there was no explanation for the proliferation of generations, the quark model suggested that within any family fermions must appear in left-handed doublets and right-handed singlets of weak isospin [11].

At the time, 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. The earliest

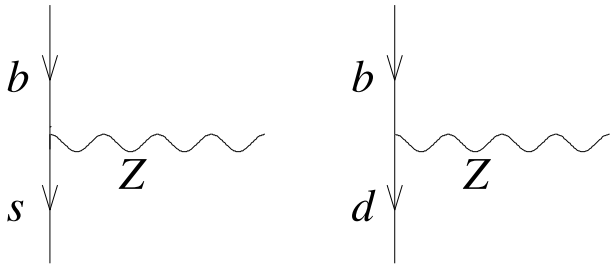


Figure 1.2: Flavor-changing neutral current vertices.

results on the b weak isospin were provided by the JADE collaboration at DESY[12] and indicated the b to be part of a weak isospin doublet. More recent efforts, concentrated on the forward-backward asymmetry, came from LEP experiments where the $e^+e^- \to Z \to b\bar{b}$ process was studied in detail. The asymmetry, $A_{FB}(b)$ is defined as follows:

$$A_{FB}(b) = \frac{\int_0^1 dz(d\sigma/dz) - \int_{-1}^0 dz(d\sigma/dz)}{\int_{-1}^1 dz(d\sigma/dz)} = \frac{N_{forward} - N_{backward}}{N_{forward} + N_{backward}}, \tag{1.32}$$

where $z = \cos\theta$, θ is the b-quark emerging angle with respect to e^- in the center-of-mass reference frame, and σ is the $e^+e^- \to b\bar{b}$ cross section at $\sqrt{s} = M_Z$. At the Z resonance, the forward-backward asymmetry described above can be expressed in terms of vector and axial-vector couplings:

$$A_{FB}(b) = \frac{3V_eA_eV_bA_b}{(V_e^2 + A_e^2)(V_b^2A_b^2)} \tag{1.33}$$

where for a given fermion, f, vector and axial-vector couplings can be written as:

$$V_f = T_f^3 - 2Q_f \sin^2\theta_W, \tag{1.34}$$

$$A_f = T_f^3. \tag{1.35}$$

In eq. 1.34 and 1.35, Q_f denotes the fermion electric charge and θ_W is the weak mixing angle. The expression in eq. 1.33 is proportional to T_b^3 and therefore in the case the b-quark was a singlet of weak isospin, $A_{FB}(b)$ would be equal to zero. On the contrary, in the case of an interpretation of the b-quark as a member of a doublet with respect to $SU(2)_L$, using the value $\theta_W = 0.23172$, the asymmetry would be $A_{FB}(b) = 0.10204$. The measurement from LEP collaborations found $A_{FB}(b) = 0.0941 \pm 0.0030$ [13], consistent with the b-quark being part of a doublet of weak isospin.

Absence of Flavor-Changing-Neutral-Current

In 1983, the observation of the $W^\pm$ and Z^0 gauge bosons by the UA1 and UA2 collaborations at CERN [14, 15] provided a direct evidence for both neutral and charged current weak interactions together with a strong confirmation of the Standard Model.

An important feature of the Standard Model is the absence of flavor-changing neutral currents, *i.e.* of processes like those illustrated in Fig. 1.2. By contrast, a theory with

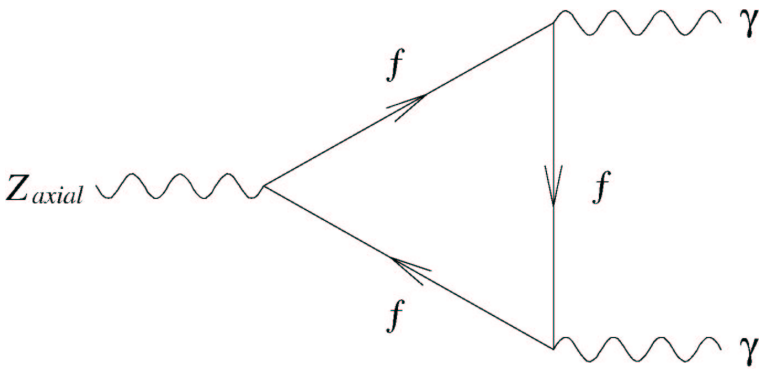


Figure 1.3: Example of a triangle loop.

only one third-generation quark predicts the existence of such kind of processes, which could be observed in B mesons decays. Kane and Peskin[16] in 1981 showed that if the b quark was a singlet of weak isospin, its decays via W and Z bosons would yield

$$\frac{\Gamma(B \to X l^+ l^-)}{\Gamma(B \to X l \nu)} \geq 0.12, \tag{1.36}$$

implying the branching ratio $BR(B \to X l^+ l^-) > 0.013$. This relation and consequently all theoretical models with five quarks were ruled out by experimental limits on the $B \to X l^+ l^-$ processes: in 1987 the CLEO collaboration indeed measured $BR(B \to X l^+ l^-) < 0.0012$ at the 90% confidence level [17].

Absence of triangle anomalies

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 spoil the renormalizability of the Standard Model (see Fig.1.3). 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:

$$\sum_L Q = 0. \tag{1.37}$$

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

$$\sum_L Q = -1 + 3 \times \left[\left(+\frac{2}{3} \right) + \left(-\frac{1}{3} \right) \right] = 0. \tag{1.38}$$

The absence of the top quark in the third generation would violate the condition of eq. 1.37.

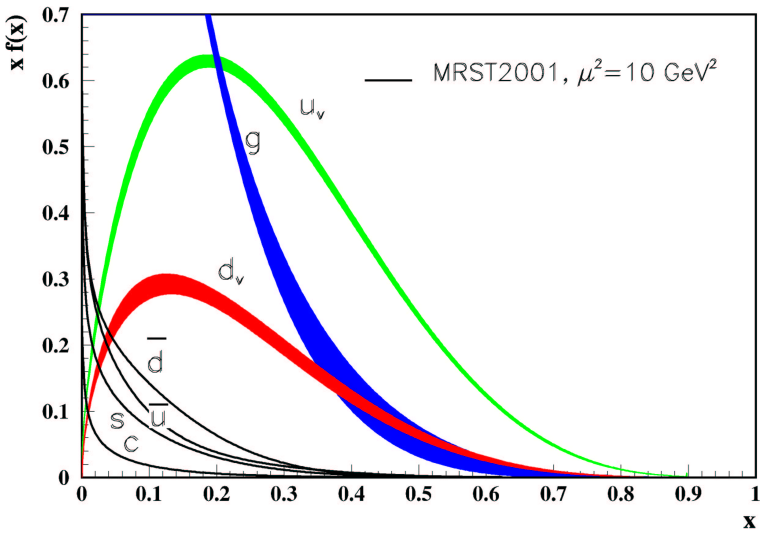


Figure 1.4: Parton distribution functions of quarks and gluons in the proton at a momentum transfer $\mu^2 = 10\ GeV^2$[18].

1.3.2 Top quark production

At hadron colliders, such as the Fermilab Tevatron, top quarks are produced mainly in pairs through strong interactions. Given that protons and antiprotons are not elementary particles, but are composed of quarks and gluons, the initial state of the actual hard scattering interaction is complicated and subject to non-perturbative effects. Thanks to the asymptotic freedom property of QCD, however, if the momenta of the initial particles are high enough ($\gg \Lambda_{QCD} \sim 200\ MeV$), it is possible to approximately 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, however, is not accessible in a event-by-event way: 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[18], where P the momentum of the proton (Fig. 1.4).

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. In order to theoretically evaluate the $t\bar{t}$ production cross section $\sigma_{t\bar{t}}$, one must sum over all the possible interactions, weighted by their probability specified by the PDF's. In the case of proton-antiproton collisions:

$$\sigma_{t\bar{t}} = \sum_{a,b} \int dx_a dx_b f_a^p(x_a, \mu^2) f_b^{\bar{p}}(x_b, \mu^2) \hat{\sigma}(ab \to t\bar{t}; \hat{s}, \mu^2, m_{top}), \tag{1.39}$$

where the summation indices run over light quarks and gluons contained in the initial proton and antiproton carrying momentum x_a and x_b of the initial hadron respectively; f_a^p and $f_b^{\bar{p}}$ are the parton distribution functions for proton and antiproton respectively, and $\hat{\sigma}$ is the parton-parton cross section. The center-of-mass energy of the $a-b$ parton

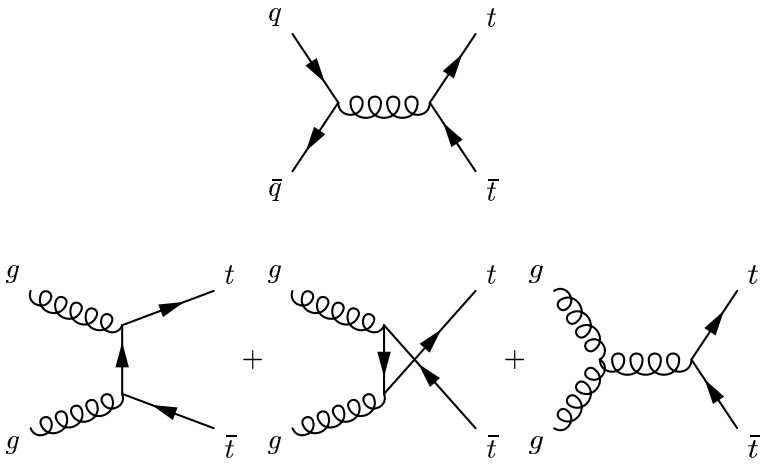


Figure 1.5: Leading order top pair production diagrams.

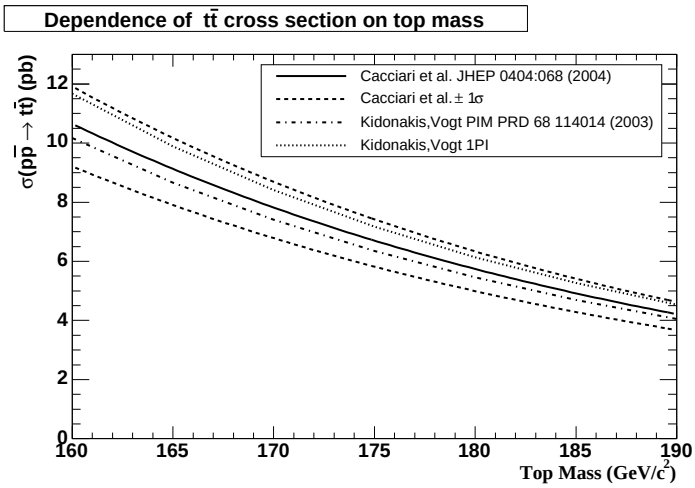


Figure 1.6: Dependence of top pair production cross section on top quark mass.

system is given by $\hat{s}$ and it is related to the $p\bar{p}$ center-of-mass energy by the relation: $\hat{s} = x_a x_b s$. The parameter μ is a factorization scale which is introduced to include resultant contributions from higher order Feynman diagrams.

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. 1.5.

The theoretical Standard Model prediction for $t\bar{t}$ production, at $\sqrt{s} = 1.96\ TeV$, is $\sigma_{t\bar{t}} = 6.7^{+0.7}_{-0.9}\ pb$ for a top mass of 175 GeV/c^2[20, 21]; $\sigma_{t\bar{t}} = 6.1^{+0.7}_{-0.9}\ pb$ for a top mass of 178 GeV/c^2. In Fig. 1.6 the theoretical cross section dependence on the top mass is shown for the Tevatron center-of-mass energy.

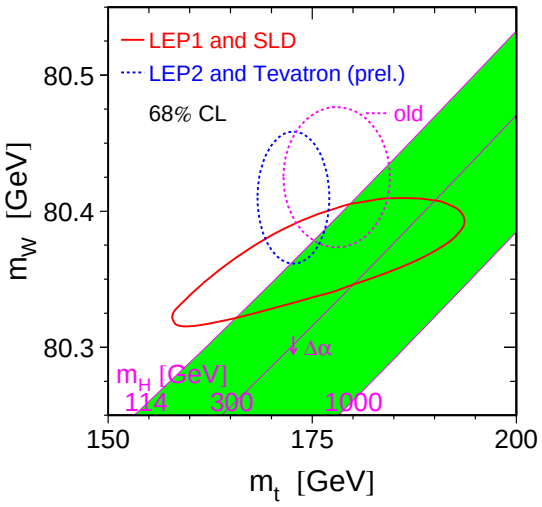


Figure 1.7: Limits on Higgs boson mass from direct top quark and W boson mass measurements, and indirect contraints from electroweak precision tests. Results are from CDF, DØ, LEP and SLD.

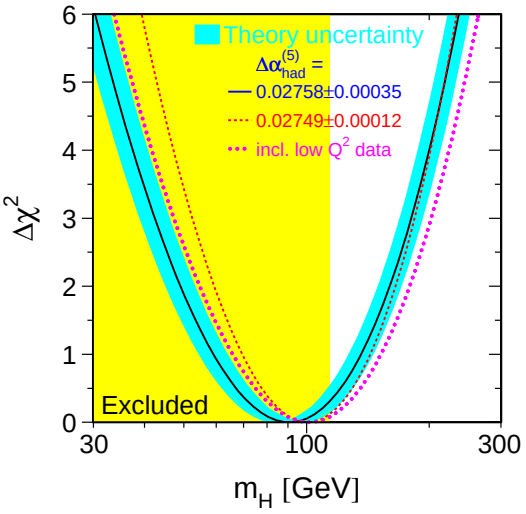


Figure 1.8: $\Delta\chi^2$ curve derived from high-Q^2 precision electroweak measurements, performed at LEP and by SLD, CDF, and DØ, as a function of the Higgs boson mass

1.3.3 Top quark mass

The top quark mass, M_{top}, is an important ingredient in Particle Physics. Its precise measurement not only serves for setting basic parameters in the calculation of the electroweak processes, but also provides a constraint on the mass of the Higgs boson, the last missing piece of the Standard Model. The current value of the top mass is set at $172.7 \pm 2.9\ GeV/c^2$ as a result of a combination of Tevatron Run I and Run II measurements[22], and it is known with an accuracy of 1.7%. Because of its large mass, the top quark causes large radiative corrections to electroweak processes whose effect can be observed in precision measurements[25].

As mentioned above, of particular interest is the constraint on the mass of the Higgs boson. Fig. 1.7 shows the limits on Higgs mass that can be derived from direct and indirect measurement of top quark and W masses. The dotted contour shows direct top and W mass measurements from Tevatron and LEP, for two values of the top mass $172.7 \pm 2.9\ GeV/c^2$, the current world average, and $178 \pm 4.3\ GeV/c^2$, the previous published world average (old)[23]. The solid contour shows the predicted top and W masses from electroweak measurement from LEP, SLD and neutral current experiments. The straight diagonal lines show the Standard Model predictions for three different Higgs masses. The consistency of the results shown in Fig. 1.7 favors a light Higgs boson mass.

On the other hand, Fig. 1.8 shows the $\Delta\chi^2$ curve derived from high-Q^2 precision electroweak measurements, performed at LEP and by SLD, CDF, and DØ, as a function of the Higgs boson mass, assuming the Standard Model to be the correct theory of nature.

The preferred value for its mass, corresponding to the minimum of the black curve, is at 91 GeV, with an experimental uncertainty of +45 and −32 GeV (at 68 percent confidence level derived from $\Delta\chi^2 = 1$ for the black line, thus not taking the theoretical uncertainty, shown by the blue band, into account).

While this is not a proof of the Standard Model Higgs boson existence, it does serve as a guideline in what mass range to look for it. The precision electroweak measurements tell us that the mass of the Standard-Model Higgs boson is lower than ∼186 GeV/c^2 (one-sided 95 percent confidence level upper limit derived from $\Delta\chi^2 = 2.7$ for the blue band, thus including both the experimental and the theoretical uncertainty). Direct searches, performed by LEP experiments, lead to the conclusion that m_H must be larger than 114.4 GeV/c^2 (95 percent confidence level limit), as indicated by the excluded area drawn in yellow in the figure.

1.3.4 Top decay width

The Standard Model predicts the top quark width to be of the order of 1.5 GeV, the value being dependent on its mass[26, 27]. This large width ($\Gamma_{top} \gg \Lambda_{QCD}$) causes the top quark to decay before hadronizing ($\tau_{top} \sim 4 \times 10^{-25} s$), 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.

1.3.5 Top quark decays

Within the Standard Model, the dominant decay signature of the top quark is $t \to W^+ b$ (or $\bar{t} \to W^- \bar{b}$). Decays to down and strange quarks are also allowed but are highly suppressed, being governed by off-diagonal elements of the Cabibbo-Kobayashi-Maskawa matrix[2] (CKM), as shown in Fig. 1.9.

For these reasons, it is natural to assume $BR(t \to Wb) \sim 1$ and consequently to classify the top pair decay signatures according to the W decay modes. The W bosons decay to either of the three generation fermions, $W^+ \to e^+\nu_e$, $W^+ \to \mu^+\nu_\mu$, $W^+ \to \tau^+\nu_\tau$, or into the lightest two generation of quarks: $W^+ \to u\bar{d}$, $W^+ \to c\bar{s}$. Because quarks come in three colors, each hadronic decay branching ratio is three times as large as each leptonic one[3].

The $t\bar{t}$ signatures are classified into three categories. The *di-lepton* category represents the case in which both W bosons decay leptonically; the *lepton+jets* signature on the other hand, arises when one of the W decays hadronically and the other into $l\nu$. Finally, the

[2] The CKM matrix arises because of the difference of mass and weak eigenstates for quarks. By convention, the matrix is often expressed in terms of a 3×3 unitary matrix operating on the charge $-e/3$ quark mass eigenstates (d, s and b):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \approx \begin{pmatrix} 0.9739 & 0.221 & 0.0029 \\ 0.021 & 0.9730 & 0.039 \\ 0.0048 & 0.037 & 0.999 \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.40)$$

[3] Neglecting the small effect due to fermion mass differences.

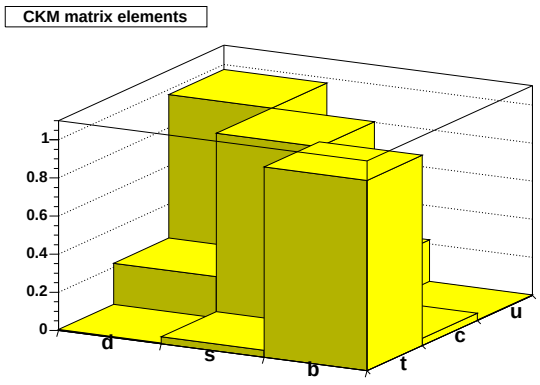


Figure 1.9: CKM matrix elements.

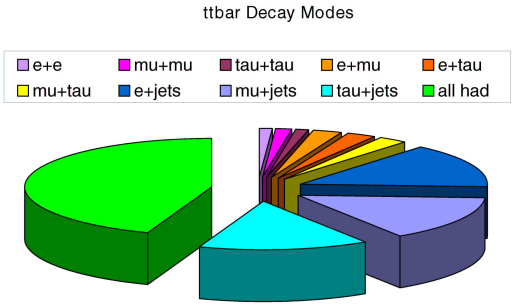


Figure 1.10: SM $t\bar{t}$ decay signatures.

Category	Decay Mode	Branching Ratio	
Di-lepton	$t\bar{t} \to e\nu b\ e\nu b$	1/81	1.2%
	$t\bar{t} \to \mu\nu b\ \mu\nu \bar{b}$	1/81	1.2%
	$t\bar{t} \to e\nu b\ \mu\nu \bar{b}$	2/81	2.5%
	$t\bar{t} \to e\nu b\ \tau\nu \bar{b}$	2/81	2.5%
	$t\bar{t} \to \mu\nu b\ \tau\nu \bar{b}$	2/81	2.5%
	$t\bar{t} \to \tau\nu b\ \tau\nu \bar{b}$	1/81	1.2%
Lepton+jets	$t\bar{t} \to q\bar{q}'b\ e\nu b$	12/81	14.8%
	$t\bar{t} \to q\bar{q}'b\ \mu\nu \bar{b}$	12/81	14.8%
	$t\bar{t} \to q\bar{q}'b\ e\tau \bar{b}$	12/81	14.8%
All-hadronic	$t\bar{t} \to q\bar{q}'b\ q\bar{q}'b$	36/81	44.4%

Table 1.3: Standard Model $t\bar{t}$ decay modes and their associated branching ratios.

all-hadronic channel corresponds to the case in which both W bosons decay into quarks. The possible $t\bar{t}$ decay modes and their corresponding branching ratios are summarized in Tab. 1.3 and Fig. 1.10.

1.3.6 The top quark beyond the Standard Model

The top quark mass is of the order of the electroweak scale, and is much larger than any other fermion mass in the SM. This is an indication that the top quark may play a special role in the SM and in the mechanism responsible for the electroweak symmetry breaking. The large top mass implies its Yukawa coupling to be of $\mathcal{O}(1)$, yielding the top quark to couple strongly to the Higgs sector.

As discussed above, the top quark decay branching ratios have not been measured to high precision so far. Hence it is not excluded that the top quark could decay to some exotic (non-SM) particle. In particular many models, such as supersymmetry, predict the existence of a charged Higgs boson ($H^{\pm}$) which could be lighter than the top quark. Current limits on $m_{H^{\pm}}$, in the framework of the Minimal Supersymmetric Standard Model

extension, set $m_{H^\pm} > 78.6\ GeV$ at the 95% confidence level[18], allowing the decay $t \to H^+ b$, where H^+ preferentially decays to $\tau\nu_\tau$ if the ratio of the two neutral Higgs fields vaccum expectation values, β, is small (in particular if $\tan\beta \gtrsim 5$ $BR(H^+ \to \tau\nu_\tau) \simeq 1$ [19]). Compared to the SM $t \to W^+ b$, this extra decay mode would enhance the presence of τ-leptons in the $t\bar{t}$ final state signature.

On the other hand, the supersymmetric partner of the top quark, the stop squark, $\tilde{t}$, could contribute either to the $t\bar{t}$ production or decay. If the $m_{\tilde{t}} < m_t$ the decay $t \to \tilde{t}\chi^0$, where χ^0 is a neutralino, is allowed. The stop squark would consequently decays either semileptonically ($\tilde{t} \to bl\nu$) or hadronically, together with the flavor changing $\tilde{t} \to c\chi^0$ mode, hence modifying the lepton content in the final state. Alternatively, if the stop is heavier than the top quark, it may be produced and decay into a top quark and a neutralino. In many theories, the lightest neutralino is stable, and would escape the detector without leaving measurable signals. Thus, exotic decays would result in a larger unbalance in the event momentum. Such contributions may significantly increase the apparent $t\bar{t}$ cross section.

Chapter 2

Accelerator complex: from H^- to $p\bar{p}$ collisions

The Tevatron[29] is a proton-antiproton synchrotron accelerator hosted at the Fermi National Accelerator Laboratory. To date it is the source of the highest energy $p\bar{p}$ collisions which occur at the center-of-mass energy of $\sqrt{s} = 1.96\ TeV$. The final proton and antiproton beams are the results of a complex acceleration apparatus which involves different stages, spanning from proton and antiproton production, their acceleration and transfer towards different sub-systems, to their actual collision in designed interaction points where particle detectors are installed. A schematic view of the Tevatron acceleration chain is provided in Fig. 2.1. In the following a description of the acceleration apparatus will be given.

2.1 The proton source

The process leading to $p\bar{p}$ collisions begins in a Cockroft-Walton chamber in which H^- gas is produced by hydrogen ionization. H^- ions are immediately accelerated by a positive voltage to an energy of 750 KeV and transported through a transfer line to the linear accelerator, the Linac.

The Linac[30, 31] picks up the H^- ions at energy of 750 KeV, and accelerates them up to the energy of 400 MeV.

The Booster[32] takes the 400 MeV negative hydrogen ions from Linac, strips the electrons off, which leaves only protons, and accelerates them 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 interspersed. The Booster loading scheme overlays the injected beam with the one already circulating in the machine. In fact, the H^- source and the overlaid beam injection were developed to increase beam intensity from the Booster: the idea exploits the fact that it is a proton beam that circulates in the Booster, while the Linac provides the incoming beam of negative H^- ions. Superimposing H^+ and H^- beams does not violate the *Liouville's Theorem*[1]. When this is accomplished, the mixed beams go through a carbon foil, which

[1] The density in phase space remains constant along a dynamic trajectory (in presence of conservative

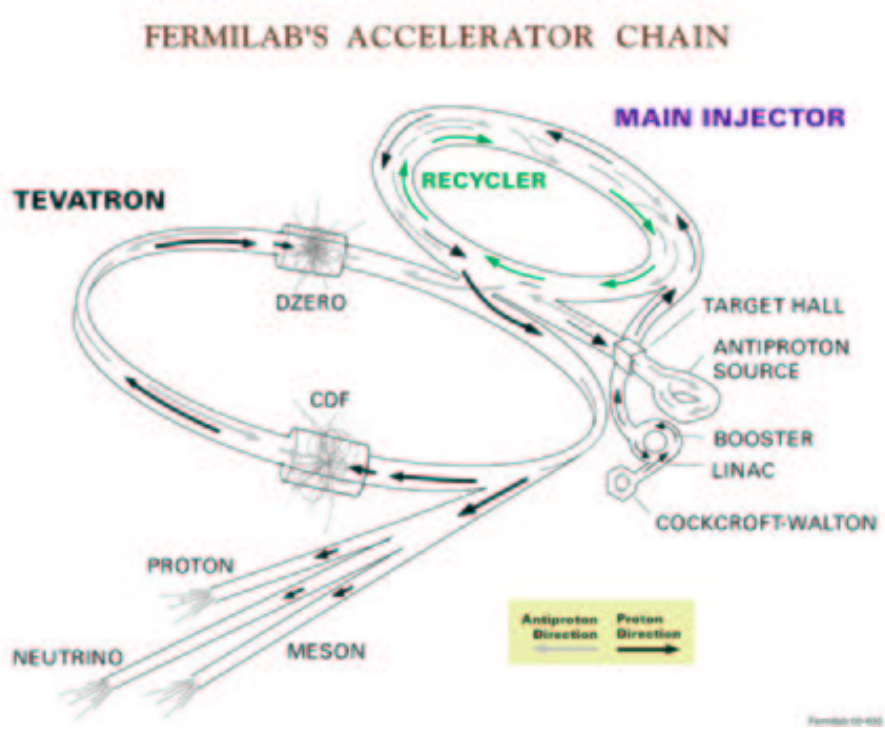


Figure 2.1: The FERMILAB's accelerator chain.

strips off the electrons turning the negative hydrogens into protons. In order to minimize the beam blow up due to stripping, the beams are forced to pass through the foil only during loading phase. 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 [29], 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.

2.2 The Main Injector

The Main Injector (MI)[33] is a circular synchrotron seven times the circumference of the Booster that plays a central 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 proton from the Booster, or 8 GeV antiprotons from the Recycler; it can accelerate protons up to 120 GeV for antiproton production or to 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.

forces), in particular two dynamic trajectories cannot merge.

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 (RF) systems it is accelerated and bunched.

In the following the Main Injector operational modes will be discussed.

2.2.1 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. Recently upgrades focused in doubling the number of protons on the target station have been completed. The new procedure, called "slip stacking" allows to merge two batches from the Booster before sending them to the target station (some details will be given in Sec. 2.5.2).

2.2.2 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.

As in the case of antiproton production, the procedure starts from 8 GeV protons from Booster, but in this case six full batches are loaded in the Main Injector in quick succession.

2.2.3 Collider operations

Collider Mode is the most complex scenario that the Main Injector has to cope with: in addition to supplying 120 GeV protons for antiproton production, the Main Injector must also feed the Tevatron protons and antiprotons at 150 GeV. To make matters more complicated, the protons and antiprotons need to be filled into super-bunches more intense than any individual bunch that can be accelerated by the Booster. A process called coalescing has been developed for this task; coalescing takes place at Main Injector *flat top* (*i.e.* the maximum energy at which the machine can keep the particles for an extended time).

The sequence of steps needed during a shot (the scientific term for loading protons and antiprotons) can be described as follows:

- One batch (84 bunches) of protons is accelerated to 8 GeV by the Booster;
- Only 7 of the batch bunches are extracted to the Main Injector to be then accelerated to 150 GeV.
- At flat top the bunches are coalesced, *i.e.* pushed together to form a narrow, high intensity bunch.
- The coalesced bunch is injected to the Tevatron.

- Previous steps are repeated until 36 coalesced bunches are sent to the Tevatron.
- Meanwhile, the 8 GeV antiprotons from the production target have been stored in the Accumulator, waiting to be injected in opposite direction with respect to the protons to the Main Injector.
- When this happens, antiprotons are accelerated to 150 GeV, coalesced and injected in the Tevatron in the opposite direction taken by protons.
- Main Injector drops back to 8 GeV for another group of antiproton bunches. The processes repeat until 36 antiproton bunches have been delivered to 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.

2.3 The antiproton source

The number of antiprotons available has always been an important limiting factor in producing the high luminosity desired for Tevatron physics. They are difficult, or at least time-consuming, to produce. The performance of the proton source, moreover, greatly affects the quality and duration of the physics run of the Tevatron. Anyway colliding protons and antiprotons, compared to protons and protons, has greater advantages even if it is more difficult to produce them in a sufficient quantity. In fact the antiprotons can be accelerated in the same ring used for protons, because of the opposite charge, thus reducing the cost of the magnets for a second ring. Moreover, the production rate for a number of interesting processes is higher in $p\bar{p}$ collisions at $\sqrt{s}$ up to 3 TeV compared with pp collisions at the same energy.

The Fermilab antiproton source[34] is comprised of a target station, two rings called the Debuncher and Accumulator, and the transfer lines between these rings and the Main Injector. In general, the following steps are taken in order to produce an antiproton beam suitable for collisions in the Tevatron.

- A single batch of protons with an intensity up to $4 - 5 \times 10^{12}$ is accelerated by the Main Injector at 120 GeV. Proton beam intensities up to 7×10^{12} have been recently achieved using slip stacking procedure (Sec. 2.5.2 and Fig. 2.8).
- After the extraction the proton beam proceeds to the target area where its spot size is reduced by means of quadrupole magnets and is made to collide with a nickel target producing showers of secondary particles. Immediately downstream of the target station is located the collection Lithium lens module, in which a solenoidal magnetic field focuses the negative secondaries. Lithium was chosen because it is the least-dense solid conductor which in turn minimizes particles scattering and absorption.
- A pulsed dipole magnet follows the lens. Its purpose is to select 8 GeV negative charged particles and to force them towards the Debuncher. Most of particles with

wrong charge-to-mass ratio are filtered out of the beam and collected by a graphite-core beam-dump.

- Surviving particles, typically 1 or 2 antiprotons for every 10^5 protons striking the target, are then injected in to the Debuncher where the momentum spread is reduced using stochastic and momentum cooling (see Sec. 2.3.2). The reduction of the momentum spread of incoming particles is needed in order to improve the Debuncher to Accumulator transfer efficiency, because of the limited momentum aperture of the Accumulator at injection. The Debuncher is a rounded triangular-shaped synchrotron with a mean radius of 90 meters.

- Just before the next pulse arrives from the target, the antiprotons are extracted from the Debuncher and injected to the Accumulator. As its name implies, the purpose of the Accumulator is to accumulate antiprotons. It is also a triangular-shaped synchrotron of radius 75 meters and is housed in the same tunnel as the Debuncher. It is the storage ring for the antiprotons; all of the collected antiprotons are stored here at 8 GeV and cooled until needed. Both RF and stochastic cooling systems are used in the momentum stacking process. The RF decelerates the recently injected pulses of antiprotons from the injection energy to the edge of the stack tail. The stack tail momentum cooling system sweeps the beam deposited by the RF away from the edge of the tail and decelerates it towards the dense portion of the stack, known as the core. Additional cooling systems keep the antiprotons in the core at the desired momentum and minimize the transverse beam size.

- When enough antiprotons have been accumulated in the Acculumator, their transfer starts. Antiproton beam destination can be either the Main Injector or the Recycler ring (Sec. 2.3.1).

Figure 2.2 shows the general layout of the antiproton source at the Tevatron. Figure 2.3 shows the antiproton production complex with the addition of the functionalities provided by the Recycler ring.

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, assuming there are no technical problems with the accelerator, motivating why a lot of effort was spent to upgrade the antiproton production at Tevatron for the Run II. In this context it is important to mention the Recycler ring [35, 36, 37]. The Recycler is designed to collect antiprotons left at the end of a collider *store*[2].

2.3.1 The Recycler ring

The Recycler[35, 36, 37] 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 solely by permanent gradient magnets and quadrupoles. Three main missions are designed for the Recycler operations: first and foremost, it allows antiprotons left over at the end of Tevatron Collider stores to be re-cooled and re-used; secondly, since the antiproton production rate decreases as

[2] A store is roughly a 20 hours period of time in which the colliding beams are retained in the Tevatron.

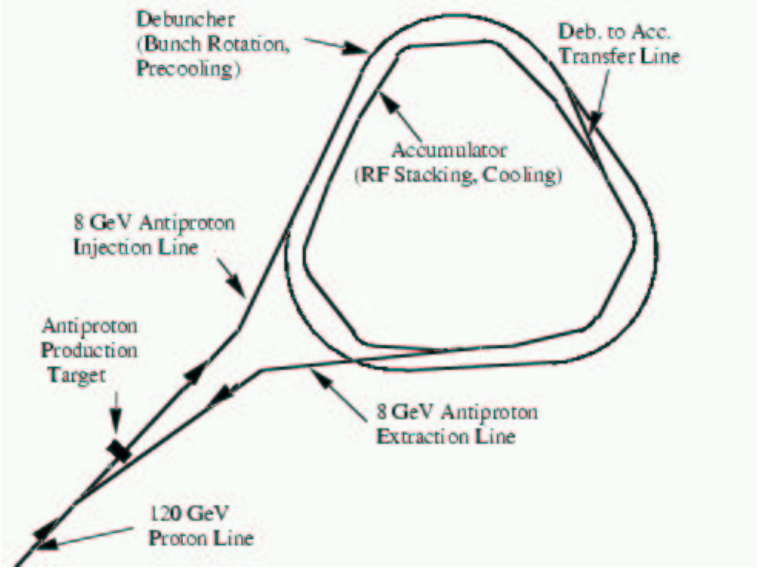


Figure 2.2: Layout of the antiproton source.

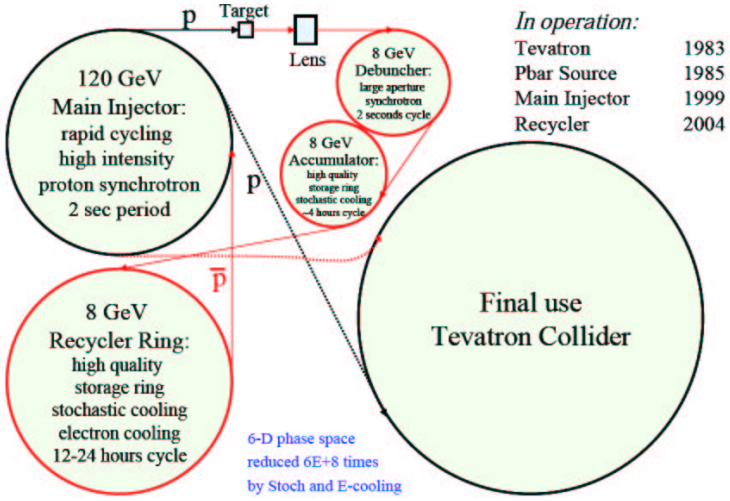


Figure 2.3: Antiproton production complex.

the beam current in the Accumulator ring rises, the Recycler is designed to act as a post-Accumulator cooler ring, allowing the Accumulator to operate optimally. Finally, permanent magnets were chosen in the construction of the Recycler construction in order 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.

Recycling the antiprotons left over after the end of Tevatron collisions is a bit involved, since the antiprotons are at 1 TeV. 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, scrapers, basically large chunks of metal, are slowly moved into the proton

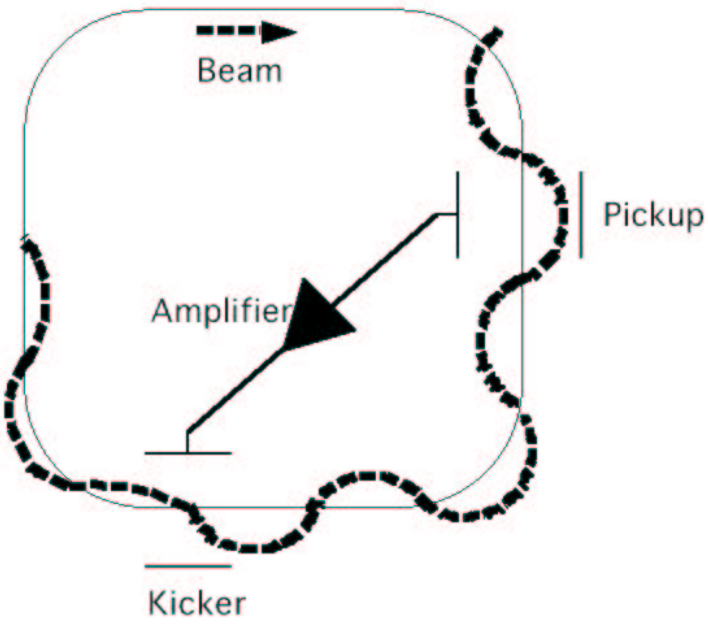


Figure 2.4: One particle model for a transverse stochastic system.

beam until only the antiprotons are left (the proton and antiproton orbits follow different paths in the Tevatron). At this point, antiprotons can be decelerated from 1 TeV to 150 GeV, using the Tevatron RF systems, coalesced and then transferred to the Main Injector. While the antiprotons are still circulating at 150 GeV, they are decoalesced; that is, 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.

The Recycler will also take up the role of the Accumulator as the final storage for 8 *GeV* antiprotons, allowing the existing Antiproton Source to perform more efficiently and to produce antiprotons with higher rate.

2.3.2 Antiproton cooling

Beam cooling[34] is a technique whereby the physical size and energy spread of a particle beam circulating in a storage ring is reduced without any accompanying beam loss. The goal is to compress the same number of particles into a beam of smaller size and energy spread, *i.e.* to increase the particle density. Phase space density can be used as a figure of merit for a particle beam, and cooling increases the density. In any cooling method no violation of the Liouville's theorem is implicit, given that phase space density is modified by the application of non-conservative forces.

The cooling electronics act on the beam through a feedback loop to alter the beam's momentum or transverse oscillations. Two types of beam cooling have been demonstrated and used at various laboratories: electron cooling, which was pioneered by G. I. Budker *et al.* at Novosibirsk, and stochastic cooling, developed by Simon van der Meer at CERN. Electron cooling gets its name from the fact that an electron beam is used to cool the

particles in question. Stochastic cooling is so named because of the stochastic nature of the corrections to the particle orbits.

Theoretically, electron cooling works on the principle of a heat exchanger. Two beams travel a certain distance parallel to each other: a "warm" beam of protons, antiprotons, or heavy ions with relatively large variation in transverse kinetic energy and a "cold" beam of electrons having much less variation in transverse kinetic energy. Both beams travel at approximately the same velocity and as the beams interact, the transverse kinetic energy of the warmer beam is transferred to the electron beam, which is then collected at the end of the cooling section. Electron cooling was developed for use in the Recycler Ring.

The Antiproton Source only employs stochastic cooling. The stochastic cooling systems used in the Antiproton Source are either betatron or momentum. Betatron and transverse all refer to systems that reduce betatron oscillations in the horizontal and vertical transverse planes.

The basic principles of the stochastic cooling technique can be described as follows: consider a single particle circulating in a storage ring as shown in the single particle model depicted in Fig. 2.4. Assume that the particle has been injected with some error in position and angle with respect to the ideal orbit (the center of the beam pipe). As the focusing system tries to restore the resultant deviation, the particle oscillates around the ideal orbit. These betatron oscillations can be approximated by a purely sinusoidal oscillations. The cooling system is designed to dump the amplitude of these oscillations. A pick-up electrode senses the position of the particle in the transverse plane, returning a signal with an amplitude proportional to the particle's deviation at the pick-up. This information will be applied to kicker electrodes located at $\pi/4$ from the pick-up, which deflect the particle by an angle proportional to its deviation at the pick-up point.

Momentum cooling systems reduce the longitudinal energy spread of a beam by accelerating or decelerating particles in the beam distribution towards a central momentum. In a momentum cooling system, the pick-up signals are used to provide kickers longitudinal fields to accelerate or decelerate the passing particles.

2.4 The Tevatron ring

The Tevatron 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 throughout, requiring cryogenic cooling and consequently a large scale production and distribution of liquid helium. During Run II the Tevatron operates at the 36×36 mode, which refers to the number of bunches in each beam.

The antiprotons are injected after the protons have already been loaded. Just before the antiproton injection a set of electrostatic separators are used to create a pair of non-intersecting helical closed orbits. When the Tevatron loading is complete, the beams are accelerated to the maximum energy and collisions begin. 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 the proton strikes an antiproton. The focusing, driven by

quadrupole magnets, reduces the beam spot size and thus increases the *luminosity*. The instantaneous luminosity, a quantity proportional to number of collisions per unit time, is given approximatively by:

$$\mathcal{L} = \frac{N_B N_{\bar{p}} N_p f}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)}, \tag{2.1}$$

where N_B is the number of bunches, $N_{(\bar{p}),p}$ is the number of the (anti)protons per bunch, f is the revolution frequency, and $\sigma^2_{(\bar{p})p}$ is the effective width of the (anti)proton beam. Clearly, the smaller $\sigma^2_{(\bar{p})p}$ the larger the rate of collisions.

For the data taking period considered in this work the instantaneous luminosity was in the range $0.1 \times 10^{32}\ cm^{-2}s^{-1}$ and $1 \times 10^{32}\ cm^{-2}s^{-1}$.

During collisions the instantaneous luminosity decreases in time as particles are lost and the beams begin to heat up. Meanwhile, new antiprotons are stored in the Accumulator. When the luminosity becomes too low (approximately after 15-20 hours) it becomes beneficial dumping the current store and start a new cycle. Table 2.1 summarizes the accelerator parameters for Run II.

Parameter	Value
Particles collided	$p\bar{p}$
Maximum beam energy	$0.980\ TeV$
Time between collisions	$0.396\ \mu s$
Crossing angle	$0\ \mu rad$
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 2.1: Accelerator parameters for Run II configuration [18].

Figure 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.

Figure 2.6, on the other hand shows the weekly and total integrated luminosity to date; while the total Tevatron luminosity delivered compared to the total luminosity recorded by the experiments is provided as a function of the store number in Fig. 2.7.

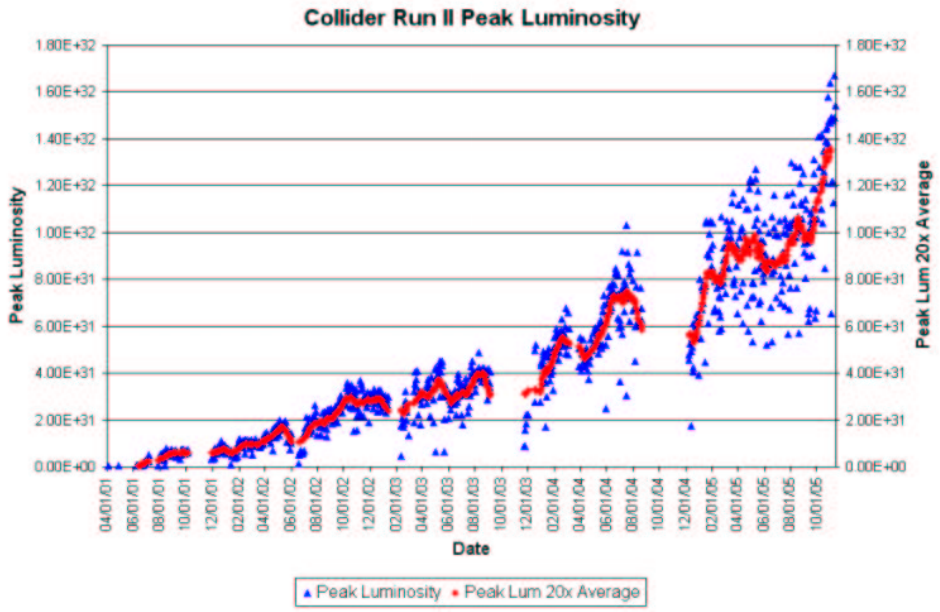


Figure 2.5: Run II peak luminosity.

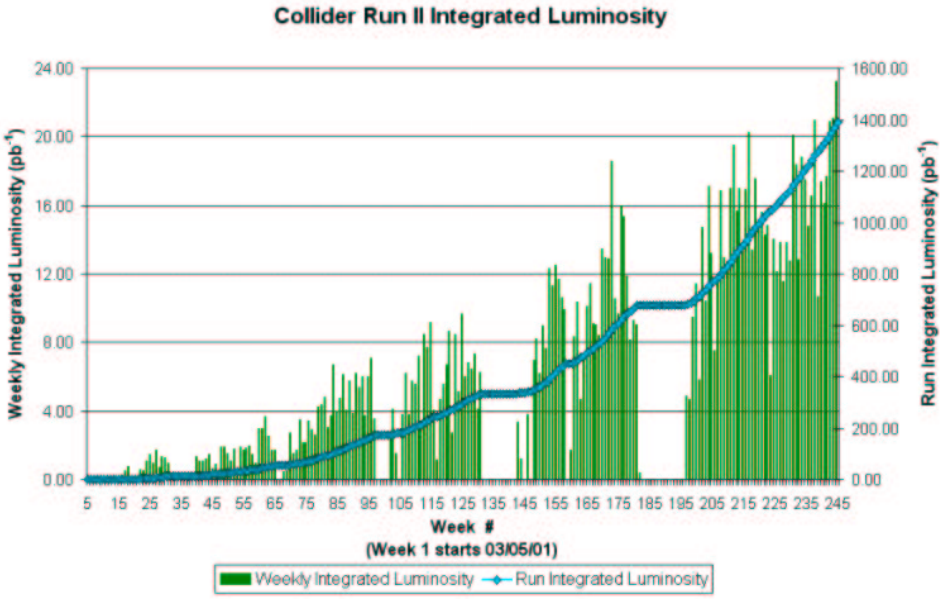


Figure 2.6: Weekly and total integrated luminosity.

2.5 Accelerator status and upgrades

The ever increasing peak luminosity observed during Tevatron shots since 2001 is the results of a complex campaign of operations, maintenance, upgrades, R&D and studies. This campaign (see [39]) was and is designed:

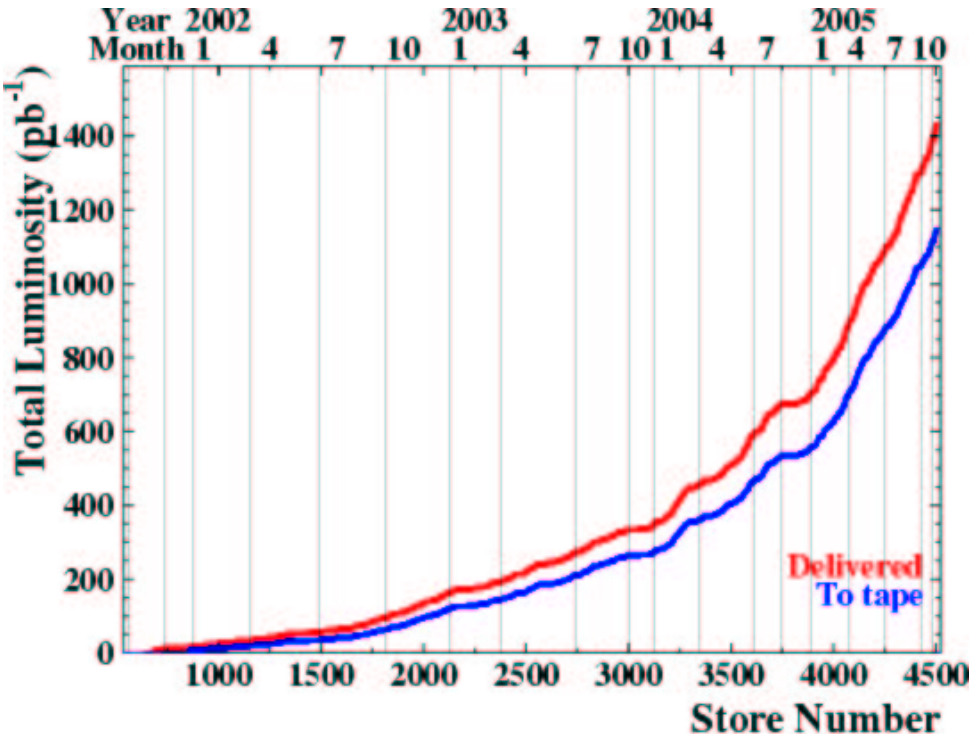


Figure 2.7: Integrated Run II luminosity.

1. to deliver increasing luminosity in the short-term, by improving the reliability and efficiency of operations and optimizing the operational parameters;

2. to implement and commission a program of upgrades to the accelerator complex to provide significant luminosity increases in the future.

The Run II upgrade program consists of two categories of sub projects: reliability upgrades and luminosity upgrades, and is closely integrated with ongoing operations. We will discuss these issues in the following sub-sections.

2.5.1 Reliability upgrades

The strategy for improving reliability is aimed at pursuing the causes of equipment failure either due to individual components or systems.

Voltage frequency converters

In the past, one of the leading causes of lost stores was a defect in voltage to frequency converters (VFC) used in the Tevatron quench protection system. All of the flawed VFC cards were replaced and the quench protection monitoring software was upgraded. The Tevatron RF power amplifiers were desensitized to power glitches. Upgrades to the beam abort logic in the Tevatron and development and implementation of a more sophisticated and robust abort system are being pursued.

Tevatron alignments and optics

Alignment and repairs of components in the accelerators, particularly in the Tevatron, have been undertaken as a major component of the upgrades. During the shutdown in the fall of 2003, an effort was mounted to re-align the components in the Tevatron. A new alignment network (TevNet) was implemented in the Tevatron to improve alignment precision and to simplify the process of alignment.

After the 2003 shutdown, dedicated studies were carried out to better understand and improve the optics in the Tevatron, to increase transfer efficiencies for proton and antiproton bunches for collider stores and to improve the helical orbits. This work resulted in the excellent and continually improving performance of the Tevatron.

2.5.2 Luminosity upgrades

The Run II luminosity upgrade program consists of a set of sub projects for upgrades throughout the accelerator complex to increase peak luminosity to about $2.8{\times}10^{32} cm^{-2} s^{-1}$. In the following, we provide a brief discussion of the instantaneous luminosity and the key parameters involved.

The integrated luminosity seen by an experiment depends on the peak (instantaneous) luminosity, the luminosity lifetime, and reliable operation. The instantaneous luminosity at each experiment is given by the formula:

$$\mathcal{L} = \frac{3\gamma f_0 B}{\beta^*} \frac{N_p}{\epsilon_p} \frac{N_{\bar{p}}}{1 + \frac{\epsilon_{\bar{p}}}{\epsilon_p}} H \tag{2.2}$$

where γ is the relativistic energy factor, f_0 is the revolution frequency, N_p and $N_{\bar{p}}$ are the numbers of protons and antiprotons per bunch and B is the number of bunches of each. β^* is the beta function at the center of the interaction region (which describes the beam divergence at the interaction point), and ϵ_p and $\epsilon_{\bar{p}}$ are the proton and antiproton 95 %, normalized, transverse emittances (which describe the beam longitudinal section). H is the hourglass factor, a function of the bunch length and β^*, and is less than unity.

While the luminosity depends on the transverse emittances explicitly and on the longitudinal emittance through the hourglass factor, the most direct way to increase the luminosity is to increase the proton and antiproton bunch intensities. The term N_p/ϵ_p (the proton brightness) is constrained by the maximum tolerable antiproton beam-beam tune shift that it causes. Therefore, the strategy for Run II upgrade is to increase the luminosity primarily by increasing the number of antiprotons and by using them efficiently in the collider operations.

The central strategy for the Run II luminosity upgrade is to increase antiproton production and stack size and to upgrade the Tevatron to handle high intensity bunches. In other words, the complex needs to produce, transport, cool and store more antiprotons and to use them efficiently in collider stores in the Tevatron. With stochastic cooling, the stacking rate decreases linearly as the antiproton stack size increases, so a key element of the upgrade plan was to implement electron cooling in a second storage ring, the Recycler, to allow stacking at high rates to very large stack sizes.

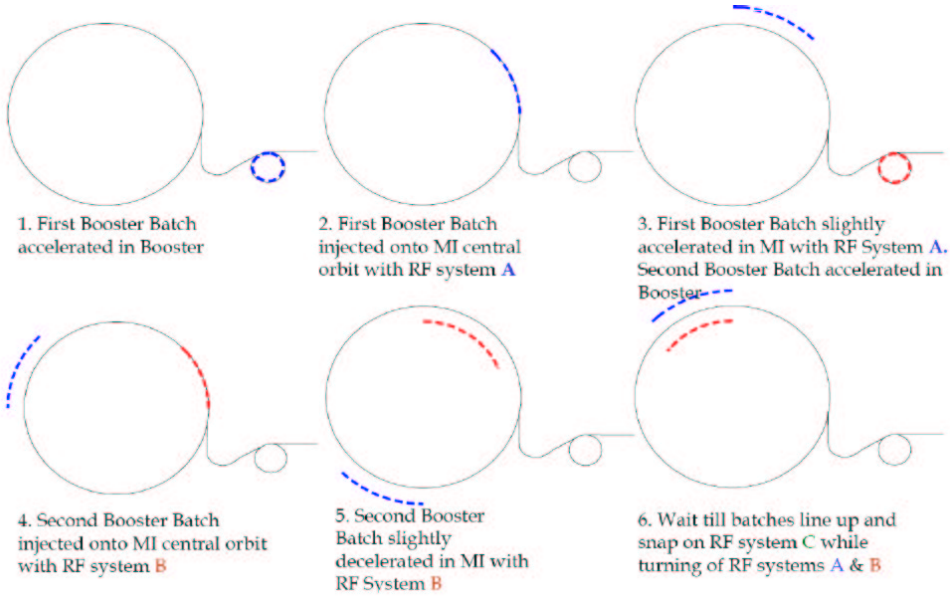


Figure 2.8: Slip stacking procedure.

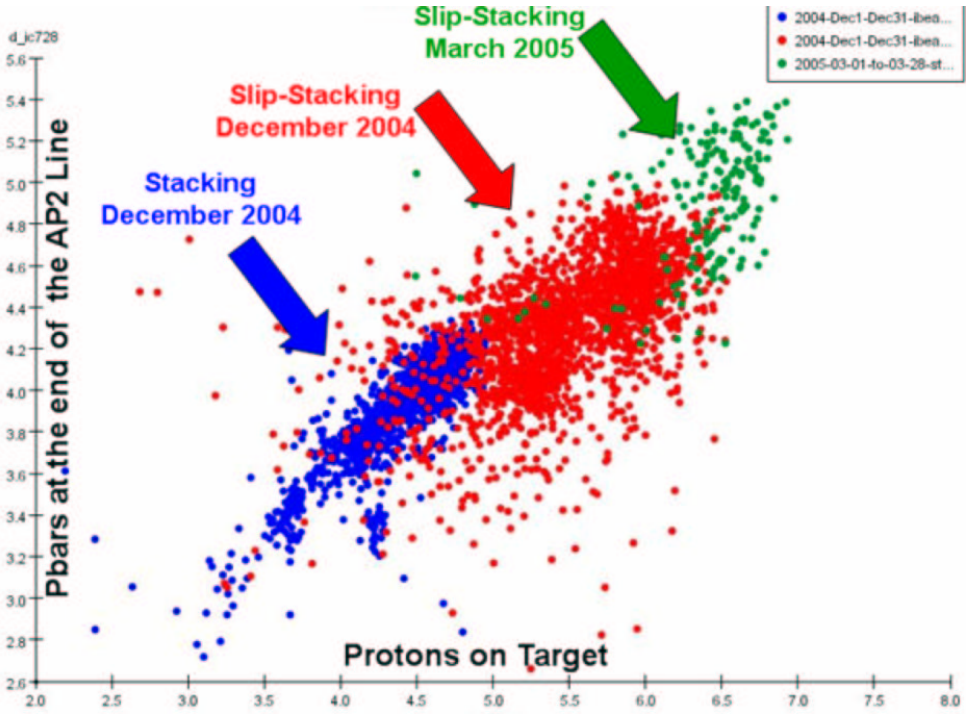


Figure 2.9: Performance of the antiproton production with slip stacking method adoption. Proton and antiprotons numbers are in 10^{12} and 10^{7}, respectively.

Increase the number of protons and antiprotons: slip stacking

The primary goal is to double the antiproton production rate by doubling the proton intensity on the production target and to make necessary upgrades to the target station

to take full advantage of the increased incident proton flux. The doubling of proton intensity for antiproton production is accomplished by a technique called *slip stacking*[28], which merges two batches of protons from the Booster. The slip stacking procedure is schematically pictured in Fig. 2.8. The design goal is to achieve sustained proton intensities of 8×10^{12} protons without degrading the beam quality. This requires upgrades to the Main Injector RF system to compensate for beam loading effects.

Slip stacking in the Main Injector has been tested for total proton intensities of the order of 7×10^{12} protons per pulse and with the specified bunch length on the antiproton production target. It was used for antiproton production and stacking since several days prior to the fall 2004 shutdown and a stacking rate increase of $\sim 15\%$ was observed. The increase in the number of protons on target observed by using the slip stacking is shown in Fig. 2.9.

To reduce damage to the antiproton target station due to increased flux delivered by slip stacking, new heat-resistant target materials (Stainless Steel targets with graphite cover) have been developed and installed. Additional protection can be provided if necessary by increasing the beam spot size by about 10%. A system for sweeping the beam spot across the target (and a system to compensate for it downstream) has been developed and retained as a back-up.

Increase the antiproton collection

This project is aimed at increasing the acceptance of antiprotons and focuses on the elements that are immediately downstream of the production target – the lithium collection lens, the transfer line, and the Debuncher ring.

The lithium lens is the first component downstream of the production target and has significant impact on the acceptance of the antiproton source. The number of antiprotons collected at the upstream end of the transfer line increases with the magnetic field gradient in the lens. A new lens design under development will allow operation at higher magnetic gradients, giving 10-15% increase in acceptance.

The transfer line transports the antiprotons from the lithium lens to the Debuncher ring where the antiproton beam is de-bunched and the momentum spread reduced before the antiprotons are stacked and further cooled in the Accumulator. The plan is to increase antiproton acceptance in both the transfer line and the Debuncher.

Increase the antiproton stacking and storing capabilities

This branch of the upgrade project encompasses stacking and cooling of antiprotons both in the antiproton source and the Recycler.

The flux capability of the Accumulator stack-tail stochastic cooling system is being upgraded. To sustain high rates for several hours and build up large stacks of 600×10^{10} antiprotons, partial stacks will be transferred every 30 minutes from the Accumulator to the Recycler. The antiprotons are stacked and stored in the Recycler until used in a collider shot.

The Recycler, which serves as the second antiproton storage ring, uses both stochastic and electron cooling systems to maintain large stacks with small transverse and longitudinal emittances.

Commissioning the Recycler, ready for installation and implementation of electron cooling, was one of the major milestones for Run II upgrades. This milestone was met, on schedule, on June 1, 2004. Electron cooling of the 8 GeV antiprotons in the Recycler required a high quality electron beam (small beam size, small angular spread) with a current of about 500 mA and kinetic energy of 4.3 MeV. The cooling of antiprotons by these electrons takes place in a 20 m section of the Recycler with ten solenoid modules and magnetic field correctors. An R&D program has been completed to produce an electron beam with the required properties, and the equipment was commissioned on July 2005. Since August 2005, electron cooling is used on every Tevatron shot.

The excellent performance of the Recycler itself and the transfers to and from the machine, has allowed the development of a mixed source mode for collider operation, where antiprotons are injected into the Tevatron from both the Accumulator and the Recycler. Mixed source operation produced record luminosity stores prior to the 2004 shutdown.

Recycler stand alone shots to the Tevatron, *i.e.* using only antiprotons stored in the Recycler to feel the Tevatron, were firstly operated on September 2005. The Recycler stack size could indeed reach 200^{10} antiprotons (the Recycler set a new record for Stored Antiprotons with a level of 247^{10} on September 28^{th}, 2005). The final transition to the Recycler-only shots is expected to be complete by the end of 2005 (Operations established the Recycler only store 4402 on September 26^{th} with an initial luminosity of $119.35 \times 10^{30} cm^{-2} s^{-1}$).

The next major steps for increased antiproton stacks are upgrading the stack-tail cooling system in the Accumulator and implementing rapid transfers of antiprotons between the Accumulator and Recycler (1-2 minutes per transfer).

The accumulator stack-tail cooling system will be upgraded in two stages to provide higher stacking rates. The changes to the stack-tail system require the large stacks no longer to be maintained in the Accumulator, but transferred in the Recycler. In the first stage, the existing cooling tanks will be reconfigured to increase the stacking rate for small stacks to $\sim 30 \times 10^{10}$ per hour. The stack will be transferred to the Recycler every 1-2 hours and stored and cooled with electron cooling.

After significant operational experience is gained with the higher stacking rates, with electron cooling in the Recycler, and with rapid transfers, the band-width of the stack-tail cooling system will be upgraded. This upgrade requires new pickups and kickers in the Accumulator, and is therefore not easily reversible. It is expected to allow zero stack stacking rates in excess of 45×10^{10} per hour, and is planned for the 2006 shutdown.

With the bandwidth upgrade it will be necessary to transfer a partial-stack of antiprotons from the Accumulator to the Recycler every half hour. A transfer time of 1-2 minutes (during which stacking in the antiproton source is interrupted) is to be achieved by automating the transfer process, which will require upgrades to the beam line instrumentation, improvements to the power supply regulation, and feedback from the Main Injector damper system. These upgrades are currently in progress.

2.5.3 Luminosity projections

Depending on the status and the perspectives on the Tevatron luminosity upgrades, projection on the total integrated luminosity achievable until 2009 can be derived. To do so, a parametric model, based on the simulation and the analysis of the actual performance parameters, as well as on the scheduled shutdowns, is used.

Current Tevatron performances and their extrapolation to 2009 are provided in Tab. 2.2. With the good status of the electron cooling upgrade, the major uncertainty left concerns the achievable antiproton stacking rate. For this reason, different luminosity scenarios can be considered, as described as follows:

- Design projection: includes electron cooling, a 15 minutes Accumulator to Recycler transfers, and peak stack rates of 30 mA/hr.
- Fall-back projection: includes electron cooling, 30 minutes Accumulator to Recycler transfers, and peak stack rates of $20 - 25$ mA/hr.
- Base projection: includes electron cooling, 45 minutes Accumulator to Recycler transfers, and peak stack rates of 15 mA/hr.

Luminosity projections, as a function of the achievable peak stack rate are illustrated in Figures 2.10, 2.11 and 2.12 in terms of weekly integrated, total integrated and peak luminosity, respectively.

Projections set as $4.2 \div 8.1$ fb^{-1} the Tevatron total integrated luminosity perspectives for 2009.

Parameter	now	2009 projection	units
Initial luminosity	110-140	270	$\times 10^{30} cm^{-2} s^{-1}$
Integrated luminosity	16	28	$pb^{-1}/week$
Total Int. luminosity	1.3	4.2-8.1	fb^{-1}
Protons/bunch	235	300	$\times 10^9$
Antiprotons/bunch	35-50	75	$\times 10^9$
Peak Pbar Accumulation rate	16.6	>20	$\times 10^{10}/hr$

Table 2.2: Present and projected accelerator complex performances.

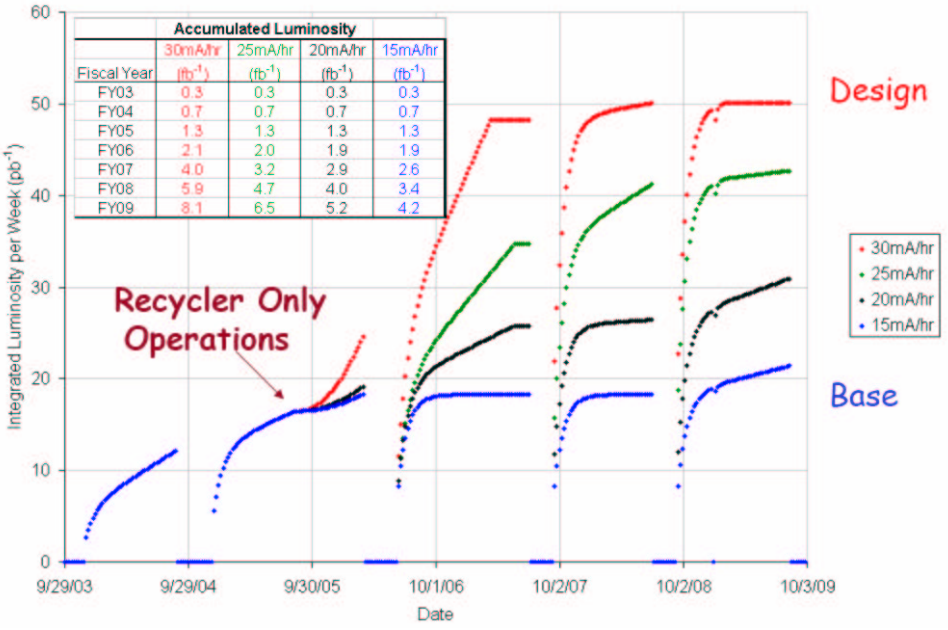


Figure 2.10: Run II integrated weekly luminosity projections.

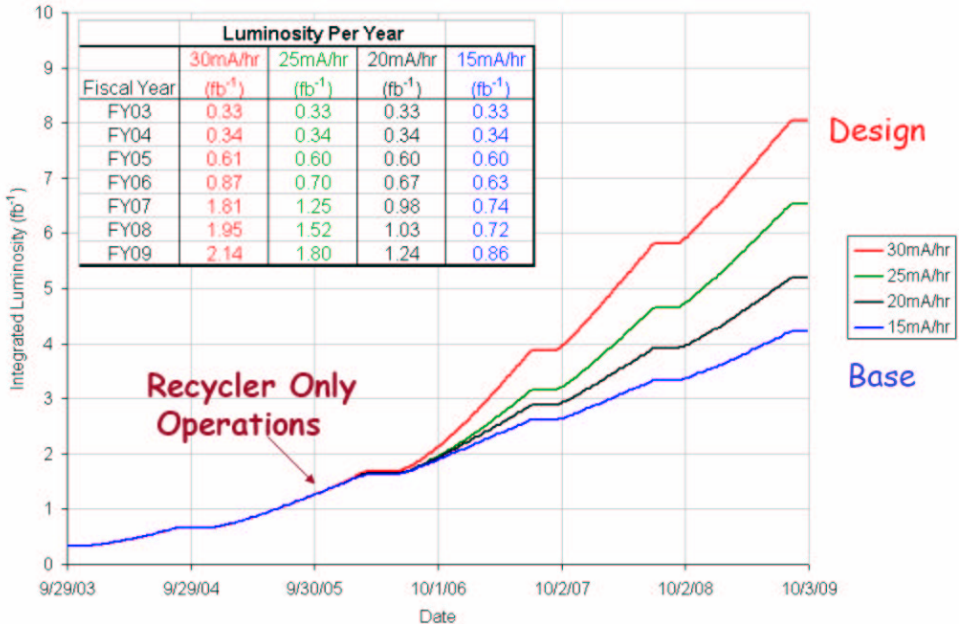


Figure 2.11: Run II total integrated luminosity projections.

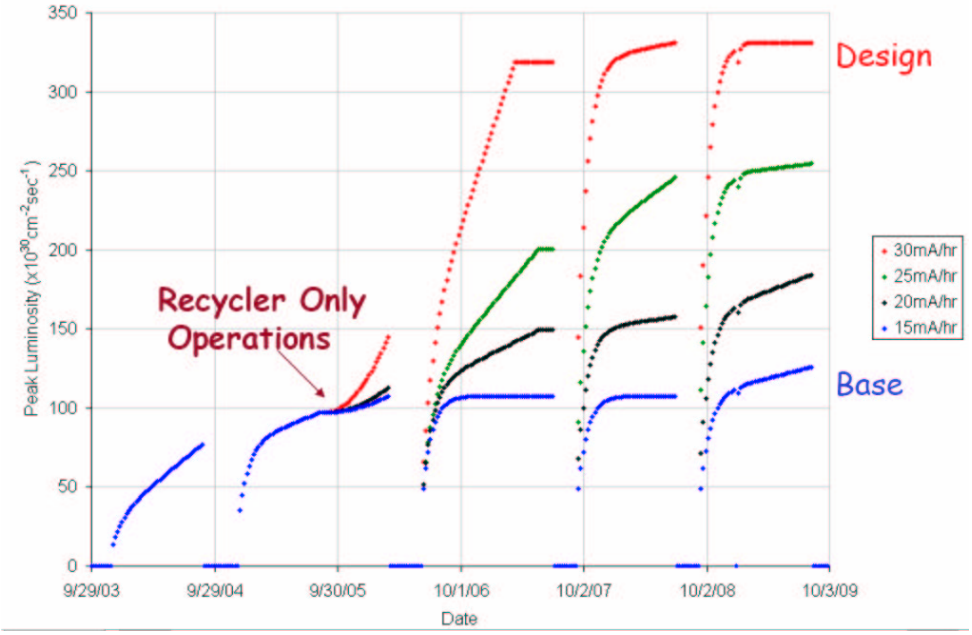


Figure 2.12: Run II peak luminosity projections.

Chapter 3

Detection of $p\bar{p}$ interactions at CDF

This Chapter is dedicated to a detailed description of the CDF detector used to study $p\bar{p}$ interactions provided by the Tevatron, whose sub-systems are crucial for the reconstruction of the physical objects this work will deal with.

Proton-antiproton interactions at the center of mass energy available at the Tevatron are interpreted in terms of collisions between the elementary partons constituting them. At this level, the phenomenology is usually described in the framework of a gauge theory based on the $SU(3)$ group, called *Quantum Chromo-Dynamics* (QCD). In such a theory, as already highlighted in Chapter 1, the interactions are carried by massless gauge bosons, the *gluons*, whose coupling to fermions is expressed in terms of three kinds of charge named *color* (by convention red, green and blue). The interaction potential grows linearly with the distance between the outgoing partons, until the creation of new color-singlet hadronic states becomes energetically favorable. At the end of this process, called *hadronization*, collimated *jets* of particles emerge, whose energies and directions carry a reminiscence of initial partons ones. The color radiation emitted from partons departing from the interaction center can lead to the creation of new jets or be included in the original one. In this case, the transverse expansion of the jet is limited by an effect known as *color coherence*: in the forward evolution, the angle between quark and emitted gluon decreases at every subsequent emission; in fact, the effects of interference between all the possible emission diagrams are destructive outside this region of angular ordering.

Aside QCD processes, electroweak production of W and Z bosons takes place in the collisions as well. For that reason, aside the detection of collimated spray of particles the capability to detect leptons and to establish the presence of neutrinos have to be considered in the design of a particle detector.

The *Collider Detector at Fermilab* (Fig. 3.1) [38] is located around one of the the two interaction points along the Tevatron ring (Fig. 2.1) and has been designed in order to perform precise measurement 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: this is a task of the calorimetry system, essentially based on the search of clusters of calorimetric towers

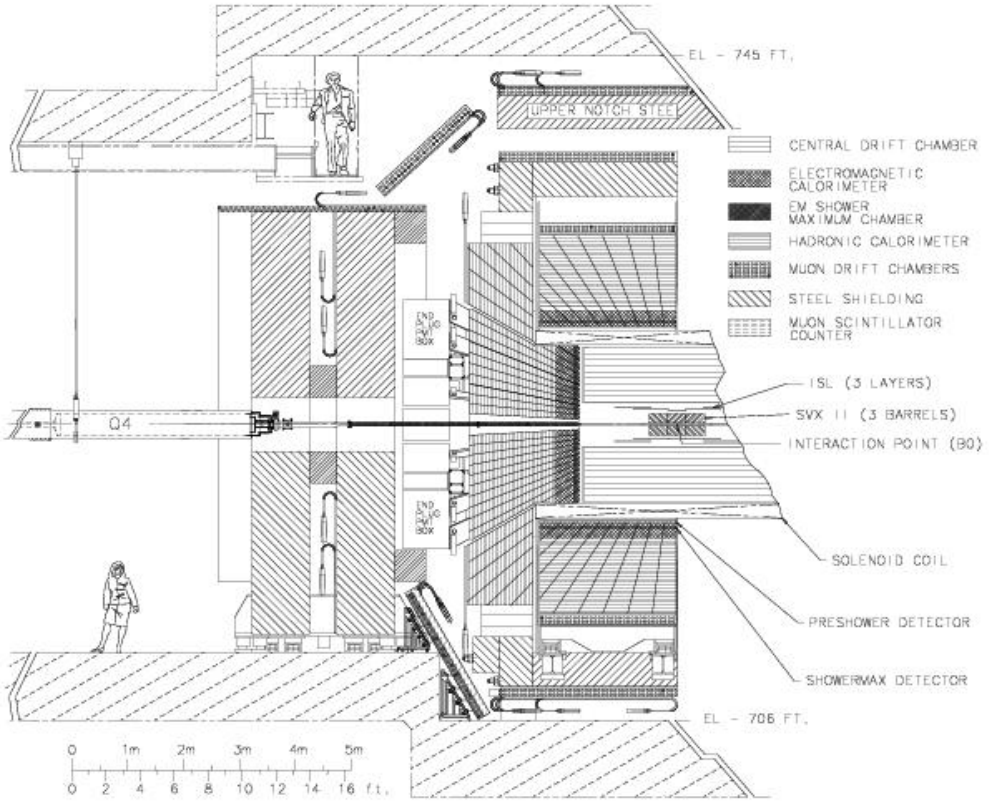


Figure 3.1: CDF II Detector Layout.

with energy above a certain threshold. In order to determine the direction of the jet momenta a precise measurement of the event interaction center is needed; besides, the identification of the jets originated by heavy quarks requires an accurate reconstruction of the secondary vertices produced in heavy flavor 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 of this Chapter, the detection and the reconstruction of the events at CDF II[1] will be examined in detail, with particular attention focused on those detector elements used in this work.

[1]We will refer to CDF II as the Collider Detector Fermilab upgraded complex installed for the Tevatron Run II.

3.1 Coordinate systems at CDF II

The description of CDF II subdetectors provided in the following sections 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 and completes the right-handed coordinate system.

The detector is essentially a cylindrically and forward-backward symmetric device. For this reason, 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 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 these boosts. For these reasons, the detector components (whenever appropriate) are chosen to be uniformly segmented along η and ϕ coordinates.

3.2 Tracking systems

Charged particles cause ionization passing through matter. This phenomenon is typically localized near the trajectory followed by the particle through the medium. Detecting ionization products give us 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 the CDF II is devoted to tracking systems, whose volume is permeated by an uniform magnetic field of magnitude $B = 1.4\,T$, oriented along the z-axis. This feature constrains charged particles to an helicoidal trajectory by means of the Lorentz force, whose radius, measured in the transverse plane ($x-y$) is directly related to the particles transverse momentum, P_T.

Particle trajectories can be completely described by five parameters [40]:

- 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;
- 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.

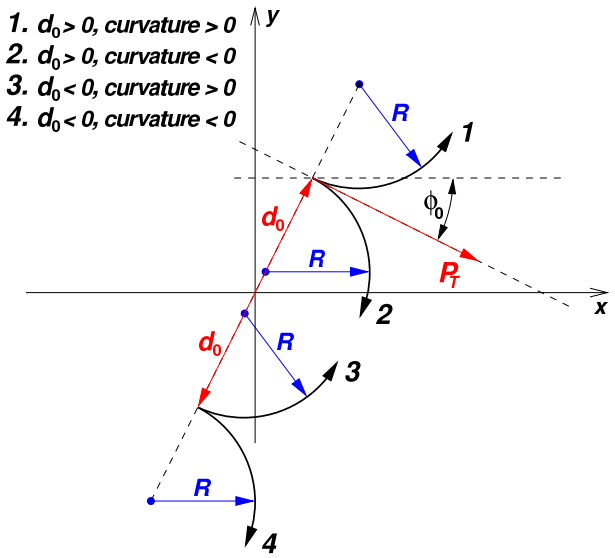


Figure 3.2: Illustration of helix track parametrization.

Actually, the impact parameter and the curvature are signed quantities defined by :

$$C = \frac{q}{2R}, \tag{3.2}$$

$$d_0 = q(\sqrt{x_c^2 + y_c^2} - R), \tag{3.3}$$

where q is the charge of the particle, $(x_c^2 + y_c^2)$ is the center of the helix as projected onto the x-y plane and R is its radius. A graphical view of these variables together with the ϕ_0 is shown in Figure 3.2.

From helix parameters one can easily derive particle transverse and longitudinal momenta as:

$$P_T = \frac{cB}{2|C|}, \tag{3.4}$$

$$P_z = P_T \cot\theta. \tag{3.5}$$

The CDF II tracking system is essentially divided in an inner silicon strip detector, and an outer drift chamber. The two sections provide precise vertex and momentum measurements respectively. Fig. 3.3 shows the overall CDF II tracking volume which covers a pseudorapidity range up to $|\eta| = 2$, and allow track reconstruction over a wide volume.

3.2.1 Silicon vertex detectors

The silicon vertex detector is crucial for precise determination of particle positions, and in particular its information can be used for the determination of secondary decay vertices produced by heavy flavor decays. This feature will be of central importance in this work.

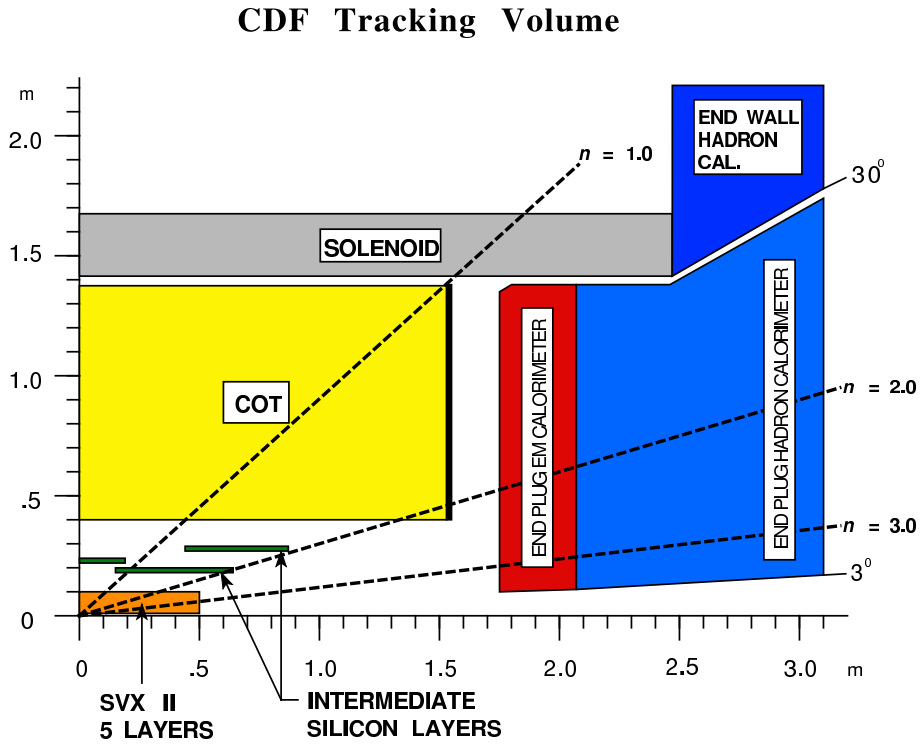


Figure 3.3: The r-z view of The CDF II tracking system.

Basic principles of silicon detectors

The basic principle on which silicon strip detectors work can be summarized as follow. A charged particle traveling across a silicon crystal produces electron-hole pairs.[2]

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. The band gap, E_g, is defined to be the minimum energy needed to excite an electron in the conduction band. In the case of silicon $E_g = 1.12\ eV$.

From another point of view, one of the most important characteristics of semiconductors is the fact that both positive and negative charge carriers exist. 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.

For a pure semiconductor the concentration of holes and electrons can be determined using thermal physics equations, but this ratio can also be manipulated by adding a certain amount of impurity atoms to the semiconductor crystals in a process called *doping*. By introducing impurities with a different number of valence electrons, the number of available charge carriers in the semiconductor can be increased. An important consequence of doping is the creation of intermediate energy levels in the forbidden region because of the excess charge carriers.

By introducing an impurity with five valence electrons (one more than those of silicon)

[2]For each electron-hole pair 3.6 eV are necessary in the silicon. This number should be compared to the 26 eV that are necessary to ionize an Argon atom in a common drift chamber.

such as arsenic or phosphorus, the semiconductor will have additional electrons for excitation into the conduction band. Electrons will then become the majority charge carriers in the semiconductor. These type of doped semiconductors are called n-type semiconductors. The extra electron is nearly free in the crystal and has an energy level that lies in the forbidden energy gap, just below the conduction band. This energy level is called the donor level and typically lies 0.05 eV below the conduction band in silicon.

On the other hand, by introducing an impurity with three valence electrons (one less than the normal four of silicon), such as gallium or indium, there will be a shortage of electrons available for covalent bonding in the crystal and an excess of "holes" in the valence level of the crystal. These are called p-type semiconductors and have holes as the majority charge carriers. These holes produce acceptor energy levels in the forbidden region which lie above the valence band (approximately 0.05 eV higher for silicon).

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*. The charge carriers will subsequently diffuse into the neighboring area until a thermal equilibrium is reached in which the thermal energy of an individual charge carrier is no longer large enough to excite it over the newly formed depletion region around the np-junction. The thermal equilibrium results in a depletion region surrounded by the remaining positive charge carriers in the p-type semiconductor and the remaining negative charge carriers in the n-type semiconductor. After the thermal equilibrium is reached, the n-type semi-conductor possesses a net positive charge and the p-type side possesses a net negative charge: before the two semiconductors were joined, each was independently neutral; after the junction, electrons in the n-type semiconductor are attracted to the positive holes in the p-type semiconductor and diffused over the junction. When the n-type semiconductor looses some of its negative charge carriers, it is left with a resultant positive charge. The same argument holds for the p-type semiconductor becoming negatively charged because it looses holes and gains electrons through diffusion. The depletion region is thereby depleted of mobile charge carriers as a thermal equilibrium is reached, and a electrical potential barrier is created between the n-type and p-type regions. At normal temperatures, the thermal energy of the semiconductors is sufficient to produce a depletion region around the np-junction. The size of the depletion region is typically 10 μm for the silicon and the potential through the junction is 0.6 eV for the silicon. However, the size of this region can be manipulated by adding external electric voltages to each side of the np-device.

For each μm of depletion region traversed by an ionizing particle typically 100 electron-hole pairs are produced. Their identification over the thermal background becomes easier as the size of the depletion region increases. For this reason, a reverse-bias is generally applied to the $n-p$ junction. To do so, electric voltages are applied such that a positive voltage is in contact with the n-type end of the device, and a negative voltage is placed in contact with the p-type semiconductor. When this electric field is set up, the positive voltage will attract the negative electrons from the n-type semiconductor, drawing them away from the depletion region. Conversely, the negative voltage will attract the positive holes away from the depletion region. These new forces of attraction result in an enlarge-

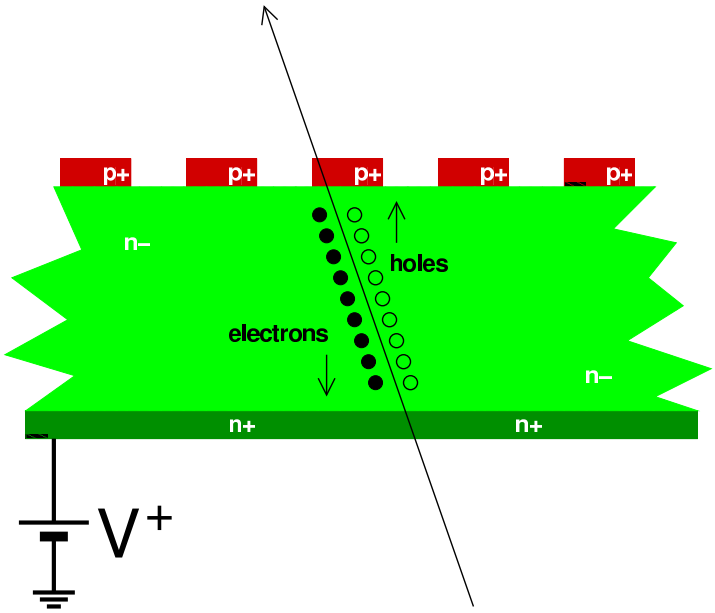


Figure 3.4: A generic silicon micro-strip detector.

ment of the depletion region and, consequently, of the energy gap between regions. This reverse-biasing results in a larger sensitivity for detecting the ionization signals produced by incoming charged particles.

Typical micro-strip silicon detectors

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^-), $\sim 300\ \mu m$ 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 and creates an electric field. With this setup, when a charged particle crosses the active volume (i.e. the n^- substrate), it creates a trail of electron-hole pairs from ionization. The presence of the electric field drifts the holes to the p^+ implanted strips producing a well localized signal. With additional effort the n^+ side can also be made in form of orthogonal electron-collecting strips and then used to measure the $r - z$ coordinate (assuming the p^+ side to measure the $r - \phi$ coordinates). 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. This is done by means of a weighting technique: the strip positions are weighted 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.

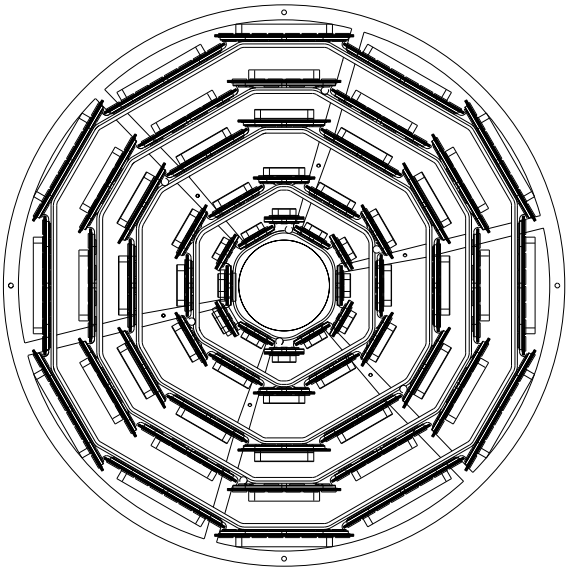


Figure 3.5: $r-\phi$ view of SVX II.

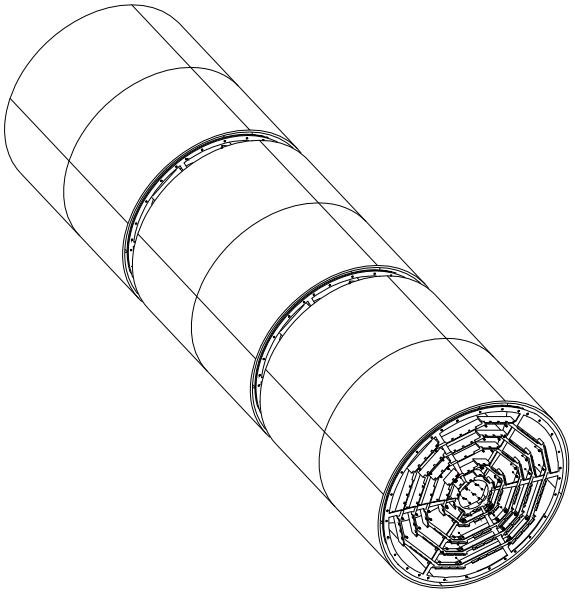


Figure 3.6: Perspective view of SVX II.

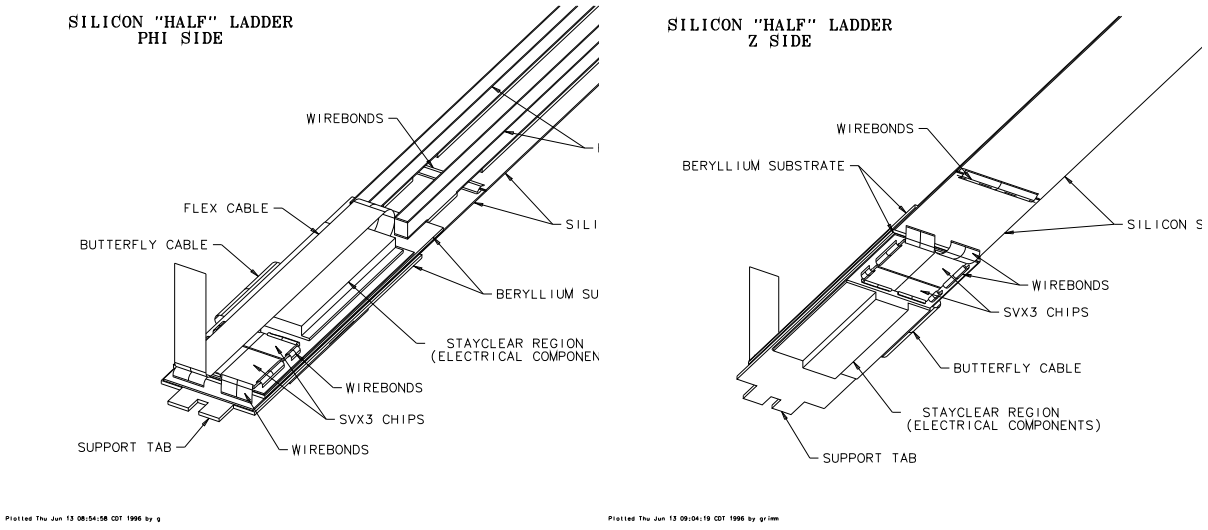
Layer	Radius $[cm]$		# of strips		Strip pitch $[\mu m]$		Stereo	Ladder Active, $[mm]$	
	stereo	r - phi	stereo	r - ϕ	stereo	r - ϕ	angle	width	length
0	2.55	3.00	256	256	60	141	90°	15.30	4×72.43
1	4.12	4.57	576	384	62	125.5	90°	23.75	4×72.43
2	6.52	7.02	640	640	60	60	+1.2°	38.34	4×72.43
3	8.22	8.72	512	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.

Constructing scheme of SVX II

The CDF II Silicon VerteX Detector is shown in Fig. 3.5 and 3.6, and it is known as SVX II[38]. Is composed of three different barrels each 29 *cm* long. Each barrel supports 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 combine an r - ϕ measurement on one side with the 90°stereo (r - z) on the other; while layers 2 and 4 combine an r - ϕ measurement with small angle stereo at 1.2°(Table 3.1).

The active silicon crystals (also called wafers or sensors) are supported by light-weight substrates (Rohacell foam) in assemblies called ladders, whose layout is depicted in Fig. 3.7. There are four wafers connected by wire-bonds in each ladder. Twelve ladders of appropriate width compose a layer. Sixty ladders are mounted between two beryllium bulkheads making an SVX II barrel.



(a) top *(r - ϕ)* (b) bottom *(r - z)*

Figure 3.7: Perspective view of the top $(r-\phi)$ and bottom $(r-z)$ side of Layer 0 ladder.

Each ladder is read out at both ends. The number of channels in the system (405,504) and the nature of the signals require much of the electronics to be physically mounted on the system as opposed to somewhere outside the detector: the amount of cables to be otherwise routed out would be impossible to manage, and on the other hand, some kind of pre-amplification is needed in order to avoid the signal to be lost in the thermal noise. The negative side is that a sizable amount of passive material is brought within the detector, increasing the effect of Coulomb scattering. Moreover, the built-in electronics dissipates significant amount of heat (over 1 kW), requiring an adequate cooling system to be incorporated into the bulkheads.

Intermediate Silicon Layer and Layer 00

To reach better performances in terms of resolution and tracking coverage two special sub-detectors are added to the silicon tracker: the Intermediate Silicon Layer (ISL) and Layer 00 (L00) [41].

- **ISL:** The Intermediate Silicon Layers (ISL) consists 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. 3.8, 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 constituted of two layers placed at $r = 20$ and 29 cm respectively (see Fig. 3.3).

  In both regions, the stereo sampling enables a full three-dimensional stand-alone silicon tracking.

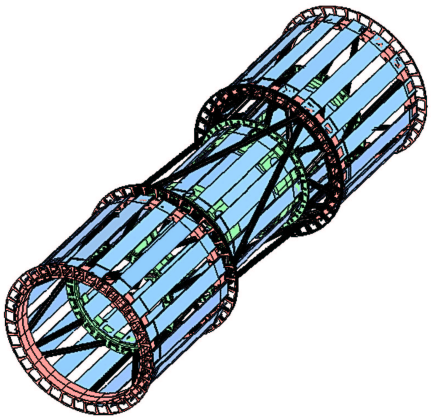


Figure 3.8: Perspective view of ISL.

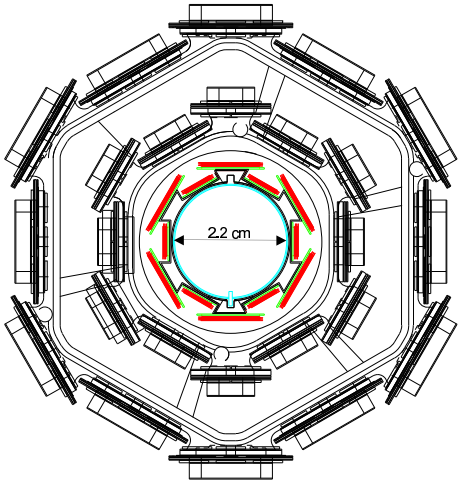


Figure 3.9: Transverse view of Layer 00, the innermost silicon layer.

- **L00:** L00 is composed of a set of silicon strips assembled directly onto the beam pipe (Fig. 3.9). 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 silicon (different from the ones used within SVX), which are single-sided with a 25(50) μm implant(readout) pitch.

The ISL is intended to improve the tracking resolution in the central region, while in the $1.0 < |\eta| < 2.0$ region it provides a useful tool for silicon stand alone tracking in conjunction with SVX layers. L00, on the other hand, allows to overcome the effects of multiple scattering for tracks passing through high density regions of SVX (where the bulkheads and readout electronics are located) making it possible to improve vertexing resolution.

3.2.2 The drift chamber

In addition to the silicon detector, the Central Outer Tracker (COT) is located at larger radii, and is used both to improve the momentum resolution and to provide useful information 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 [38] 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 Figure 3.10). 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

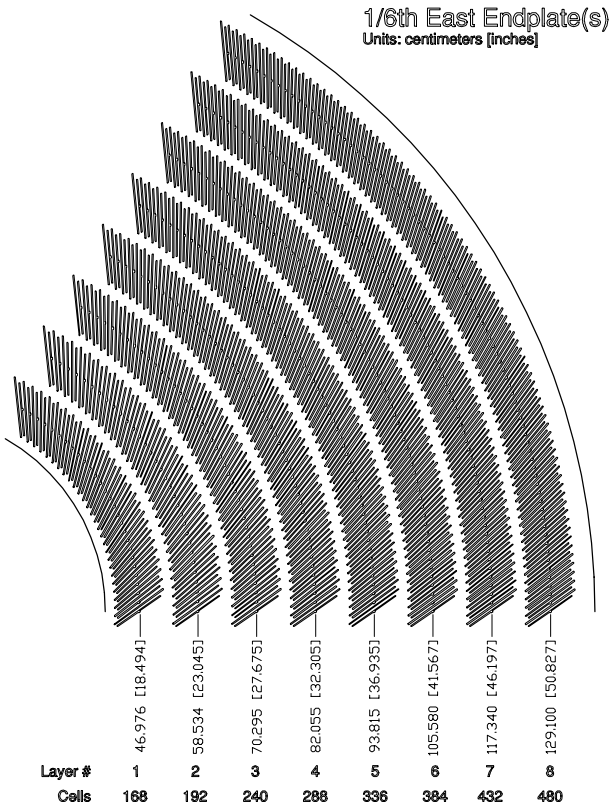


Figure 3.10: Layout of wire planes on COT endplate.

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. 3.11).

Resolution-wise the optimal situation 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. In brief an "avalanche" of ionization electrons is produced, creating a sizable signal (a hit) on the sense wire, which is subsequently amplified and shaped by the electronics attached at the end of the wire. The r - ϕ position of the track with respect to the sense wire is inferred by the signal arrival time.

There are other advantages related to the wire-plane tilt. For example, it allows wire-

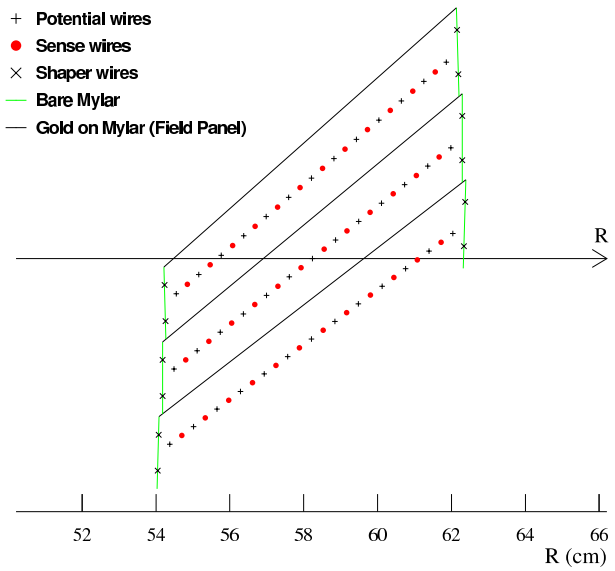


Figure 3.11: Transverse view of three cells from the COT superlayer 2. The arrow shows the radial direction or, equivalently, a very high p_T track trajectory. The electric field is roughly perpendicular to the field panels; the drift velocity would be vertical in this picture. The angle between wire-plane of the central cell and the radial direction is 35°.

planes to overlap in the radial view, with the result that high P_T tracks pass very close to at least one wire in each superlayer. This feature is useful for the eXtremely Fast Tracker (XFT) at the Level-1 trigger, as will be described in Section 3.6.1. Moreover, the large tilt provides a lever in resolving the left/right ambiguity: a particle passing through a wire plane leaves several hits, then associated to track segments, for which the reconstruction algorithm needs to decide whether they come from the left or right part with respect to the wire. This ambiguity creates two track segments, one being the mirror image of the other. The angle between the two track segments ($\tan^{-1}(2 \cdot \tan\alpha) \approx 54°$) is large, which renders pattern recognition to distinguish between them easier.

A measurement of COT performance is given by the single hit position resolution. It 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 Track reconstruction

The detectors described in the previous sections provide milestone information for the trajectory reconstruction of charged particles. 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.

Track reconstruction requires an excellent alignment between COT and silicon detectors. The global CDF II coordinate system is anchored to the center of the COT, i.e. the COT cylinder axis is the z axis of the global CDF II coordinate system, and the center of the COT is its origin. Positions of other detector components are measured[3] with re-

[3]Measurements are performed mechanically, by means of optical survey, or, when the ultimate precision

spect to the COT reference frame and encoded in so-called *alignments tables*. In general, whenever a spatial measurement is done by a system other than the COT, it is usually provided in the local reference frame and then converted into the global CDF II (w.r.t. COT) reference frame, via the corresponding alignment table. This alignment procedure underlies the track reconstruction procedure we are going to describe in the following.

Outside-In algorithm

The CDF standard track reconstruction is performed by the so called *Outside-In* algorithm [43]. 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 [44]. Matched segments and hits are them used to produce a track candidate. When the r - ϕ projection of the track is reconstructed, curvature (C), impact parameter (d_0), and azimuthal angle (ϕ_0) can be derived. The remaining track parameters can be calculated using hits in the stereo superlayers. Hits from stereo superlayers appear with an offset with respect the r - ϕ plane. The amount of the offset is proportional to the track z coordinate. Proper matching of stereo superlayer segments/hits with existing r - ϕ track projection allows to extract z information of the track at 4 radial regions, thereby providing a measurement of $\cot\theta$ and z_0 parameters of the track helix. The resulting track, because of the parameter uncertainties, is more like a tube of radius σ, the latter determined by the errors on the track parameters. As an additional complication, σ does not have to be necessarily the same in ϕ and z directions.

To complete track helix determination with the highest accuracy, SVX and COT information have to be combined. This is done extrapolating COT-only tracks through the silicon detector. As extrapolation proceeds from the outermost SVX layer toward the beampipe, 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[4] 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. At 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.

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*[45], has been developed. The algorithm essentially works in a reverse mode with respect to the Outside-In one: it uses silicon stand alone reconstructed

is needed and the possibility exists, with data.

[4] Often the radius is defined as some multiple of σ rather than an absolute number because σ changes as hits are added.

tracks to define a search road through the COT chamber. The hits in the road form a COT track that is fit using the silicon d_0 and z_0 constraints. Afterwards, silicon hits are refit using the new COT track as a seed. *Inside-Out* tracks can be used in conjunction to the standard *Outside-In* tracks to increases the CDF II 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.

The determination of the absolute momentum scale is usually obtained using clean and well known signals. In particular, the decay $J/\psi \to \mu^+\mu^-$ is used for the momentum calibration: the estimate of the momentum scale is performed by scaling the observed $J/\psi(1s)$ mass peak to the world average value of the $J/\psi(1s)$ mass.

3.2.4 Primary vertex reconstruction

Primary vertex reconstruction proceeds through an iterative algorithm that combines all the information on the reconstructed tracks[46]. The algorithm starts with an input seed (x_0, y_0, z_0) where x_0, y_0 are the x and y position of the run-average beamline and z_0 is the pre-determined highest sum-P_T vertex in the event. Then a P_T-weighted fit is performed and tracks with large impact parameter with respect to the candidate vertex are discarded. The procedure is repeated until stability is reached. The fitting algorithm can be based on x-y information only or it can take advantage of tracking upgrades and perform a three-dimensional reconstruction. 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 flavor 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.

3.3 Calorimetric systems

The calorimeter system, together with the muon and tracking systems, is one of the main sub-detectors of CDF II. 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. In order to supply information on particle positions, the calorimeter is also segmented in $\eta - \phi$ sections, called towers, projected towards the geometrical center of the detector. Each tower consists of alternating layers of passive ma-

	Thickness	number of layers	Resolution (E in GeV)
CEM	19 X_0	20-30 Pb: 3 mm, Scint.: 5 mm	$13.5\%\sqrt{E_T} \oplus 2\%$
PEM	21 X_0	22 Pb: 4.5 mm, Scint.: 4 mm + Preshower	$16\%\sqrt{E_T} \oplus 1\%$
CHA/WHA	4.7/4.5 λ_0	32/15 Fe: 25/50 mm, Scint.: 10 mm	$75\%/\sqrt{E_T} \oplus 3\%$
PHA	7 λ_0	23 Fe: 51 mm, Scint.: 6 mm	$80\%/\sqrt{E_T} \oplus 5\%$

Table 3.2: Basic quantities characterizing CDF II calorimetry.

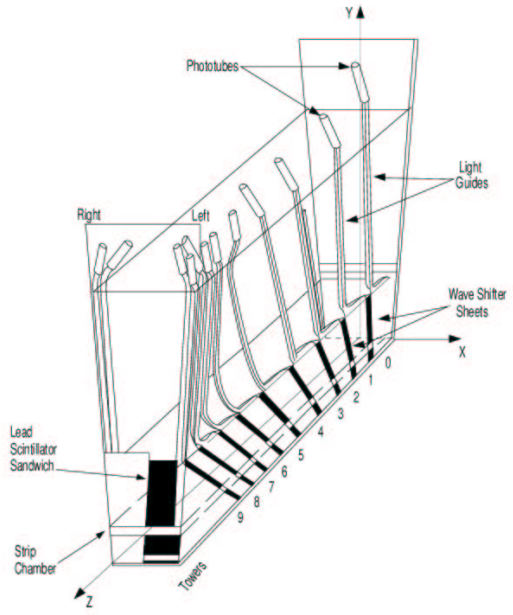


Figure 3.12: Perspective view of a CEM module.

terial 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 mechanically subdivided into three regions: central, wall and plug, in order of increasing pseudorapidity ranges. The naming convention is set as follows: Central Electromagnetic (CEM), Central Hadronic (CHA), Wall Hadronic (WHA), Plug Electromagnetic (PEM) and Plug Hadronic (PHA). In the following, a detailed description of the various calorimeter sectors will be provided, while Tab. 3.2 summarizes their most important characteristics.

3.3.1 The central calorimeter

The Central Electro-Magnetic calorimeter (CEM)[38] 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

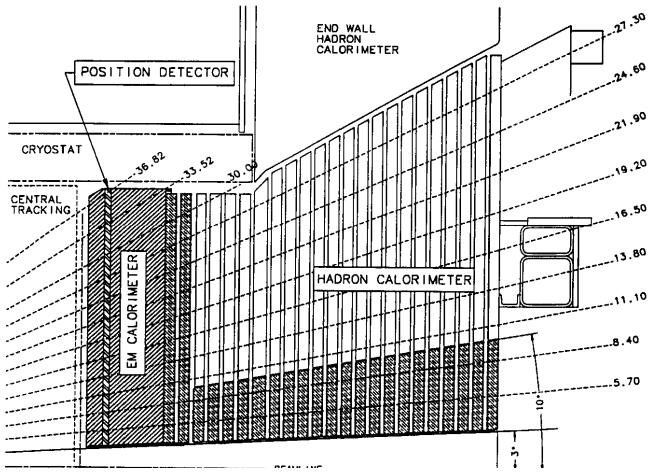


Figure 3.13: $r-z$ view of the plug calorimeter.

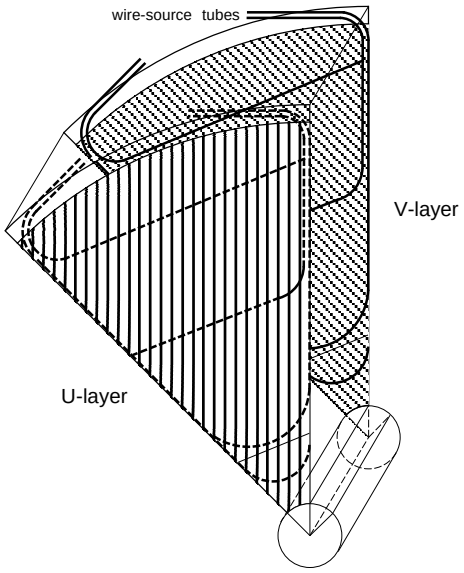


Figure 3.14: Perspective view of SMD.

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 3.12, where both the arrangement in projected towers and the light-gathering system are visible.

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

- The Central Electromagnetic Strip chamber (CES) is a two-dimensional strip/wire chamber located at the radial distance corresponding to the maximum shower development ($\sim 5.9 X_0$). 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 by using the tracker and the solenoid coil material as radiators, resulting to be a very useful tool for the rejection of background electrons and photons.

3.3.2 The plug calorimeter

The plug calorimeter, shown in Fig. 3.13, 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.

Projective towers consist of alternating layers of absorbing material (lead and iron for electromagnetic and hadronic sectors respectively) and scintillator tiles. The first layer of the electromagnetic calorimeter acts as a pre-shower detector; to this scope, the first scintillator tile is 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, at a radial depth of $\sim 6\ X_0$. The SMD consists of eight 45^o azimuthal sectors, each covering six (or three) calorimetric towers in ϕ. Each sector is segmented in two η regions in order to reduce the detector occupancy. Within each region, 5 mm pitch scintillating strips are arranged on two layers (called U and V). Being parallel to either boundary of the sector, U and V strips form an angle of 45^o among them (see Fig. 3.14), providing a two-dimensional measurement of the electromagnetic shower. As the CES, the SMD is an useful tool to discriminate photons and electrons from pion background.

3.3.3 Jet reconstruction

In general jets are the results of the fragmentation process of partons outcoming from $p\bar{p}$ collision. The fragmentation yields a stream of energetic colorless particles emitted spatially collimated 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 jet reconstruction algorithm at CDF uses the energy depositions in the calorimetric towers in a *fixed opening cone* procedure, in compliance with the directives of the 1990 Snowmass conference[47] for a common operational description of jets in $p\bar{p}$ collisions. 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 has to be chosen according with 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.

The reconstruction procedure, named *jet clustering*, starts by assigning a vector in the (r, η, ϕ)-space to each calorimetric tower, 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).

Calorimeter towers are grouped together in clusters of size R around towers (*seeds*) with $E_T > 1\ GeV$, where the tower energy is defined as the sum of both electromagnetic and hadronic energy deposits.

The 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. To be included in the cluster, a tower must have $E_T > 100\ MeV$. Once a list of clusters is formed the center of the cluster is calculated accordingly to the following definitions:

$$E_T^{jet} = \sum_{i=1}^{N} E_T^i \tag{3.6}$$

$$\eta^{jet} = \sum_{i=1}^{N} \frac{E_T^i \eta_i}{E_T^i} \tag{3.7}$$

$$\phi^{jet} = \sum_{i=1}^{N} \frac{E_T^i \phi_i}{E_T^i} \tag{3.8}$$

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[5]. It may happen that two jets overlap; in this case, 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, they 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 $(E_T^{raw}, P_x^{raw}, P_y^{raw}, P_z^{raw})$, is calculated using the final list of towers associated to the cluster:

$$E_{jet} = \sum_{i=1}^{N} E_i \tag{3.9}$$

$$P_{x,jet} = \sum_{i=1}^{N} E_i \sin\theta_i \cos\phi_i \tag{3.10}$$

$$P_{y,jet} = \sum_{i=1}^{N} E_i \sin\theta_i \sin\phi_i \tag{3.11}$$

$$P_{z,jet} = \sum_{i=1}^{N} E_i \cos\theta_i \tag{3.12}$$

$$P_{T,jet} = \sqrt{P_{x,jet}^2 + P_{y,jet}^2} \tag{3.13}$$

$$\phi_{jet} = \tan\frac{P_{y,jet}}{P_{x,jet}} \tag{3.14}$$

$$\sin\theta_{jet} = \frac{P_{T,jet}}{\sqrt{P_{x,jet}^2 + P_{y,jet}^2 + P_{z,jet}^2}} \tag{3.15}$$

$$E_{T,jet} = E_{jet} \sin\theta_{jet} \tag{3.16}$$

In general, jets with $E_T < 8\ GeV$ are not used in physics analyses. All the quantities previously defined are referred as *raw* since they are affected by mismeasurements for

[5]In the re-calculation of the jet direction, new towers may be added to the cluster list while others may finish out of the cluster cone, making the tower list to change.

a variety of reasons due both to physical and detector effects, as it will be clear in the following section (3.3.4).

Before treating the jet correction, the energy scale of calorimetric towers must be defined. For the Central Electro-Magnetic (CEM) calorimeter the absolute scale is set by reconstructing $Z \to e^+e^-$ decays, and imposing the reconstructed mass peak to be consistent with the world average mass of the Z boson. Moreover, the ratio of the calorimeter energy and track momentum measurements, E/p, for well-identified electrons is used to apply additional relative calibration for each tower in order to further improve the energy measurement resolution. For plug electromagnetic calorimeter (PEM) the absolute scale is set by looking at $Z \to e^+e^-$ events with one electron in the CEM and the other in the PEM. As far as the hadronic calorimeters scale is concerned, the calibrations derived from a test beam with 50 GeV charge pions are used. The definition of energy scale of the calorimeter is often referred as the Level 0 correction to jets.

3.3.4 Jet energy corrections

According to CDF definition, jets emerge from a vector sum of the momenta associated to calorimeter towers enclosed in a certain region (cone), under the assumption that all tower contributions come from massless particles originated in the primary vertex which are detected in the energy barycenter of each tower. However, a lot of factors contribute to a mis-measurement of the real parton energies.

- Due to the transversal dimensions of the physical jet, some particles can fall outside the cone of the reconstructed jet yielding an under-estimated energy measurement (*out-of-cone energy*).
- Very penetrating particles can be present in the jet, such as muons whose energy is not completely detected, or neutrinos which escape from the calorimeter.
- Contribution of low-P_T charged particles can also be lost due to the deflection they experiment within the magnetic field which permeate the tracking systems.
- Imperfect calorimeter coverage and un-instrumented detector regions (so-called *cracks*), in particular at boundaries between the central and plug calorimeters, also 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*) introduce a flux of soft hadrons that can interfere with the jet definition, since some particles generated in the underlying event can be included in the clustering procedure. The same phenomenon can happen when more events occur during the same bunch crossing (*multiple interactions*).

In order to take in to account all these effects, specific procedures have been studied [48, 49]. The generic[6] jet energy scale corrections consist of many steps which correct for

[6] This means that the flavor of the hadronizing quark yielding the jet is not taken into account.

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 3.3: Naming convention for the different jet corrections.

detector as well as for physics effects. Depending on the physics analysis, a subset of these corrections can be applied.

Correction algorithms are functions of E_T and η of the jet and have been optimized to estimate the corrected average energy, but are not meant to reduce the fluctuations around this value.

The most general form for the corrections to be applied to the P_T of a jet reconstructed within cone radius R is the following [48]:

$$\begin{aligned} P_T(R, P_T, \eta) &= [P_T^{raw}(R) \times f_\eta(R, P_T^{raw}, \eta) - Mp\bar{p}I(R)] \times f_{jes}(R, P_T^{raw}) - \\ &\quad - UE(R) + OOC(R, P_T^{raw}), \end{aligned} \tag{3.17}$$

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 so-called 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.

We will describe in the following the various correction stages and procedures. Tab. 3.3 shows the current naming convention for the different type of corrections.

η-dependent correction

Even after the calorimetric absolute scale calibrations described in Sec. 3.3.3, the response of the calorimeter is not uniform in pseudorapidity. 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, the two leading jets in the event are classified as follows: the jet in the range $0.2 < |\eta_{jet}| < 0.6$ is defined as "trigger jet", while the other is defined as "probe jet". In the case both jets in the events are

within $0.2 < |\eta_{jet}| < 0.6$, trigger and probe jets are assigned randomly. The P_T balancing fraction, $\Delta P_T f$, is then defined as:

$$\Delta P_T f = \frac{\Delta P_T}{P_T^{ave}} = \frac{P_T^{probe} - P_T^{trigger}}{\left(P_T^{probe} + P_T^{trigger}\right)/2}. \tag{3.18}$$

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

$$\beta_{dijet} = \frac{2+ < \Delta P_T f >}{2- < \Delta P_T f >}. \tag{3.19}$$

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 (eq. 2.1) 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. The number of $p\bar{p}$ interactions per bunch crossing, N, follows a Poisson distribution with a mean $< N >$ which depends on the instantaneous luminosity. 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.

These extra $p\bar{p}$ interactions increase the energy of the jets from the hardest scatter if their final state hadrons accidentally overlap with the jet originating from the primary interaction. This extra energy needs to be subtracted from the jet energy.

In order to compute the corrections, the number of primary vertices in the event, N_{vtx}, is taken into account. Indeed, N_{vtx} is a good indicator of additional interactions occurring in the same bunch crossing. The transverse energy in a random cone is measured in min-bias data and parametrized as a function of N_{vtx} in the event. 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

The *absolute* energy of a jet is defined to be the energy of the corresponding particle jet. The *absolute correction* transforms the energy measured in the calorimeter into a particle-level jet energy. After this correction the energy scale of a jet becomes independent from the CDF II detector. Since the calorimeter simulation has been tuned to reproduce the single particle response measured in the data, the absolute correction is obtained from simulation allowing the correction to cover the P_T range $[0, 600]\ GeV$.

The absolute correction is defined maximizing the probability $d\mathcal{P}(P_T^{particle}, P_T^{calo})$ of measuring a jet with P_T^{calo} given a jet with a fixed value of $P_T^{particle}$. This probability density function is parametrized by the difference between parton-level jets and calorimeter-level jets; where the former quantity is defined by means of stable particle information at generator level, the latter is defined by the standard jet-clustering algorithm.

Out-of-cone energy and underlying event

It is often desirable to reconstruct the energy of the original parton rather than the energy of the jet, in particular in the case of top quark mass measurements or Higgs boson searches.

The reconstruction of the parton energy from the jet energy, however, is subject to several difficulties. Portions of the parton energy can be lost from the jet cone due to large angle final state QCD radiation or due to particles exiting the cone during the fragmentation process. The energy associated to these processes is called *Out-of-cone energy*, and it is produced by leakage radiation outside the cone used for the jet definition.

On the other hand, a jet can receive energy contributions from particles others than the one purely associated with it due to initial state QCD radiations or interactions of the spectator partons in the proton due to the color connection mechanism. The latter are the so-called "beam-beam remnant" interactions. Initial state radiation and beam-beam remnant constitute the *underlying event*.

Both out-of-cone and underlying event corrections are derived from Monte Carlo dijet samples. Particle-level jets are matched to partons using $\Delta R < 0.4$. The difference between particle-level jet energy and the energy of the corresponding parton are parametrized using the same method as for the absolute corrections.

3.3.5 Jet energy response systematic uncertainties

The application of jet corrections is subject 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 P_T is shown in Fig. 3.15, and it results from the sum in quadrature of several contributions, as described in the following.

In the case of the η-dependent correction, the residual discrepancy from a flat response in η after the correction application on dijet data and Monte Carlo samples is taken as systematic uncertainty, which values range between 0.5% and 7.5% for central high P_T and forward low P_T jets, respectively.

On the other hand, the application of the absolute correction relies on how well the Monte Carlo simulation models the jet response in data. The uncertainties on the absolute correction is thus defined as the quadratic sum of the difference between data and simulation in the calorimeter response to single particles ([1.8, 2.8]% and 0.5% for hadronic and electromagnetic particles respectively), the difference in the momentum spectrum (1%), and the stability of the calorimeter calibrations in data (0.5%). The overall uncertainty on the absolute correction ranges between 2% and 3% for low and high P_T jets, respectively.

The uncertainties related to the multiple interactions correction (15%) depend on the primary vertex fake rate and the vertex reconstruction efficiency that can differ event by event depending on the event topology.

Moreover, the uncertainties on the out-of-cone corrections are estimated using $\gamma + jets$ samples, for which the photon P_T serves as an estimator of the corrected jet P_T. The sys-

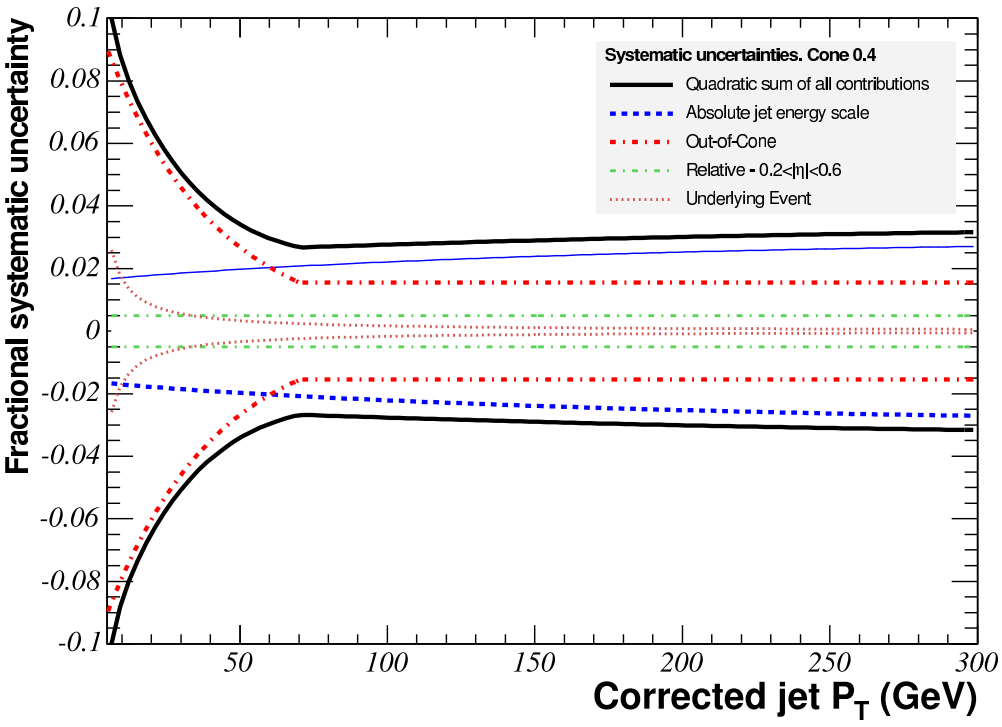


Figure 3.15: Total systematic uncertainties to corrected jet P_T.

tematic uncertainty is defined as the largest difference between data and either PYTHIA or HERWIG Monte Carlo samples in the transverse energy content of calorimeter towers close to the jet cone.

Finally, the uncertainty on the underlying-event correction is estimated comparing the underlying-event modelling in data and Monte Carlo using the average P_T of tracks in the transverse region with respect to the jet ($60° < \Delta\phi(jet, track) < 120°$).

The total systematic uncertainty to the jet response is depicted in Fig. 3.15 as a function of the jet P_T, and is obtained by summing in quadrature each of the systematic uncertainties described above. For $P_T > 60\ GeV/c$ 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 (out-of-cone) around the jet cone between data and Monte Carlo samples.

The total uncertainty on the jet energy scale varies between 8% at low jet P_T and 3% at high jet P_T.

3.3.6 Missing energy measurement

Neutrinos interact only through weak interactions and therefore cannot be directly detected as they traverse the detector material. Anyway, neutrinos production in an event

can be inferred by the presence of large 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 subject 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:

$$\vec{\not\!\!E}_T^{\,raw} = -\sum_{towers} (E_T^i)\vec{n}_i \tag{3.20}$$

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 jet 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. 3.3.5). In addition, mismeasurements of $\not\!\!E_T^{corr}$ may result from jets traversing poorly instrumented detector regions or may arise from cosmic rays background, muon misidentification or mismeasurements of the muon track momentum.

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 parametrized in terms of the total scalar transverse energy $\sum E_T$, which is defined as:

$$\sum E_T = \sum_{towers} E_T^i. \tag{3.21}$$

The $\not\!\!E_T$ resolution in the data is measured using minimum bias events[50], 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 accordingly to a Gaussian distribution with zero mean and $\sigma_x = \sigma_y = \sigma$ so that:

$$\frac{dN}{d\not\!\!E_{Tx}} \sim e^{-\frac{\not\!\!E_{Tx}^{\,2}}{2\sigma^2}} \tag{3.22}$$

$$\frac{dN}{d\not\!\!E_{Ty}} \sim e^{-\frac{\not\!\!E_{Ty}^{\,2}}{2\sigma^2}}. \tag{3.23}$$

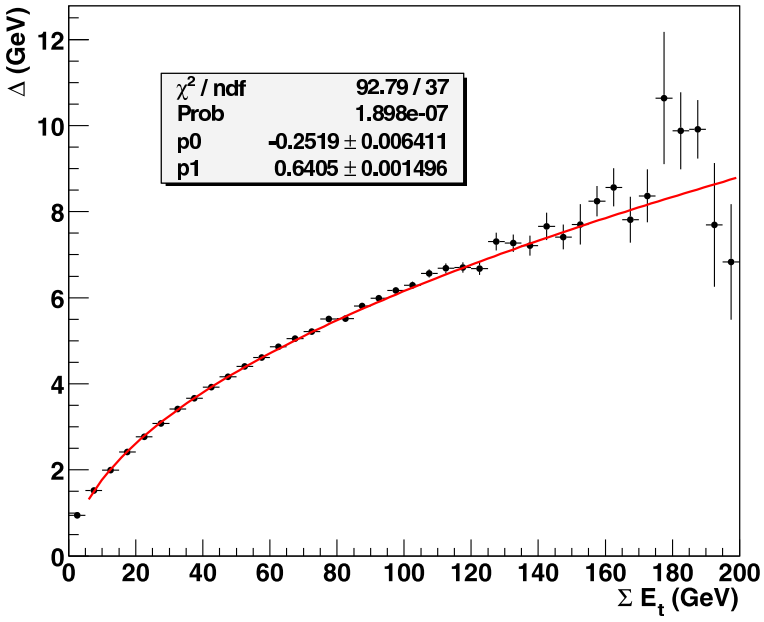


Figure 3.16: $\not\!\!E_T$ resolution as a function of $\sum E_T$ measured in minimum bias events[50].

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 $\Delta \sim 0.64 \sum E_T$[50], as shown in Fig. 3.16.

3.3.7 *b*-jet identification

The high position resolution provided by the silicon vertex detectors can be exploited to identify secondary vertices originated inside a jet by decays of long lifetime particles produced in heavy quark hadronization. This approach has been followed by the SECondary VerTeX (SECVTX) tagging algorithm [51, 52, 53, 54] and results to provide a great discrimination of high P_T b-jets from jets originated by light quarks or gluons.

The B hadrons produced by bottom quark hadronization have a lifetime of the order of a picosecond [55]; and at the typical energy of the bottom quark originating by top quark or Higgs boson decay, they travel some mm before decaying.

The SECVTX[54] algorithm relies on the displacement of secondary vertices relative to the primary event vertex to identify b hadron decays. The Run II algorithm is essentially unchanged from Run I[53], but the track selection cuts have been re-tuned for the CDF-II detector.

In order to select displaced tracks coming from decays of long-lived hadrons, a precise knowledge of the collision point is necessary. The event by event primary vertex location is determined using the procedure described in Sec. 3.2.4.

In the following the SECVTX algorithm will be described; while its performance on b-jet originating from top-quark decays will be treated in detail in Sec. 4.3.3.

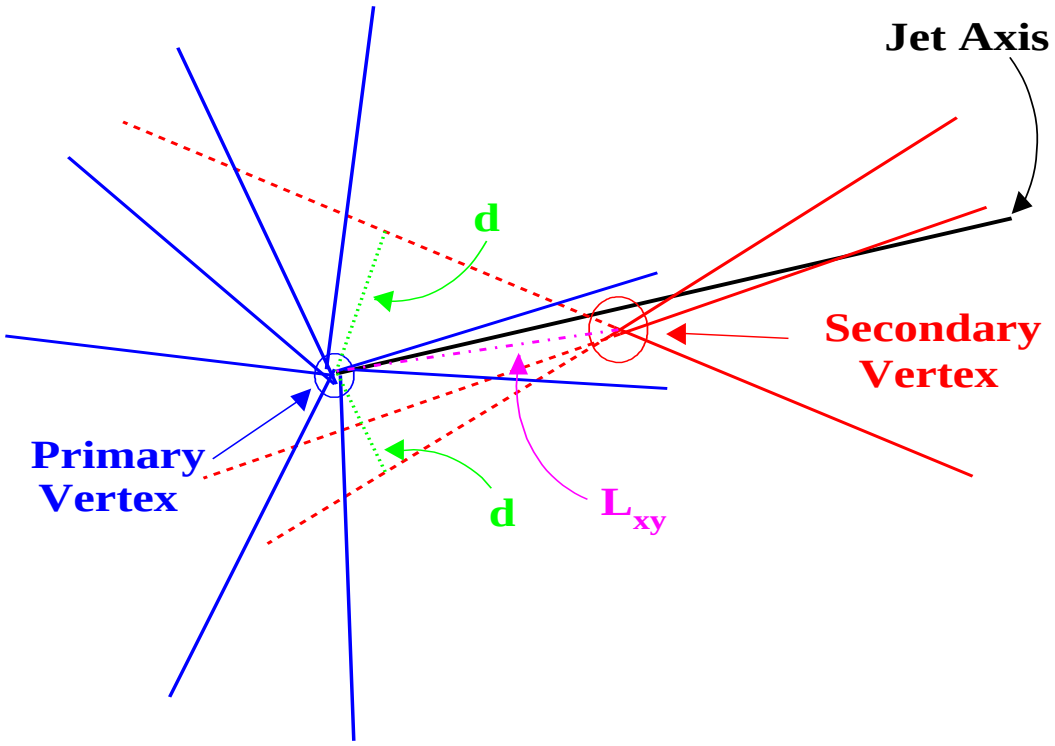


Figure 3.17: Reconstruction of the primary and secondary vertices in the r-ϕ plane. The impact parameter d for some representative tracks and the distance L_{xy} (or L_{2d}) between the vertices in the transverse plane are shown.

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. Track selection will be described in the following subsection.

Only jets with at least two of these tracks can produce a displaced vertex; a jet is defined as "taggable" if it has two good tracks. Displaced tracks in the jet are selected based on the significance of their impact parameter with respect to the primary vertex and are used as input to the SECVTX algorithm (Fig. 3.17).

There are two options for the tagging algorithm: tight and loose; in the latter option, not used in the analysis presented in this work, some of the requirements are loosened in order to retain efficiency in double tags analyses. In both cases, SECVTX uses a two-pass approach to find secondary vertices: in the first pass, using tracks with $P_T > 0.5\ GeV/c$ and $d_0/\sigma_{d_0} > 2.0$, it attempts to reconstruct a secondary vertex which includes at least three tracks (at least one of the tracks must have $P_T > 1\ GeV/c$). If the first pass is unsuccessful, it performs a second pass which makes tighter track requirements ($P_T >$

Variable		Loose		Tight	
		Pass 1	Pass 2	Pass 1	Pass 2
Use L00		yes		yes	
Use IO tracks		yes		yes	
SVX layers	$>$	2	4	3	
Track-χ^2/nDOF	$<$	8.0		8.0	
track-$\delta(z_0)$ (cm)	$<$	2.0		2.0	
Track-d0 (cm)	$<$	0.15		0.15	
d0 attachment significance	$<$	6.0		4.0	
Track-pT (GeV)	$>$	0.5	1.0	0.5	1.0
Track-d0 Sign.	$>$	2.0	3.0	2.0	3.5
Seed Vertex-χ^2	$<$	50			
At lest one Track-pT (GeV)	$>$	1.0	1.5	1.0	1.5
Track prune χ^2	$<$	90	1000	45	30
Vertex fit χ^2	$<$	120	2000	50	
Vertex L_{xy} Significance	$>$	6.0		7.5	

Table 3.4: Secvtx selections for the tight and loose options.

1 GeV/c and $d_0/\sigma_{d_0} > 3.5$) and attempts to reconstruct a two-track vertex (one track must have $P_T > 1.5$ GeV/c). 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[7]. The sign of L_{2d} is defined relative to the jet direction, specifically by the absolute difference $|\Delta\phi|$ between the jet axis and the secondary vertex vector (positive for $< 90°$, negative for $> 90°$). 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 mismeasured tracks are expected to be less displaced from the primary vertex. To reduce the background from false secondary vertices (mistags), a good secondary vertex is required to have $|L_{2d}/\sigma_{L_{2d}}| > 7.5$ (or 6.0 for the loose tagger option), where $\sigma_{L_{2d}}$, the total estimated uncertainty on L_{2d} including the error on the primary vertex, is estimated vertex-by-vertex and is of $\mathcal{O}(100\ \mu m)$. A tagged jet is defined to be a jet containing a good secondary vertex. Additionally, in order to reject secondary vertices due to material interactions or from long lived neutral particles, 2-track vertices obtained using pass-2 tracks are rejected if reconstructed at radii between 1.2 and 2.5 cm. Moreover, all secondary vertices with a radius greater than 2.5 cm are vetoed.

SecVtx track selection

The vast majority of reconstructed tracks are prompt, *i.e.* they originate from the primary interaction point whose location is estimated by reconstructing the primary vertex. Tracks that are displaced from the interaction point can be:

[7] L_{2d} is sometimes referred as to L_{xy}.

- mis-reconstructed tracks (mostly due to multiple scattering processes in the detector material);
- secondary particles produced by nuclear interaction in the detector material;
- the decay product of long-lived particles (b and c hadrons, but also strange particles, such as K_S and Λ).

For the b-jet identification we are obviously interested in selecting tracks coming from a heavy flavor hadron. Decays from K_S and Λ are removed explicitly by reconstructing the invariant mass of track pairs. Nuclear interactions are produced in the presence of material and thus an upper bound on the track impact parameter at 1.5 mm removes the majority of them. The track selection studied for the SECVTX algorithm aims at reducing the fraction of mis-reconstructed tracks and nuclear-interaction secondary tracks passing the selection requirements. Such tracks tend to have lower momentum, a large fit χ^2, and fewer or poorer quality attached hits in the silicon detector than genuine tracks from b or c hadron decays. Tracks are selected on the basis on these criteria:

- tracks must be contained in the jet cone:

$$\Delta R = \sqrt{(\eta_{track} - \eta_{jet})^2 + (\phi_{track} - \phi_{jet})^2} < 0.4;$$

- $P_T > 500\ MeV$;
- $d_0 < 1.5\ mm$;
- $\delta(z_0) = |z_0 - z_{prim.vertex}| < 5\ cm$;
- number of axial superlayers with 5 on more hits ≥ 2 and number of stereo superlayers with 5 or more hits $\geq$2;
- number of SVX $r - \phi$ hits ≥ 3;
- track fit $\chi^2/d.o.f. < 8.0$;

At this point, tracks with an oppositely charged partner track[8] which have $0.4876 < M_{\pi\pi} < 0.5076\ MeV$ (consistent with $K_0^S \to \pi^+\pi^-$) or $1.10963 < M_{\pi p} < 1.12163\ GeV$ (consistent with $\Lambda \to \pi p$) are removed.

3.3.8 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. At energies of $\mathcal{O}(10\,GeV)$ the dominant energy loss for electrons is bremsstrahlung. When electrons traverse the lead absorbers in the electromagnetic calorimeter, they interact with the nuclei of the material emitting photons which in turn produce e^+e^- pairs. Secondary particles are also very

[8] To each track is assigned the $\pi^\pm$ mass.

energetic and lead to the production of an electromagnetic cascade, called *shower*. The shape and the position of the electromagnetic shower is measured by the shower maximum detectors (CES and SMD), and shower characteristics are used for electron identification. In general an hadronic shower is longer and much broader compared to the electron shower whose energy deposition is in general contained in a single electromagnetic tower; on the contrary, hadronic showers are observed to continue into hadronic calorimeter section and to deposit energy into multiple adjacent calorimetric towers. In addition to shower properties, energy deposition in the calorimeter have to be matched to a track of comparable momentum pointing towards the center of the shower for a proper electron identification.

Electron (positron) and photon identification relies on the combination of tracking and calorimetric information. Electrons and photons leave a characteristic signature in the calorimeter, their electromagnetic shower. Electrons can be distinguished from photons in part by the slight difference of the shape of the EM 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.

Charged hadrons can mimic an electron signature if they shower early in the solenoid or in the EM calorimeter, or is they undergo to the process of charge exchange: $\pi^+ n \to \pi^0 p$ with $\pi^0 \to \gamma\gamma$, where the γ's give a signal in the EM calorimeter. Moreover, an electron can be mis-identified as a photon if the electron track is not reconstructed. Alternatively, a photon can be mis-identified as an electron if it converts to an e^+e^- pair as it goes through the material, or if a track is wrongly associated to the EM cluster. Photon conversions are identified by looking for pairs of COT tracks satisfying the following requirements:

- opposite charge;
- $|\Delta(xy)| < 2\ mm$;
- $|\Delta(\cot\theta)| < 0.04$;

where $\Delta(xy)$ is the distance between the tracks in the transverse plane calculated at the point where they are tangent in that plane; $\Delta\cot\theta$ is the difference between the polar angle cotangent of the two tracks. If a candidate electron combined with a partner track is consistent with a photon conversion, it is rejected. However, if a third track can be combined with the electron partner track to a form a conversion as well (trident), the conversion is likely to be due to a high-energy bremsstrahlung photon emitted by the initial electron as it goes through matter; in this case, the candidate electron is real and not rejected.

A baseline electron candidate is considered if a tower in the central electron calorimeter has $E_T > 2\ GeV$ and a track points to this tower. Adjacent towers in the same wedge (*i.e.* towers with the same ϕ) are added to the cluster. The energy attributed to the electron is the total cluster energy, while its momentum is inferred by the highest P_T track associated to the cluster; the direction of this track defines the electron direction, and it is used to compute the electron transverse energy, $E_T = E\sin\theta^{track}$. The basic selection cuts, shown in Tab. 3.5, use the variables described below, efficiently select electrons, and reject the background.

Variable	cut value
Fiduciality	Requirement to be in the active region of CES/CEM and away from calorimeter towers boundary
$\|z_0\|$	$< 60\ cm$
# of COT axial segments	≥ 3 with ≥ 7 hits each
# of COT stereo segments	≥ 3 with ≥ 7 hits each
E/P	< 2.0 if $P_T \leq 50\ GeV/c$
E_{HAD}/E_{EM}	$\leq 0.055 + 0.00045 \cdot E$
L_{shr}	≤ 0.2
Δz	$\leq 3.0\ cm$
$Q \cdot \Delta x$	≥ -1.5 and $\leq 3.0\ cm$
χ^2_{CES}	≤ 10

Table 3.5: Basic electron identification cuts. All energies (momenta) are in GeV (GeV/c)

- The track is required to satisfy some quality requirements: it must come from the region $|z_0| < 60\ cm$, and have a large number of hits attached to it (at least three axial and three stereo superlayers must provide segments with at least seven hits to the track reconstruction). The track must point to an instrumented region of the detector, and in particular must be away from the calorimeter tower edges;

- E is the total energy of the electron calorimeter cluster. P and P_T are the electron momentum and transverse momentum respectively, measured from the associated electron track.

- E_{had} (E_{em}) is the energy of the electron in the hadronic (electromagnetic) calorimeter.

- L_{shr} is a variable relative to the electromagnetic shower shape. The energy deposited in the two towers adjacent to the tower at the center of the cluster is compared to that expected from test beam data:

$$L_{shr} = 0.14 \times \frac{\sum_i (E_i^{obs} - E_i^{exp})}{\sqrt{(0.14\sqrt{E_{em}})^2 + \sum_i \delta(E_i^{exp})^2}}.$$

  L_{shr} is required to be less than 0.2.

- The shower maximum detector (CES) is used to reject possible hadron contaminations. The track is required to match a CES cluster in both axial ($|\Delta z| < 3\ cm$) and azimuthal ($-1.5 \leq Q \times \Delta x \leq 3\ cm$, where Q is the electron charge) directions. In the azimuthal direction, the shower asymmetry caused by the electron bremsstrahlung is taken into account. The shape of the CES cluster is required to be similar to the one evaluated from test beam data using a χ^2 requirement ($\chi^2_{CES} < 10$).

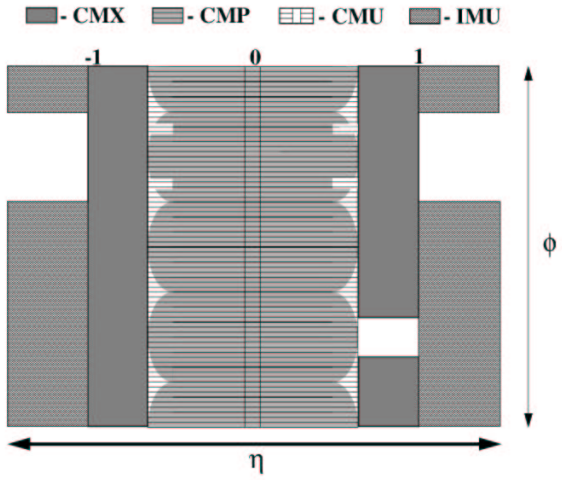


Figure 3.18: $\eta - \phi$ coverage of the Run II muon system.

3.4 Muon detectors

Thanks to their high penetration power, muons are separated from charged hadrons by the calorimeter, that acts as a shield for strongly and EM interacting particles. Muon identification can therefore be performed by extrapolating the tracks outside the calorimeter and matching them to tracks segments (called *stubs*) reconstructed in an external muon detector.
Figure 3.18 provides an overview the muon detectors coverage. The Muon system has been equipped with several devices able to provide coverage up to $|\eta| < 2.0$:

- Central MUon detector (CMU) consists of a set of 144 modules, each containing four layers of rectangular drift cells, operating in proportional mode. It is placed immediately outside the calorimeter and supplies a global coverage up to $|\eta|$ <0.6; the ϕ measurement of muon position is guaranteed by azimuthal segmentation of the detector, while the z coordinate is estimated on the basis of charge division.

- 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 3.18) according to the cylindrical structure of the inner detector. Contrary to CMU, this device supplies only ϕ information. For Run II, CMP benefits from an increased acceptance of $\sim$ 17%.

- 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. Azimuthal acceptance has been improved by 45% for Run II; only a 30°ϕ-gap remains to be used by the solenoid cryogenic system.

- 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) replaces the old Forward Muon Detector (FMU) to exploit the improved tracking capabilities and perform muon momentum measurement based only on the central solenoid field. In fact, during Run I the momentum of forward muons had to be measured by the FMU itself through a toroidal magnet; steel toroids are now used to supply mechanical support and shielding to new detector. IMU consists of four staggered layers of proportional drift tubes and two layers of scintillator tiles, arranged as for the CMP/CSP system to extend triggering and identification of muons up to $|\eta| \leq 1.5$ and $|\eta| \leq 2$, respectively.

3.4.1 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 candidate muon is reconstructed if such a stub is found in one of the muon systems and if an extrapolated COT track matches with the stub. The muon signature can be mimicked by hadrons that shower unusually late or not at all in the calorimeter and manage to escape the detector. Another source of background is due to muons from cosmic ray radiation, which can be vetoed by using timing information from muon and tracking systems. Tab. 3.6 shows the basic selection cuts applied for muon identification.

- The requirements on the COT track quality are the same as for electron reconstruction (Sec. 3.3.8).
- Additionally, in order to reject cosmic ray background, the track is required to have low impact parameter (the distance between the track and the beam line at point of closest approach). The cut on d_0 is more stringent if hits from the silicon detectors are attached to the track (in order to benefit from the higher position resolution achievable).
- For CMX muons (*i.e.* for muons produced at high pseudorapidity) the track is required to go through all COT superlayers.
- The energies, E_{em} and E_{had}, deposited in the electromagnetic and hadron calorimeters respectively along the muon trajectory are required to be small.
- The track is required to match the muon stub in the axial direction: Δx. The cut on this quantity depends on which muon system is considered.

Variable	cut value
$\|z_0\|$	$< 60\ cm$
$\|d_0\|$ (no Silicon hits)	$\geq 0.2\ cm$
$\|d_0\|$ (with Silicon hits)	$\geq 0.02\ cm$
# of COT axial segments	≥ 3 with ≥ 7 hits each
# of COT stereo segments	≥ 3 with ≥ 7 hits each
COT exit radius	$\geq 140\ cm$
E_{em}	$\leq \max(2, 2 + 0.0115(p - 100))$
E_{had}	$\leq \max(6, 6 + 0.0280(p - 100))$
Δx_{CMU}	$\leq 3.0\ cm$
Δx_{CMP}	$\leq 5.0\ cm$
Δx_{CMX}	$\leq 6.0\ cm$

Table 3.6: Basic muon identification cuts. All energies (momenta) are in GeV (GeV/c)

3.5 Other systems

3.5.1 Time of flight detector

The Time-Of-Flight system (TOF) expands the particle identification capability of CDF II in the low P_T region. 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

$$m = \frac{P}{c}\sqrt{\frac{(ct)^2}{L^2} - 1}. \tag{3.24}$$

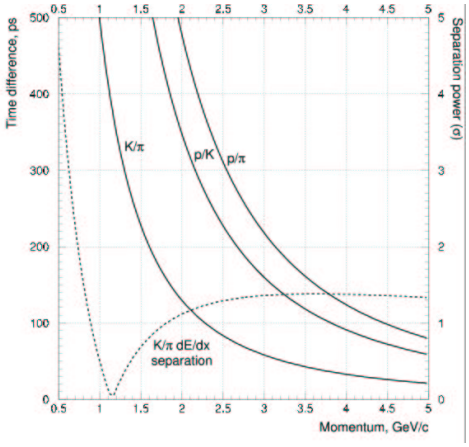
The TOF consists of 216 scintillator bars installed at a radius of about 138 cm (from the z axis) in the 4.7 cm space between the outer shell of the COT and the cryostat of the superconducting solenoid (see Figure 3.3). Bars are approximately 279 cm long and $4 \times 4\ cm^2$ in cross-section. With 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 PMTs, capable of providing adequate gain even if used inside the 1.4 T magnetic field. Usage of long attenuation length fast rises time scintillator in conjunction with these PMTs allowed to achieve specified the time measurement resolution of 100 ps.

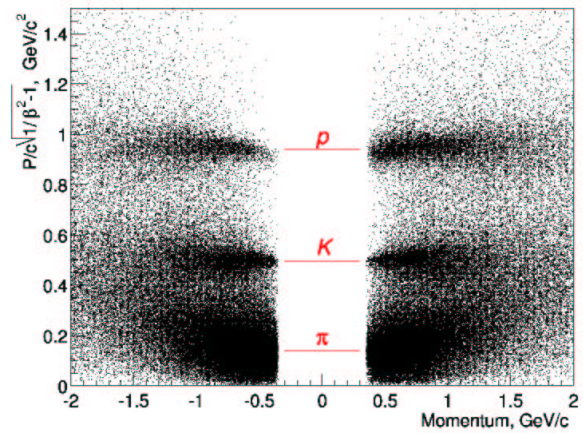
Figure 3.19 (a) shows the time-of-flight difference for K/π, p/π and K/p hypotheses and the separation power assuming resolution of 100 ps. In Figure 3.19 (b), early TOF performance is illustrated. More details on TOF construction and performance can be found in [41].

3.5.2 The Cherenkov luminosity counter

The Cherenkov Luminosity Counter (CLC) measures the average number of interactions per bunch crossing, μ, which is used to provide a measurement of the instantaneous



(a) K/π, p/π and K/p difference over a path of 140 cm as a function of momentum, expressed in ps and separation power(σ), assuming resolution of 100 ps. The dashed line shows the K/π separation power from dE/dx measurement by the COT.



(b) TOF reconstruction mass vs. momentum for positive and negative tracks. Cluster of data points corresponding to π, K and p are clearly visible. Data are from Tevatron store 860 (12/23/2001).

Figure 3.19: Time-Of-Flight system performance: design (a) and data (b).

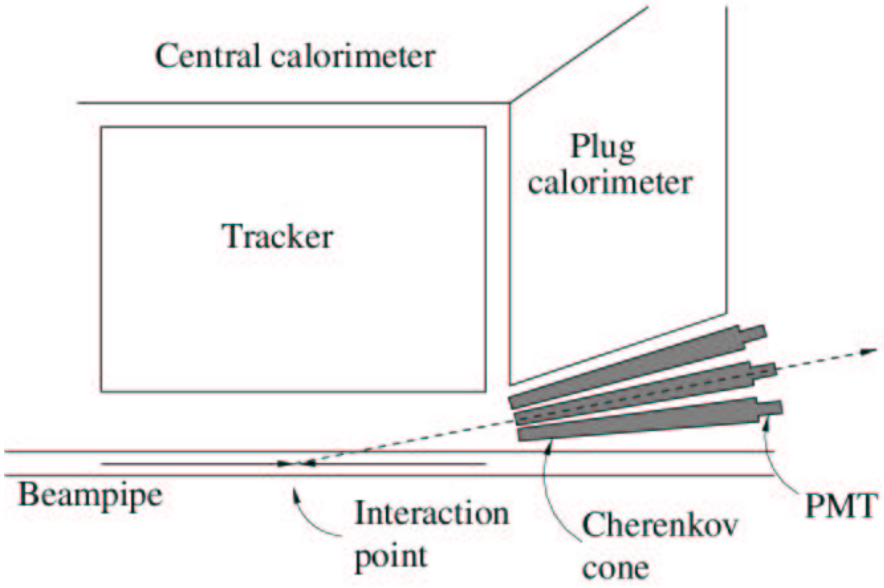


Figure 3.20: CLC geometry.

luminosity, $\mathcal{L}$, by means of the following relation:

$$\mu \cdot f_{bc} = \sigma_{p\bar{p}} \cdot \mathcal{L}, \tag{3.25}$$

where $\sigma_{p\bar{p}}$ is relatively well known 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.

The CLC 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.

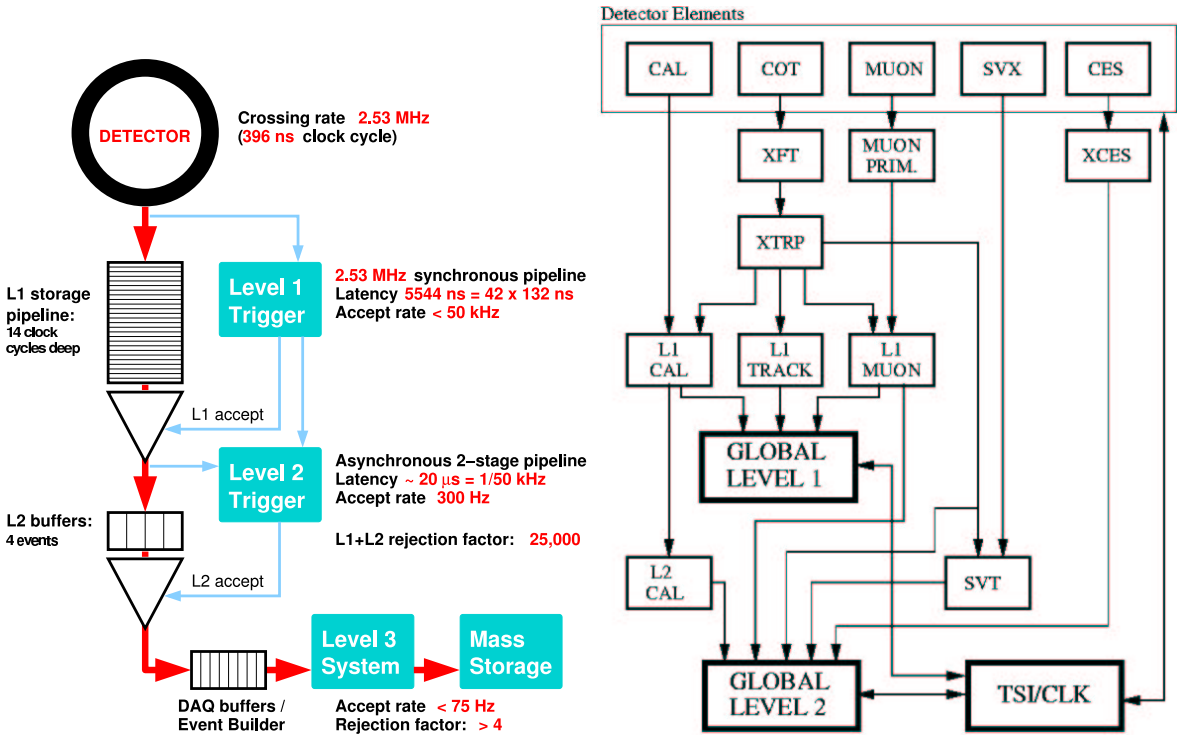
The idea is to use an assembly of long Cherenkov counters positioned in the Plug Calorimeter 3° gap, so that they point toward the interaction region, as schematically shown in Figure 3.20. This arrangement allows to make the detector much more sensitive to the particles coming directly from the interaction point because their flight path in the scintillator bars is the longest, and therefore the amount of the light produced the largest.

Excellent timing resolution and clever design allow the CLC to discern multiple interactions within the same bunch crossing and achieve an overall accuracy on the luminosity measurement better than 6%. Further information on the CLC design and performance is given in [60].

3.6 The trigger and data acquisition system

The trigger plays an important role in hadron collider experiments because 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, for a typical instantaneous luminosity of $10^{32}\ cm^{-2}s^{-1}$, results in a collision rate of 6.1 MHz, while the tape writing speed is only of $\sim$ 75 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 with each level providing a rate reduction sufficient to allow a more sophisticated event processing in the next level with minimal deadtime (Fig. 3.21). To allow time for transmission and processing of the detector signals to make the trigger decision, a Level 1 latency time of 5.5 μs has been chosen; each detector element has been provided of a local data buffering for the 14 event (at 396 ns bunch separation) that are expected to occur during such a period. Events passing the Level 1 trigger requirements are then moved to one of four on-board Level 2 buffers. This is sufficient to allow a 40 kHz Level 1 accept rate with $\leq 10\%$ deadtime for the successive 20 μs Level 2 processing time. Each separate Level 2 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 Level 2 trigger accept rate of $\sim$ 300 Hz. The Level-1/Level-2 trigger systems are managed by the Trigger Supervisor Interface (TSI) as shown in Fig. 3.21(b) which provides an interface



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

Figure 3.21: Block diagram detailing CDF DAQ and L1/L2 trigger system.

between the triggers and the Data Acquisition System as well as with the global clock and the bunch crossing signal.

Events satisfying both Level-1 and Level-2 requirements are transferred to the Level 3 trigger processor farm where they are reconstructed and filtered using the complete event information, with ≤ 75 event per second 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 Level 1, a Level 2, and a Level 3 trigger; in other words, a trigger path establishes a logic AND between selection procedures at different levels. Datasets (or data streams) are then formed by merging the data samples collected via different trigger paths; hence, datasets are defined by a logic OR between trigger paths. A graphical representation of the CDF trigger layout is depicted in fig. 3.22.

In the following, CDF trigger primitives available for each of the three trigger levels will be shortly presented.

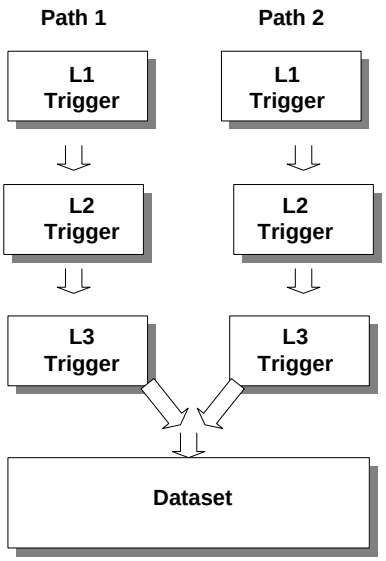


Figure 3.22: Schematic diagram of trigger paths and datasets.

3.6.1 Level 1 primitives

Tracks

The most significant tool for Level 1 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 results are then sent to an extrapolation unit (XTRP), whose task consists in extrapolating XFT tracks to the central calorimeter wedges and to the muon chambers (CMU and CMX). This allows a track to be matched to an electromagnetic calorimeter cluster for a first electron identification, a track to be matched to a stub on the muon detectors for improved muon reconstruction, and tracks to be used alone for specific triggers.

Calorimetric primitives

At Level 1, calorimetric towers are merged in pairs along η to define *trigger towers*, which are the basis for two type 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. In turn, the object trigger can be divided in single-object triggers and di-object triggers according to the number of trigger towers above threshold needed to generate a Level 1 accept.

Leptons

As already mentioned above, Level 1 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.

3.6.2 Level 2 primitives

Calorimeter clusters

Since jets are expected not to be fully contained into a single calorimeter trigger tower, the energy threshold on Level 1 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 Level 3. An effective rate reduction can be obtained at Level 2 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. 3.23:

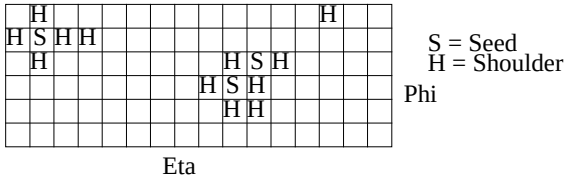
- electromagnetic and hadronic transverse energy of the trigger towers are checked to see if they are above predetermined *seed* and *shoulder* thresholds;
- all trigger towers whose energy has been found above the seed threshold are ordered according to increasing ϕ and η values.
- Cluster finding begins 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.
- Tower merged in the cluster are disabled from being merged in 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.

Different sets of thresholds can be used accordingly to the physics objects (typically electrons or jets) one wants to trigger on. To define a jet cluster, seed and shoulder thresholds are set to 3.0 and 1.0 GeV respectively both for electromagnetic and hadronic towers.

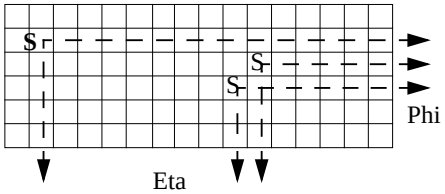
Level 2 clusters can be used to build object triggers by applying a cut on their transverse energy and position (provided from η-ϕ address of the seed towers), and global triggers by selecting on the number and $\sum E_T$ of clusters.

SVT tracks

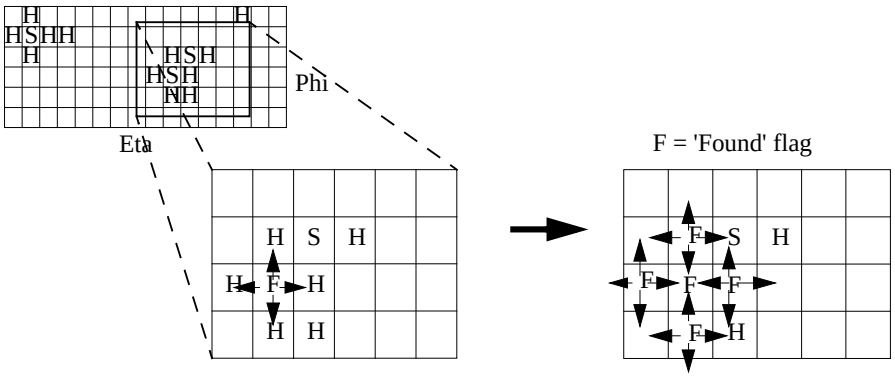
One of the most significant tools for the Level 2 trigger system is the **S**ilicon **V**ertex **T**racker (SVT) [56] which, for the first time in a hadron collider experiment, exploits the potential of a high precision silicon vertex detector to trigger on tracks with large impact parameter. This can make accessible a large number of important processes involving decays of b-hadrons with a long lifetime. Triggering on impact parameters offers remarkable



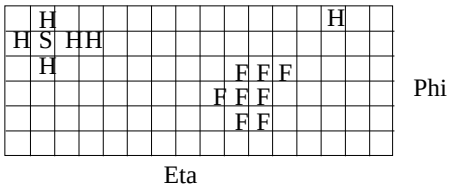
Step 1: seed and shoulder trigger tower tagging



Step 2: seed and shoulder trigger tower eta-phi address recording



Step 3: contiguous shoulder towers are merged into clusters



Step 4: trigger towers assigned to a cluster are excluded from further iterations

Figure 3.23: Cluster finding procedure.

improvements also to jet physics. For example, it can provide a $Z \to b\bar{b}$ signal, whose dijet mass peak can be used to measure resolution and mass-scale systematics on top-quark mass determination. Finally, hopes are that SVT can provide an high-efficiency collection of $HV \to b\bar{b}q\bar{q}^{(\prime)}$ events [57].

The overall architecture of SVT is shown in Fig. 3.24. First, each of 72 SVX II sectors (12 ϕ-wedges $\times$ 6 semi-barrels) is read out by a *Hit Finder*, that performs a hit clustering on each layer contained in the sector. For each cluster found, the Hit Finder computes the centroid, representing the most probable intersection point between the trajectory of

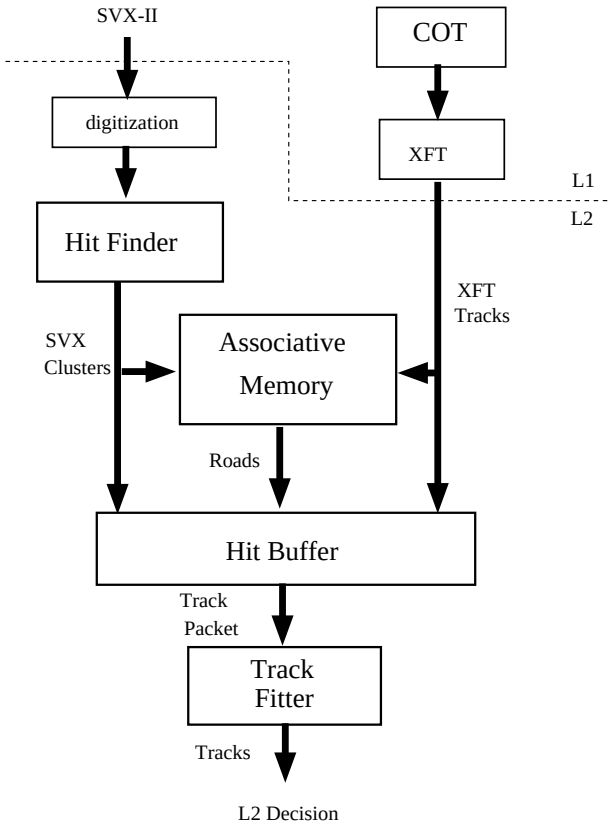


Figure 3.24: The SVT architecture.

a particle and the layer the cluster belongs to. The outputs from the Hit Finders of each SVX II wedge are merged into one stream and fed both into an *Associative Memory* (AM) and an *Hit Buffer*, together with track information from XFT.

The task of the Associative Memory is to perform the first stage of the pattern recognition: cluster centroids from the Hit Finders are mapped in *superstrips*; then, all possible combinations of stacks of superstrips and XFT tracks are compared to a pre-established set of admissible combinations (*roads*), each corresponding to a set of four SVX II superstrips and an outer XFT track. Superstrips size results from a compromise between a small size which would provide a precise pattern recognition but require a large memory, and a large size that would output a lot of fake track candidates but require a small memory. The present choice is 300 μm. At the same time the clusters and XFT tracks are sent to the Associative Memory, they are also sent to the *Hit Buffer*, where they are stored by superstrip number. Successively, Hit Buffer board receives the road number of candidate tracks and finds the superstrips and the XFT track corresponding to them. These sets of information, called *track packets*, are sent to the *Track Fitters*, where they are fitted by means of a linear approximated algorithm, consisting of a set of scalar products. For each track, P_T, ϕ_0 and d_0 are computed.

SVT efficiency studies[58], both on data and Monte Carlo, showed that for tracks with $P_T > 2\ GeV/c$ the resolutions were $\sigma_{d_0} \simeq 48\ \mu m$, resulting from the convolution of the actual beam profile ($\simeq 33 \mu m$) and the impact parameter resolution ($\simeq 35 \mu m$);

$\sigma_\phi \simeq 1\ mrad$ and $\sigma_{P_T} \simeq 0.3\%$ (with $[P_T] = GeV/c$).

Leptons

Level 2 muon primitives are essentially unchanged with respect to Level 1, the only difference consisting 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 Level 2 thanks to the information from central and plug shower maximum detectors.

3.6.3 Level 3 primitives

Events in input to Level 3 are loaded into a Linux PC farm, where they are almost fully reconstructed. In particular, jets, COT tracks and leptons are identified. The algorithms used for the reconstruction are the same used in offline analysis. 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.

If the event passes a Level 3 trigger, it is permanently stored on tape. An further offline processing is then performed on the selected events.

3.7 Offline data handling

The data flow from Level-3 triggers is segmented into ten streams, which are denoted with letters from A to J. All the streams are written to tape in real time, as the data are collected, *i.e.* in an on-line regime.

All other manipulations with data are referred to as off-line data handling. The importance 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 Level-3, 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 35 data sets which are used as input to physics analyses.

Chapter 4

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

As already highlighted in Chapter 1, top quark pairs are produced in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,TeV$ through $q\bar{q}$ annihilation and gluon fusion ($\sim 85\%$ and $\sim 15\%$ respectively). Since $|V_{tb}| \sim 1$ and $M_t > M_W + M_b$ the $t \to W^+ b$ decay dominates so that the different $t\bar{t}$ search channels are classified according to the W boson decay modes. When the produced W's both decay into quark pairs, the final state is called "all-hadronic" and nominally contains 6 jets. When one W is hadronic and the other decays into $e\bar{\nu}_e$ (*or c.c.*) or $\mu\bar{\nu}_\mu$ (*or c.c.*) the "lepton+jets" channel arises. Finally, when both W bosons decay in electron or muon modes one has the so-called "di-lepton" channel; a so-called "tau dilepton" category was added to accommodate $e\tau$ and $\mu\tau$ topologies in [61].

In this work, following a novel approach, we perform an inclusive search of the $t\bar{t}$ production process in the $\not\!\!E_T + jets$ final states. The isolation of this decay channel focuses on the selection of neutrino signatures from W boson leptonic decays, rather than in lepton identification, to isolate the $t\bar{t}$ production. This choice provides high acceptance with respect to leptonic W decays regardless of the lepton type, thus allowing for a sizable collection of un-observed $\tau + jets$ top pair decays. Indeed if standard lepton identification procedures, for both electrons and muons, have allowed so far to isolate high purity samples of $t\bar{t} \to e/\mu + jets$ events, on the other hand, the standard τ identification is found to have a too low efficiency to allow for an exclusive $\tau + jets$ search within the statistics of the data sample available at the time of this work (See Appendix A for more details). Moreover, $\not\!\!E_T + jets$ $t\bar{t}$ decays, that were not isolated so far, provide complementary results with respect to standard lepton+jets, di-lepton, and all-hadronic top pair searches, as well as new tools for the $t\bar{t}$ signal extraction.

In the following of this Chapter, collider data and Monte Carlo samples will be introduced, in order to describe the main handles available, as far as background rejection is concerned, for the analysis setup.

4.1 Monte Carlo samples

The simulation of $t\bar{t}$ events relies on the PYTHIA version v6.216[64] and HERWIG v6.504[65] which employ leading order QCD matrix elements for the hard process calculation, followed by parton showering to simulate gluon radiation and fragmentation

with the CTEQ5L[66] parton distribution functions. For heavy flavor jets, the decay algorithm QQv9.1[67] is used to provide proper modeling of b and c hadron decays.

The CDF II detector simulation reproduces the response of each detector subsystem to particles produced in $p\bar{p}$ collisions. Tracking of particles through the detector material is performed using the GEANTv3.2[68] package. Charge deposition in the silicon detectors is calculated using a simple geometrical model based on the path length of the ionizing particle[69]. The drift model for the COT uses the GARFIELD[70] package tuned to match COT data. The calorimeter simulation uses GFLASH[71] parametrization package interfaced with GEANT. The GFLASH parameters are tuned to test beam data for 50 GeV electrons and high-P_T pions, and further checked by comparing calorimeter energy deposits to the momenta measured in the drift chamber for isolated tracks in collision data.

The trigger system simulation is performed by a set of C++ packages built in the TRIGSIM++ executable emulating bit-by-bit the various trigger level decision steps.

The base Monte Carlo sample adopted for this work consists of $1,021,924$ PYTHIA generated inclusive $t\bar{t}$ events with $M_{top} = 178\ GeV/c^2$. The corresponding integrated luminosity of this sample, assuming $\sigma_{t\bar{t}} = 6.1\ pb$, is calculated to be $167.5\ fb^{-1}$.

Finally, in order to characterize the QCD background, $3M\ b\bar{b} + 4$ light jets and $2M$ 6 light jets ALPGEN+HERWIG events are used. ALPGEN[72] generates high multiplicity partonic final states using exact leading order matrix elements. Parton level events are then processed trough HERWIG and QQ for parton showering and b and c hadron decay. An additional sample of $200,000$ ALPGEN+HERWIG $Wb\bar{b}$+jets events is used for the electroweak background characterization.

4.2 Datasets

Several of the CDF datasets can in principle contain a detectable amount of $\not\!\!E_T + jets\ t\bar{t}$ events.

Different trigger paths were analyzed in order to evaluate the trigger efficiencies they provided on signal events. Multijet and missing transverse energy based triggers were studied: in particular, TOP_MULTI_JET, HIGGS_MULTI_JET, MET35_TWO_JETS, MET45 and MET_BJET trigger paths were considered and used to evaluate trigger acceptances using the output of the trigger simulation on inclusively decaying $t\bar{t}$ events. For each of these triggers the efficiency on signal events is reported before any kinematical requirements and after a minimal selection consisting in the requirement of at least four jets, one of which is asked to be identified as originating from a b-quark by the SECVTX algorithm (Sec. 3.3.7). The results of this investigation, with additional considerations, concerning the knowledge of sample composition as well as the understanding of trigger-induced biases, are used to choose the best trigger path to be adopted for our analysis.

Top-Multi-Jet trigger

The TOP_MULTI_JET trigger is specifically designed for the all hadronic $t\bar{t}$ decays, for which the nominal final state consists of six hadronic jets. Trigger requirements, among the three-level trigger architecture of the CDF data acquisition system, are set as follows:

Trigger Path	TOP_MJ	HIGGS_MJ	MET_2JET	MET45	MET_BJET
L1 eff. (%)	99.84 ± 0.02	99.84 ± 0.02	97.46 ± 0.07	97.46 ± 0.07	97.72 ± 0.07
L2 eff. (%)	64.35 ± 0.03	66.53 ± 0.21	97.04 ± 0.08	97.46 ± 0.07	66.06 ± 0.21
L3 eff. (%)	63.35 ± 0.04	65.24 ± 0.21	45.48 ± 0.22	35.26 ± 0.21	43.88 ± 0.22
L1 Ana eff. (%)	99.93 ± 0.02	99.93 ± 0.02	97.38 ± 0.09	97.38 ± 0.09	98.10 ± 0.08
L2 Ana eff. (%)	66.46 ± 0.27	81.30 ± 0.23	97.11 ± 0.10	97.38 ± 0.09	80.62 ± 0.23
L3 Ana eff. (%)	64.75 ± 0.27	79.78 ± 0.24	45.12 ± 0.29	34.91 ± 0.28	53.67 ± 0.29

Table 4.1: Trigger efficiencies on $t\bar{t}$ inclusive Monte Carlo events calculated for different trigger paths, before any kinematical selection and after (Ana) the requirement of at least four jets, one of which is required to be identified as originating from a b-quark.

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

All trigger level efficiencies, ϵ_{L1}, ϵ_{L2} and ϵ_{L3}, for inclusive $t\bar{t}$ events, are calculated from the trigger emulation output and are found to be $(99.84 \pm 0.02)\%$, $(64.35 \pm 0.03)\%$ and $(63.35 \pm 0.04)\%$, respectively. Trigger studies are reported in Tab. 4.1, where the efficiencies are reported for $t\bar{t}$ inclusive events before any kinematical selection as well as after the requirement of at least four jets and at least one positive SECVTX tagged jet.

Higgs-Multi-Jet trigger

The HIGGS_MULTI_JET trigger, although it has being specifically designed for the collection of events characterized by Higgs boson production in association with vector bosons in the all-hadronic final state, can provide a sizable acceptance for $t\bar{t}$ inclusive decays. Along with calorimetric requirement the trigger performs a selection based on SVT tracks with large impact parameter. The latter requirement allows an enrichment of the heavy flavor component of the data sample. Trigger requirements are set as follows:

- at Level 1: at least one calorimetric tower with $E_T \geq 10\ GeV$;
- at Level 2: at least two SVT tracks with $P_T \geq 2\ GeV/c$ and impact parameter $120\ \mu m \leq d_0 \leq 1\ mm$, and a total calorimetric $\sum E_T \geq 90\ GeV$.
- at Level 3: at least three jets with $E_T \geq 10\ GeV$ and an overall calorimetric $\sum E_T \geq 100\ GeV$.

The SVT track requirements at Level 2 enrich the genuine heavy flavor fraction of the sample and the presence of at least two displaced tracks at trigger level enhances the probability for an offline identification of jets originating from b-quarks. The trigger efficiencies for each level of selection are shown in the second column of Tab. 4.1 and

are found to be: $\epsilon_{L1} = (99.84 \pm 0.02)\%$ (the L1 requirement being the same as for the TOP_MULTI_JET trigger), $\epsilon_{L2} = (66.53 \pm 0.21)\%$, and finally $\epsilon_{L3} = (65.24 \pm 0.21)\%$. On the other hand, the trigger efficiency, calculated over the subsample of events with at least four jets and one positive b-tag are found to be: $(99.93 \pm 0.02)\%$, $(81.30 \pm 0.23)\%$ and $(79.78 \pm 0.24)\%$ for L1, L2, and L3 respectively. Despite the higher efficiency of these trigger requirements with respect to the TOP_MULTI_JET trigger, the requirement of displaced tracks produces a large bias in the sample which is difficult to model and to account for. In particular, even if the HIGGS_MULTI_JET trigger is found to collect $\sim 23\%$ more signal events, which results in a lower statistical error on a cross section measurement with respect to the TOP_MULTI_JET, the SVT trigger-induced bias and its related uncertainty, that need to be determined from an independent data sample, will overcome any measurement improvement produced by the higher signal statistics available. Moreover, any impact parameter based track requirements affect the SECVTX algorithm efficiency determination (see Sec. 4.3.3), thus requiring lot of efforts to be spent for the b-tagging algorithm efficiency determination on SVT triggered samples, a work that has not been carried out for high-P_T analyses.

Missing E_T + 2 jets trigger

The missing E_T+2 jets trigger is primarily designed for new physics searches, in particular for supersymmetric processes including squarks and gluinos. This trigger requirements, specified in the following,

- Level 1: missing transverse energy, $\not\!\!E_T$, exceeding 25 GeV,
- Level 2: at least two calorimetric clusters with $E_T \geq 10$ GeV and $\not\!\!E_T \geq 25$ GeV,
- Level 3: missing transverse energy above a 35 GeV threshold. At this stage of the trigger level the $\not\!\!E_T$ is recalculated using primary vertex information[1],

are found to be $(97.46 \pm 0.07)\%$, $(97.04 \pm 0.08)\%$, and $(45.48 \pm 0.22)\%$ efficient on $t\bar{t}$ events, for L1, L2, and L3 respectively. Contrarily to the cases already analyzed, the trigger efficiencies for events before any kinematical selection as well as for the subsample of them with at least four jets and one b-tag are found to be similar. The efficiency of each trigger level on $t\bar{t}$ events is reported on the third column on Tab. 4.1.

Inclusive Missing E_T trigger

The inclusive $\not\!\!E_T$ trigger serves essentially as a backup to the various missing E_T+object(s) triggers. It inclusively selects events with high missing transverse energy. Trigger requirements are set as follows:

- at Level 1: missing transverse energy, $\not\!\!E_T$, exceeding 25 GeV;

[1] The missing transverse energy measurement is here calculated using the position of the calorimeter towers with respect to the primary interaction, instead of the geometric center of the detector as for the first and second trigger levels (see Sec. 4.3.1).

- at Level 2: auto accept;
- at Level 3: missing transverse energy above a 45 GeV threshold.

Trigger efficiencies on $t\bar{t}$ inclusive events are reported in the fourth column of Tab. 4.1. The higher missing E_T requirement with respect to the missing E_T + two jet trigger yields a lower overall trigger efficiency: $\epsilon_{L3} = (35.26 \pm 0.21)\%$.

Missing E_T + Heavy Flavor trigger

This trigger is mainly designed for searches of non-SM signals, and places side by side missing E_T and displaced tracks requirements. The trigger requirements, defined as follows:

- at Level 1: missing transverse energy, $\not\!\!E_T$, exceeding 15 GeV and two $P_T \geq 2\ GeV/c$ XFT tracks,
- at Level 2: $\not\!\!E_T \geq 15\ GeV$ and two SVT tracks with $P_T \geq 2\ GeV/c$ and impact parameter, d_0, exceeding 100 μm,
- at Level 3: missing transverse energy above a 20 GeV threshold,

result in a final efficiency of $(43.88 \pm 0.22)\%$ which is slightly enhanced when evaluated on inclusive $t\bar{t}$ events with the additional requirement of four jets and one positive SECVTX tagged jets.

4.2.1 The multijet data set

In principle, any of the triggers described above could be used to pre-select a data sample in which to perform our search.

Anyway, as already mentioned before, SVT track requirements at trigger level, if on one hand provides an enrichment of the heavy flavor fraction of the data sample, on the other, yields sizable biases as far as the b-tagging algorithm is concerned which are difficult to model and to account for in the analysis. For this reason, our choice is to not employ as initial dataset any of the datasets collected using displaced track requirements.

Moreover, even if missing E_T requirements could in principle reduce the initial background amount in the triggered sample, it enhances the EWK+jets component in the sample and reduces the QCD-like fraction of events which is essential, as it will be discussed in Chapter 5, to parametrize the background b-tagging rates.

Both these effects create the necessity of long and detailed sample composition studies in order to identify the best way to compute the overall background amount. On the contrary, the TOP_MULTI_JET trigger provides, at the first order, a QCD-like sample in which background prediction tools are being developed in parallel to this work and which can directly be used, with opportune modifications, in the estimate of the background to $\not\!\!E_T + jets\ t\bar{t}$ decays.

From the previous reasoning, the results reported in this work are based on data collected from March 2002 to August 2004 by the Collider Detector at Fermilab using the

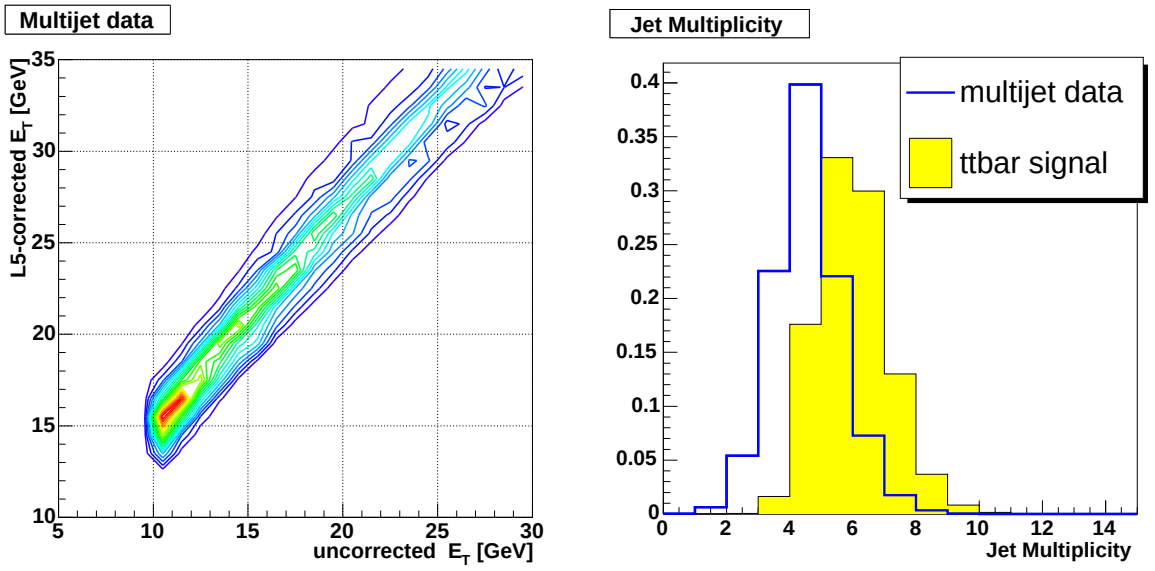


Figure 4.1: Correlation between uncorrected and L5-corrected jet energy observed in multijet data events (left panel) and number of jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ in $t\bar{t}$ inclusive and multijet data events (right panel) after TOP_MULTI_JET trigger cuts.

TOP_MULTI_JET trigger. The total integrated luminosity for this period is 417 pb^{-1}, of which 311 pb^{-1} has silicon detectors, calorimeters and muon systems fully operational, and are actually used for the analysis. The initial data sample consists of $4,249,644$ events.

4.3 Primary characteristics of signal events

Three main features distinguish the $t\bar{t}$ decay topology we are looking at from background processes: first of all, the leptonic decay of the W boson from the t-quark yields a considerable amount of missing transverse energy, $\not\!\!E_T$; on the other hand, the $\not\!\!E_T$ direction in the transverse plane $r - \phi$ is expected to be uncorrelated with respect any jet direction in the event. Finally, each $t\bar{t}$ event contains two b-jets whose presence can be established by using the SECVTX tagging algorithm described in Sec. 3.3.7. By exploiting the background rejection power provided by these three main features we will be able to isolate the signal we are interested in.

It is useful, at this point, to define the basic characteristics of jets we are going to consider in the following. We use jets reconstructed within the pseudorapidity range $|\eta| < 2.0$ whose L5-corrected energy (Sec. 3.3.4) exceeds the 15 GeV threshold. The E_T^{L5} cut value is set in order to enforce the jet energy threshold of $E_T^{raw} \geq 10\ GeV$ acting at trigger level, according to the correlation between uncorrected jet energies and the L5-corrected values observe for multijet data events, as shown in the left panel of Fig. 4.1. For comparison, the jet multiplicity for both multijet data and inclusive $t\bar{t}$ Monte Carlo events is shown in the right panel of Fig. 4.1.

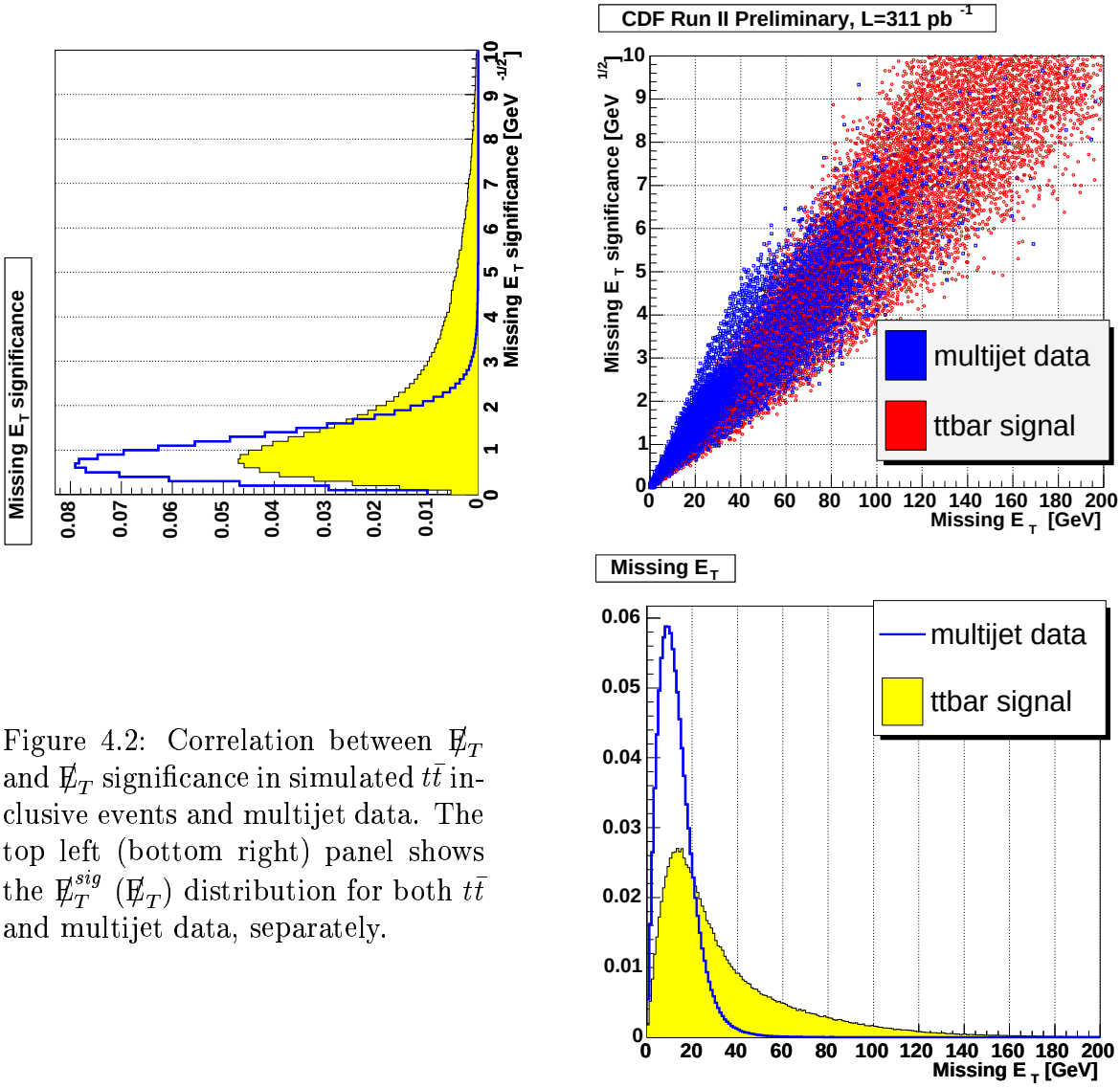


Figure 4.2: Correlation between $\not\!\!E_T$ and $\not\!\!E_T$ significance in simulated $t\bar{t}$ inclusive events and multijet data. The top left (bottom right) panel shows the $\not\!\!E_T^{sig}$ ($\not\!\!E_T$) distribution for both $t\bar{t}$ and multijet data, separately.

4.3.1 $\not\!\!E_T$ and $\not\!\!E_T$ significance

As highlighted in Sec. 3.3.6, 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 calorimeter towers:

$$\vec{\not\!\!E}_T^{\,raw} = -\sum_{towers} (E_T^i)\vec{n}_i, \tag{4.1}$$

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 sum extends up to $|\eta| < 3.6$.

The $\not\!\!E_T$ is the only observable signature that neutrinos leave in CDF II detector. In this section we will describe the main missing transverse energy sources and the correlated variables we adopted for the $\not\!\!E_T + jets\ t\bar{t}$ signal isolation.

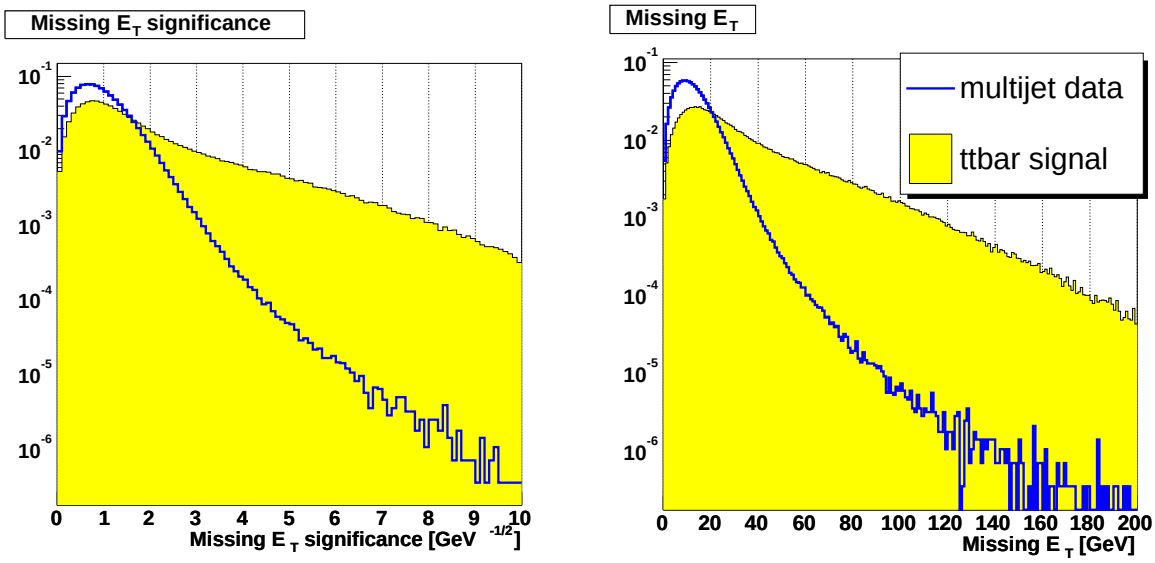


Figure 4.3: Logarithmic scale $\not\!\!E_T^{sig}$ ($\not\!\!E_T$) distributions for both $t\bar{t}$ and multijet data.

In general, apart from genuine neutrinos from W leptonic decays, missing transverse energy can be produced by jet energy mismeasurement and by b-quark semi-leptonic decays. The former is an instrumental effect that can be in part accounted for with the application of jet energy corrections as described in Sec. 3.3.4. On the other hand, b hadrons decay $(23.1 \pm 1.5)\%$ of the time into $\nu + X$[73], yielding some missing energy oriented along with the jet direction which needs to be accounted for by specific b-jet energy corrections which are still under development[74].

The resolution on the $\not\!\!E_T$ measurement is observed to scale as the square root of the total transverse energy, $\sum E_T$[50], for this reason the $\not\!\!E_T$ significance defined as:

$$\not\!\!E_T^{sig} = \frac{\not\!\!E_T}{\sqrt{\sum E_T}} \tag{4.2}$$

is found to be more discriminant than the $\not\!\!E_T$ as an analysis cut. Here and in the following, unless otherwise stated, we refer to $\sum E_T$ as the sum over all jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ in the event. The correlation between $\not\!\!E_T$ and its significance is shown in Fig. 4.2 for inclusive $t\bar{t}$ simulated events as well as for multijet data, together with the individual $\not\!\!E_T$ and $\not\!\!E_T^{sig}$ distributions for both samples (see also Fig. 4.3).

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

Several corrections are in general 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. In the following a short description of the corrections applied to the $\not\!\!E_T$ measurement will be given:

- Vertex correction: the $\not\!\!E_T$ is recomputed using the primary vertex instead of the geometric center of the CDF detector, which is used for the raw $\not\!\!E_T$ evaluation.

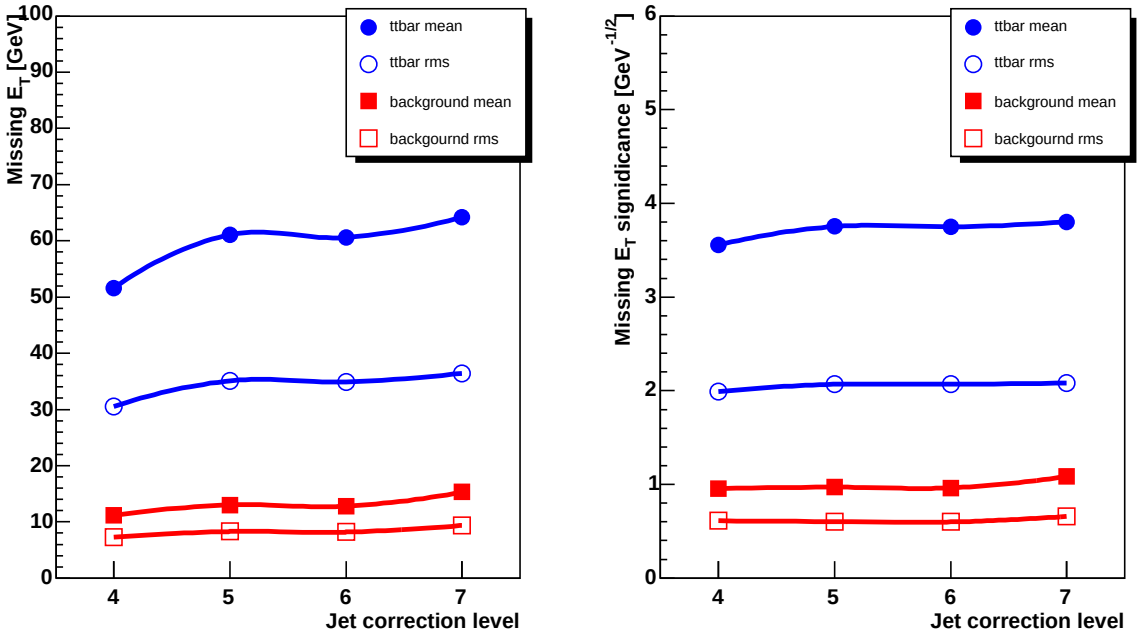


Figure 4.4: $\not\!\!E_T$ (left) and $\not\!\!E_T$ significance (right) distribution mean and RMS variation as a function of the applied jet energy correction level for $t\bar{t}$ and multijet data events.

- Muon corrections are applied in order to account for the modest calorimetric energy deposits released by high-P_T muons. Essentially, the $\not\!\!E_T$ is recalculated after the P_T^{μ} subtraction, and the addition of the hadronic and electromagnetic part of the energy deposit associated with the μ.

- Track-lepton unbalancing correction: this correction applies to those tracks which do not have energy in the calorimeter commensurate with their momentum (*i.e.* when $E_T/P_T < 0.7$, E_T being calculated in a 3×3 box of calorimetric towers around the track direction).

- Jet corrections propagation to the missing E_T measurement: the $\not\!\!E_T$ is calculated on top of the previous corrections, according with the corrections applied to jets.

The application of jet energy corrections alters the shape and the characteristics of the $\not\!\!E_T$ and $\not\!\!E_T$ significance distributions for both $t\bar{t}$ signal and multijet data. Distribution means and RMS's for $\not\!\!E_T$ and its significance are shown in Fig. 4.4 as a function of the applied jet correction level. According to these features we choose to adopt the jet correction level providing the best discrimination of signal against the background. Fig. 4.5 shows cut S/N and $S/\sqrt{N}$ optimization studies performed using both $\not\!\!E_T$ and $\not\!\!E_T^{sig}$ variables[2] for signal and multijet data as a function of the jet correction level applied.

[2]In this case the same jet energy correction level applied for the $\not\!\!E_T$ correction is used for jets entering in the $\sum E_T$ definition.

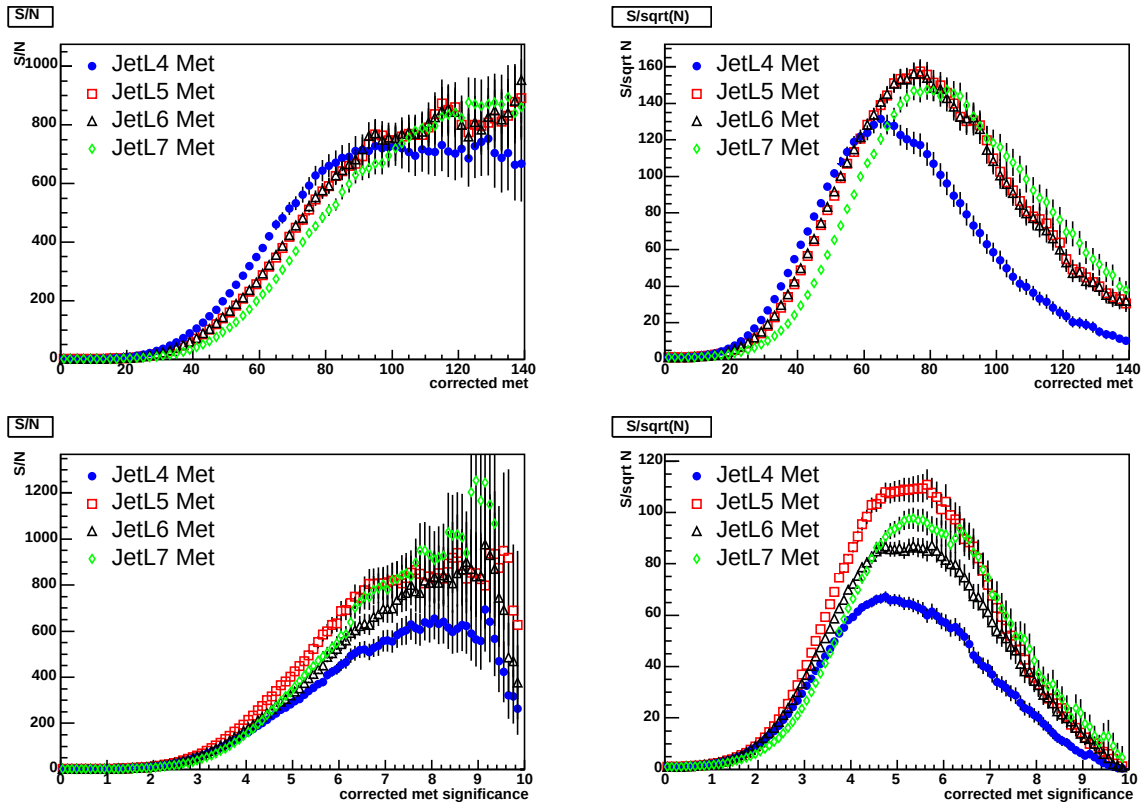


Figure 4.5: Cut S/N and $S/\sqrt{N}$ optimization studies performed using both $\not\!\!E_T$ and $\not\!\!E_T^{sig}$ distributions for signal and multijet data as a function of the jet correction level applied.

Level 5 corrections yield the best rejection of background with respect to the signal we are interested in.

4.3.2 $\not\!\!E_T$ geometrical properties

The direction of the $\not\!\!E_T$ in the $r - \phi$ plane also provides an handle to discriminate the various sources of missing transverse energy on a geometrical basis: the $\not\!\!E_T$ yielded by neutrino production in leptonic W boson decays is in general uncorrelated with any jet direction in the event. So if we define the $\min \Delta\phi(\not\!\!E_T, jet)$ as the minimum difference in the ϕ coordinates of the $\not\!\!E_T$ with each jet in the event, we expect the events to have large values of $\min \Delta\phi(\not\!\!E_T, jet)$ in the case of $W \to l\nu$ decays, and more in particular in the case of $t\bar{t} \to \not\!\!E_T + jets$ events. On the other hand, for multijet data events, for which the main source of missing transverse energy is represented by jet energy mismeasurement, we expect the $\not\!\!E_T$ to be aligned with the jet direction thus providing values of $\min \Delta\phi(\not\!\!E_T, jet)$ peaked around zero.

Fig. 4.6 shows the $\min \Delta\phi(\not\!\!E_T, jet)$ distributions for inclusive Monte Carlo $t\bar{t}$ and multijet data events. On the left panel, distributions are shown before any requirement on the $\not\!\!E_T^{sig}$. In this case no large differences are observed between data and signal distributions. On the contrary, requiring a large value of the $\not\!\!E_T^{sig}$, the expected behavior is observed for

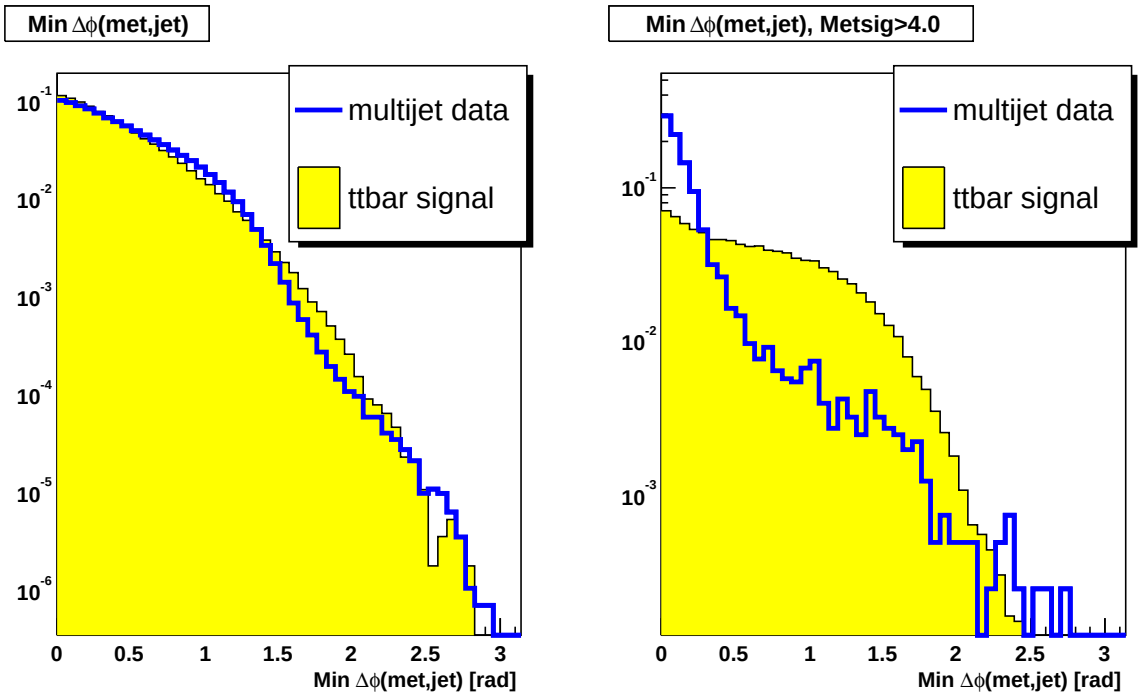


Figure 4.6: $\min \Delta\phi(\not\!\!E_T, jet)$ distributions for $t\bar{t}$ inclusive and multijet data events. On the left panel, distributions are shown before any requirement on the $\not\!\!E_T^{sig}$. On the right panel, distributions are shown after $\not\!\!E_T^{sig} \geq 4.0\ GeV^{1/2}$ cut.

both data and $t\bar{t}$ events (right panel of the figure).

4.3.3 *b*-jet identification rates

In the previous sections the main characteristics of the $\not\!\!E_T$ and its geometrical properties were described and studied in terms of their discrimination power against the background. However, high $\not\!\!E_T$ significance uncorrelated with jet direction can still be produced by processes other than the $t\bar{t}$ production: $W \rightarrow l\nu$ produced in association with jets can indeed yield the same missing energy signature as the $\not\!\!E_T$+$jets$ $t\bar{t}$ decays. The requirement for at least one identified *b*-jet can help to further reject these concurring processes. The *b*-jet identification is performed using the SECVTX algorithm described in Sec. 3.3.7. In the following *b*-jet identification efficiency and fake rate will be discussed.

b-jet identification efficiency determination and scale factor

The *b*-jet identification efficiency of the SECVTX algorithm can not be determined on a Monte Carlo basis only: indeed, imperfect detector descriptions, tracks from underlying event interactions which are difficult to model, multiple interactions which are not reproduced in the Monte Carlo, and different heavy flavor contents of the various samples, yield the necessity to measure the *b*-tagging efficiency directly from data and then to introduce a data-to-Monte Carlo scale factor to account for differences, if any, in the efficiencies

measured on the data and on Monte Carlo events.

In order to estimate the efficiency of the SECVTX tagger directly using data events, it is necessary to identify a control sample of pure b-jets. The following step consists in examining the ratio of the b-tagging efficiencies as measured in the data and the Monte Carlo and correct accordingly the Monte Carlo derived efficiency (*i.e.* applying the so-called SECVTX scale factor, or SF). In this way, 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.

As a b-enriched control sample, dijet events which have a lepton within one jet (henceforth referred to as the "lepton-jet") are used[75]. The sample is further selected by requiring at least one tagged jet back-to-back with respect to the lepton jet (henceforth referred to as the "away-jet"). The lepton-jet is in this way enriched in heavy flavor content both by the requirement of a lepton within the jet (consistent with a semileptonic b-quark decay), and by tagging the away side jet (hence preferentially selecting $b\bar{b}$ events). In the selected sample b-tagging rates on the lepton-jet are examined to determine the b-jet identification efficiency. Some complications arise due to the bias yielded by possible residual light flavor contaminations 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[76] makes use of conversions in order to calculate the residual light flavor contribution to the lepton-jet tags; on the other hand, the muon method[77] uses a Monte Carlo template of the transverse momentum of the muon relative to the jet axis to fit to data distribution. Both methods rest on the assumption that 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. Moreover, both methods assume the tagging on the lepton side to be uncorrelated with the tagging on the away side. Finally, both methods also consider the scale factor to be the same for b- and c-jets.

A combination of the two scale factor measurements[78] has been performed by minimizing a generalized likelihood. This procedure requires the determination of the correlation between the two scale factor measurements and the associated systematics. The combined result provide a SECVTX scale factor determination of $SF = 0.909 \pm 0.060$.

Mis-identifications

A SECVTX tag on a jet that does not contain heavy flavor is referred to as a fake tag or mistag. Mis-identification by the b-tagging algorithm are due to several sources. Firstly, some false tags can originate from tracking resolution effects: any given track has a probability to be displaced simply due to the Gaussian nature of its parameter uncertainties. When several tracks randomly have large displacement significances d_0/σ_{d_0}, they can combine to form a mistag. These mistags can be reduced by selecting good-quality vertices with large L_{2d} displacement. Secondly, some mistags can be produced by long-lived particles, such as K_0^S and Λ, in light-flavor jets. These can be mitigated to some extent by requiring tags to fall outside these particles mass windows, specifically by asking the total mass of the tracks inside the tag to fall outside opportune mass windows around these particles (Sec. 3.3.7). Finally, b-jet mis-identifications can be yielded by material interactions and conversions on the beampipe and on inner silicon detector layers. These can be

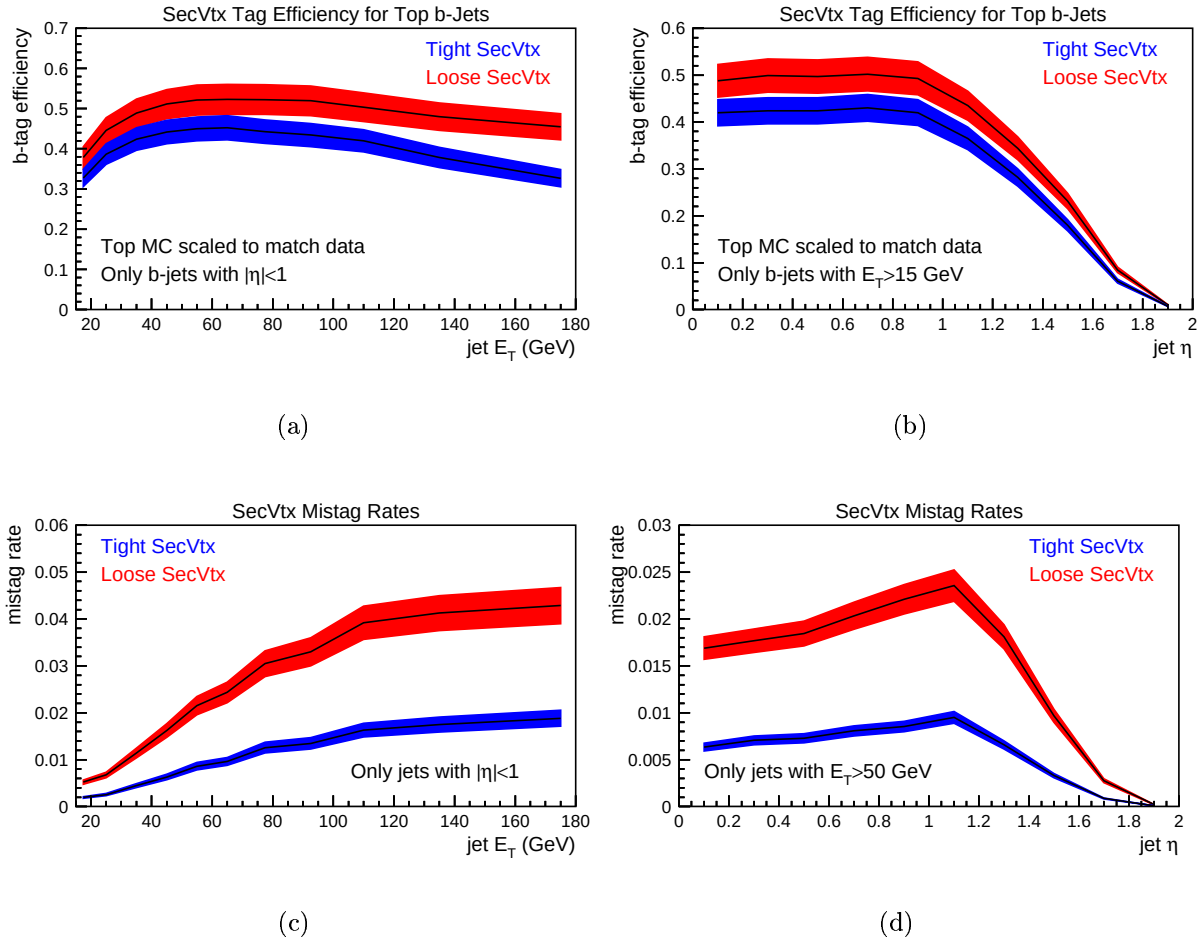


Figure 4.7: SECVTX tagger efficiency and fake tag rate. Both tight (blue) and loose (red) options of the SECVTX algorithm are shown. See the text for details.

mitigated by disallowing two-track vertices reconstructed within the region occupied by the detector material.

Although the amount of mis-identification can be reduced in some sense, no method is 100% effective. In order to estimate the residual mistag rate negative-L_{2d} tagged jets are used.

Overall tagging efficiency and mistag rate

The overall efficiency to tag a fiducial b-jet in $t\bar{t}$ events is about 44%. Fiducial jets are defined according to the following requirements: $E_T^{raw} > 10\ GeV$ and $|\eta| < 2.0$. On the other hand, the overall negative tag rate for jets in Monte Carlo $t\bar{t}$ signal events is found to be 0.5%. Figures 4.7(a) and (b) show the SECVTX efficiency times scale factor in $t\bar{t}$ events versus jet E_T and η, respectively. Figures 4.7(c) and (d) show the SECVTX negative tag rates versus jet E_T and η, respectively. Both performances for the tight and loose versions of SECVTX are shown, although only the tight version (blue) 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 rises as a function of jet E_T, and then eventually falls back off. This is due to the imposed cuts on the maximum vertex radius allowed, 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 rates also rise as a function of jet E_T. However, because we only plot negative tags due to resolution, it does not have the same drop-off as the efficiency. The negative tag rate essentially reaches a plateau at around 100 GeV. The negative tag rates also increase with jet $|\eta|$, and then fall off 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.

We will adopt the requirement of at least one positive tagged jets in order to define our final sample by which to measure the $t\bar{t}$ production cross section.

4.4 Other kinematical characteristics of signal events

Besides the missing E_T, $\min \Delta\phi(\not\!\!E_T, jets)$ and the b-tagging requirements, other kinematical variables could be used to characterize the $t\bar{t}$ production with respect to background processes. Even if not directly applied for the kinematical selection, which will be described in detail in Sec.5.2.1, we found interesting to introduce in this section some other kinematical variables in which the top pair production manifests differently compared to background processes; these kinematical variables will be used to check our the background prediction method (Sec. 5.1.1 and 5.1.3), as well as to characterize the final data sample by means of two-component likelihood fits (Sec. 6.2).

Topology- or energy-related variables distinguish $t\bar{t}$ production from the multijet background. Topological variables in some sense describe the jet activity in the event, and are defined to be the aplanarity, the centrality and the sphericity[79]. For each event, a normalized momentum tensor, M_{ab}, is defined as follows:

$$M_{ab} = \frac{\sum_j P_{ja} P_{jb}}{\sum_j P_j^2}, \tag{4.3}$$

where a, b run over the three space coordinates and P_j is the momentum of the jet j. M_{ab} is a symmetric matrix and can be diagonalized. Using unit eigenvectors $\mathbf{n_1}$, $\mathbf{n_2}$, $\mathbf{n_3}$, the eigenvalues Q_i are normalized and ordered by:

$$Q_1 + Q_2 + Q_3 = 1, \qquad 0 \leq Q_1 \leq Q_2 \leq Q_3. \tag{4.4}$$

The M_{ab} eigenvalues are used to quantify 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_i are used to define topological variables.

The Sphericity, S, is defined as:

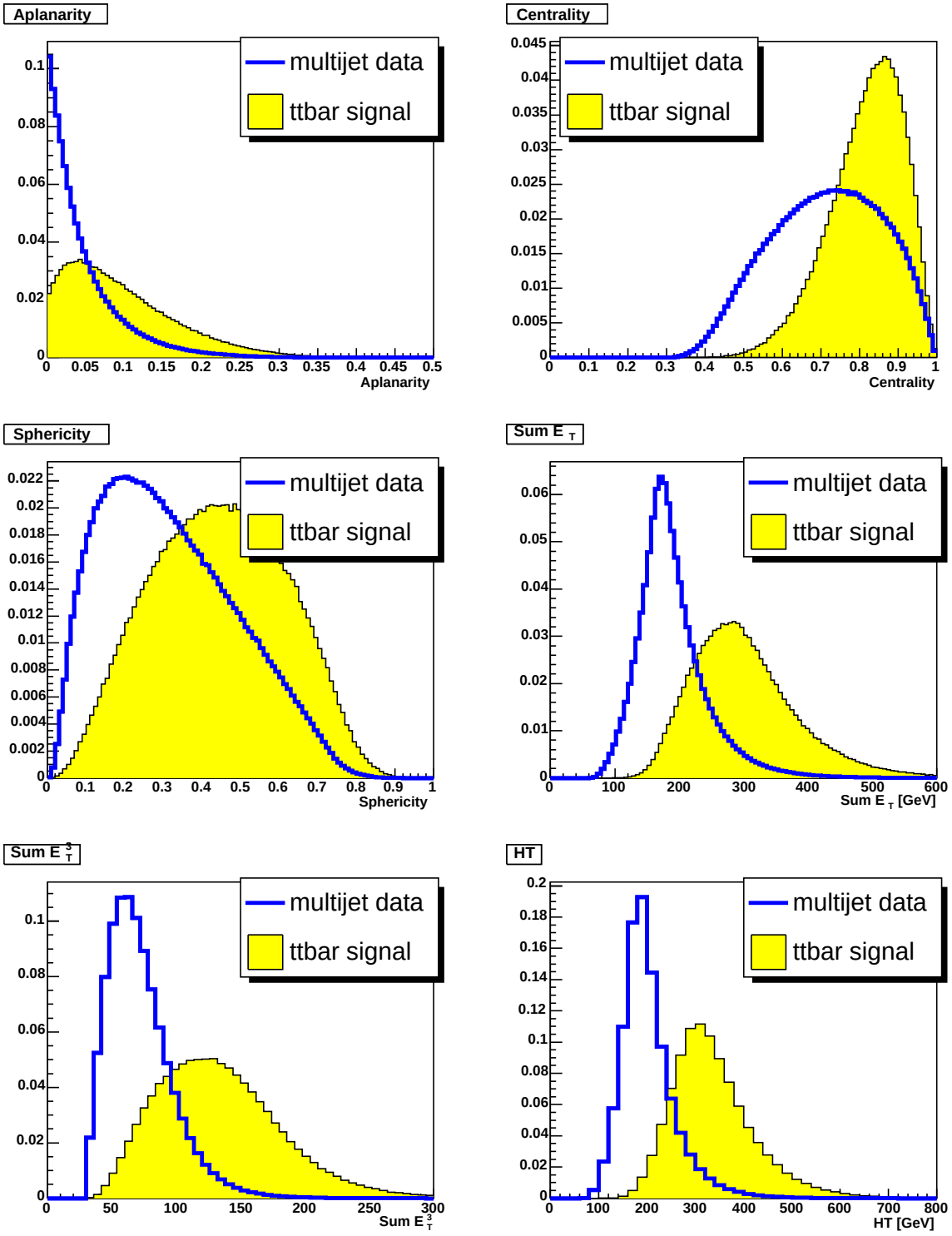


Figure 4.8: Aplanarity, Centrality, Sphericity, $\sum E_T$, $\sum E_T^3$ and $H_T = \not{E}_T + \sum E_T$ distributions for multijet and inclusive Monte Carlo $t\bar{t}$ events after the requirement of at least four jets with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$.

$$S = \frac{3}{2}(Q_1 + Q_2). \tag{4.5}$$

Sphericity values lie in the range $[0, 1]$: events with $S \approx 1$ are rather spherical, while on the contrary, events characterized by $S \ll 1$ are cigar-shaped, looking like a pair of back-to-back jets.

The Aplanarity, A, is defined by

$$A = \frac{3}{2}Q_1, \tag{4.6}$$

and it is normalized to lie in the range $[0, 1/2]$. A is small for coplanar (include collinear) events. More in particular, extreme values of A are reached in the case of two opposite jets and in the case of evenly distributed jets, respectively.

Moreover, distributions of energy-related variables, such as Centrality, $\sum E_T$, $\sum E_T^3$ and H_T, for $t\bar{t}$ events are expected to differ with respect to multijet data events. Centrality is defined as the ratio of the event $\sum E_T$ and $\sqrt{\hat{s}}$ where $\sqrt{\hat{s}}$ is the center-of-mass energy in the hard scattering reference frame[3]. The total transverse clustered energy $\sum E_T$ is defined as the sum over all jets, with $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$, in the events. On the other hand, the $\sum E_T^3$ is defined by the sum over all $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ jets in the event except the two leading ones. Finally, H_T is defined by $\sum E_T + \not\!\!E_T$.

Fig. 4.8 provides the comparison of these kinematical variables for inclusive Monte Carlo[4] $t\bar{t}$ and multijet data events, with at least $4(E_T^{L5} \geq 15\ GeV;\ |\eta| \leq 2.0)$ jets.

Even if the signal to background discrimination power provided by these variables is large in the data sample with at least 4 $(E_T^{L5} \geq 15\ GeV;\ |\eta| \leq 2.0)$ jets, the effective background rejection of each kinematical variable is correlated with the selections already performed in the data sample. For example, we will see that the gain provided by a cut on aplanarity is negligible after having applied tight cuts on the $\not\!\!E_T^{sig}$, $\min \Delta\phi(\not\!\!E_T, jets)$ and having required at least one positive b-tagged jet (Sec. 5.2.1).

4.5 Event clean-up

Before treating in detail the event selection optimization, it is useful to define some clean-up cuts which reject those events collected under sub-optimal detector 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, events with well reconstructed high-P_T leptons are rejected in this preliminary selection step, in order to guarantee orthogonality with respect to other $t\bar{t}$ cross section analyses relying on the lepton+jets decay signature.

The following prerequisites are applied to the multijet and inclusive $t\bar{t}$ signal samples.

[3] $\hat{s}$ 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 $x_2 = (\sum E - \sum P_z)/(1.96\ TeV)$ and $\sum E$, $\sum P_z$ are the sum of the energy and of the z component of the momentum of all jets in the event, respectively.

[4] Monte Carlo events are required to satisfy the TOP_MULTI_JET trigger selections.

- All events collected under the same detector conditions and on a time scale of $6 \div 12$ hours belong to a so-called *run*. We adopted the Good run list v7.0 provided by the CDF Data Quality Group for which runs are requested to have silicon, muon and calorimeter detectors fully operational (there is an exception regarding the CMX muon subdetector whose status is ignored)[80].

- We reject events with a good, $high - P_T$ electron or muon to avoid overlaps with other top lepton+jets analyses. The number of tight leptons in the Monte Carlo samples is corrected applying the high-P_T leptons identification scale factors: $SF_e = 0.996 \pm 0.005$ [81], $SF_\mu^{CMUP} = 0.874 \pm 0.009$ and $SF_\mu^{CMX} = 0.989 \pm 0.006$ [82] which are used to rescale the identification efficiency of leptons measured on Monte Carlo events to the one measured on data. Selection cuts for high-P_T electrons and muons are listed in detail on Appendix B and as well as in [81, 82].

- We require the events to have the primary vertex location well centered in the CDF II detector, in particular:

  - the number of good quality vertices in the event is required to be greater than zero. A vertex is defined to be of good quality if it is formed with at least three COT tracks[83].
  - 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; this cut selects well-centered events.
  - The distance between the event primary vertex[46] and the vertex used for jet reclustering $|z_{jet} - z_{primvtx}|$ is required to be less than 5 cm. Indeed for jet reclustering and the missing transverse energy calculation, several different event z vertex positions can be used[84]:

    0. nominal CDF origin at $z = 0.0$ cm;
    1. z_0 of the good quality highest-P_T vertex;
    2. z_0 of the tight lepton with the highest transverse energy;
    3. z_0 of the good quality vertex closest to the z_0 of the tight lepton with the highest transverse energy;
    4. average z_0 of the tight leptons;
    5. same as 3 but if no tight isolated leptons are found, tight non-isolated leptons are used;
    6. same as 2 but if no tight isolated leptons are found, tight non-isolated leptons are used.

    For our analysis we used option (1.) of the above list. The $|z_{jet} - z_{primvtx}| < 5$ cm condition ensures the the vertex used for jet reclustering and then for the secondary vertex search to be close to be primary vertex found in the event (see Sec.3.2.4).

The impact of these preliminary selections on both inclusive Monte Carlo $t\bar{t}$ and multijet data events, is shown in Tab. 4.2. For reference the effect of the clean up selections

N evt	$MC_{\tau+jets}$	MC_{incl}	MJ
Good Run	1,021,924	1,021,924	4,249,644
$BR(t\bar{t} \to \tau + jets)$	149,323	–	–
Trigger	82,200	647,365	4,249,644
N tight leptons = 0	78,084	583,697	4,245,940
$\|Z_{vert}\| < 60\ cm$	75,035	559,342	4,051,242
$\|Z_{jet} - Z_{primvtx}\| < 5\ cm$	74,912	558,539	3,897,756
N_{vert} good quality ≥ 1	74,904	558,528	3,897,755

Table 4.2: Events surviving the clean-up requirements for both inclusive Monte Carlo $t\bar{t}$ and multijet data samples. The $\tau+jets$ exclusive channel is obtained by asking, at HEPG level, for a W to decay hadronically, and the other in the $\tau + \nu_\tau$ mode.

is also shown for the exclusive τ+jets top pairs decay channel. In the case of Monte Carlo events, prerequisites are applied after having accounted for the trigger efficiency as given by the trigger simulation.

Chapter 5

Background parametrization and event selection

In this Chapter, the background prediction method and the event selection optimization aimed at measuring the $t\bar{t} \to \not\!\!E_T + jets$ production cross section will be described.

The overall analysis setup is based on the idea that we can distinguish $t\bar{t}$ production from background processes, in a given kinematically selected sample, by means of their different b-jet identification rates. For this reason, the cross section measurement will exploit the excess in the number of b-tagged jets over the background expectation, as follows:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{exp}}{\epsilon_{kin} \cdot \epsilon_{tag}^{ave} \cdot L}, \tag{5.1}$$

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, $\mathcal{L}$ is the integrated luminosity of the TOP_MULTI_JET triggered data sample.

The first part of this Chapter will be devoted to obtaining a reliable prediction of the total amount of background b-tags, which will be used for a kinematical selection optimization procedure on the 311 pb^{-1} data sample. The complete sample of multijet data can not be directly used for the analysis cuts optimization since, even if small, signal contaminations are present. Instead, as it will be described, the subsample of events with exactly 3 jets with $E_T \geq 15\ GeV$ and $|\eta| \leq 2.0$, where the $t\bar{t}$ fraction is totally negligible, constitutes the basis for a per-jet b-tagging probability parameterization to be used for the background prediction. Checks of the reliability of the background estimate will be performed in several control samples of data.

5.1 The background prediction: positive b-tagging rate parametrization

As already mentioned in the introduction of this Chapter, the estimate of the number of b-tagged jets yielded by background processes constitutes one of the milestones of the

analysis described in this work. Indeed not only the background rejection power is used to set the analysis cuts on kinematical variables like $\not\!\!E_T^{sig}$ and $\min \Delta\phi(\not\!\!E_T, jets)$ but also the amount of b-tagged jets expected from background processes, in a given kinematical selected sample, is in input to the event selection optimization. As it will described in Sec.5.2.1, the kinematical selection optimization is aimed at minimizing the expected statistical uncertainty on a cross section measurement and it relies on an estimate of both the amount of expected b-tagged jets from inclusive Monte Carlo $t\bar{t}$ and multijet background events.

In this section we describe 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 on which the background prediction method rests is that b-tag rates for $t\bar{t}$ signal and background processes differ due to the different characteristics of the b-jets produced by the top quark decays compared to the b-jets arising from QCD and vector boson plus heavy flavor production processes. In that hypothesis, parameterizing the b-tag rates as a function of jet characteristics, in events depleted of signal contamination, will allow to predict the number of b-tagged jets from background processes populating a given kinematically selected region.

The steps needed to successfully finalize this approach can be summarized as follows:

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

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

3. Construct 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 over all jets in the selected events their associated b-tagging probabilities.

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 kinematical selection + $\geq$ 1 SECVTX tag).

In general, the use of tagging rate parameterizations rests on the assumption of a negligible $t\bar{t}$ contamination in the sample in which b-tag rates dependencies are studies. Indeed, a $t\bar{t}$ presence in the sample used to parametrize the tagging rate has a sizable impact in the amount of background tags prediction. This effect is further enhanced also considering the different tagging probabilities associated to top-like events and to multijet data respectively. For this reason, we need to choose as base sample a data region depleted as much as possible of signal. We chose to use the $3(E_T^{L5} \geq 15\ GeV; |\eta| \leq 2.0)$ jet events contained in the TOP_MULTI_JET triggered dataset for the background tagging rate parametrization.

Number of events	3 jets	ϵ_{tag}^{ave}	4 jets	ϵ_{tag}^{ave}
multijet data	879,187	0.065	1,553,525	0.087
expected inclusive $t\bar{t}$	16.88	0.582	182.92	0.743
multijet data, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	2,317	0.108	2,412	0.166
expected inclusive $t\bar{t}$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	6.54	0.593	56.92	0.773

Table 5.1: Multijet and expected inclusive $t\bar{t}$ Monte Carlo events and average number of positive tagged jets per event in the sample used for positive tagging rate parameterizations. Events with 3 and 4 jets are displayed separately. The numbers of observed and Monte Carlo expected events in the sample with $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$ $GeV^{1/2}$ are also shown.

Tab.5.1 shows the number of 3- and 4-jet events in the multijet data sample and the $t\bar{t}$ contamination expected from Monte Carlo assuming the theoretical production cross section of 6.1 pb, corresponding to a top mass of $M_{top} = 178$ GeV/c^2. The number of events in the same jet multiplicities is also provided after a tight cut on the $\not\!\!E_T^{sig} = \not\!\!E_T/\sqrt{\Sigma E_T}$. Besides the number of events, the average number of positive L_{2d} tagged jets per event (ϵ_{tag}^{ave}) is also shown. Even if the TOP_MULTI_JET trigger acts by requiring 4 ($E_T^{raw} \geq 10$ GeV) jets, once we ask the jets to have $E_T^{L5} \geq 15$ GeV and $|\eta| < 2.0$ a sizable part of the dataset is found to contain 3-jet events (879,187 events against 1,553,525 with exactly 4 jets). From Tab.5.1 the relative $t\bar{t}$ contribution is found to be a factor ~ 5 lower in 3-jet than in 4-jet events, regardless of the $\not\!\!E_T$ significance requirement.

Therefore by using 3-, instead of 4-jet events as a baseline for the tagging parameterizations (which would be the natural choice given the trigger requirements), we limit the impact of top-like events in the background estimate, decreasing the top fraction from 1.1×10^{-4} to 2×10^{-5} before any kinematical requirements except the clean up cuts (Sec.4.5).

5.1.1 b-tagging rate parametrization

The b-tagging probability can be defined as the ratio of the number of positive SECVTX tagged jets to the number of taggable jets in the subclass of multijet data events with exactly 3 jets. We define as taggable a jet with at least two good SECVTX tracks (see Sec.3.3.7), $E_T^{L5} \geq 15$ GeV, and $|\eta| < 2.0$.

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. The per-jet tagging probability is found to be mainly dependent on jet characteristics, such as E_T, the number of good quality tracks contained in the jet cone, and the missing E_T projection along the jet direction. The $\not\!\!E_T$ projection along the jet axis, $\not\!\!E_T{}^{prj}$, is defined by:

$$\not\!\!E_T{}^{prj} = \not\!\!E_T \cos\Delta\phi(\not\!\!E_T, jet). \tag{5.2}$$

Fig. 5.1 shows the tagging rate dependence on jet E_T, N_{trk} and $\not\!\!E_T{}^{prj}$ obtained from the subsample of multijet events with exactly three ($E_T^{L5} \geq 15$ GeV, $|\eta| \leq 2.0$) jets. Both

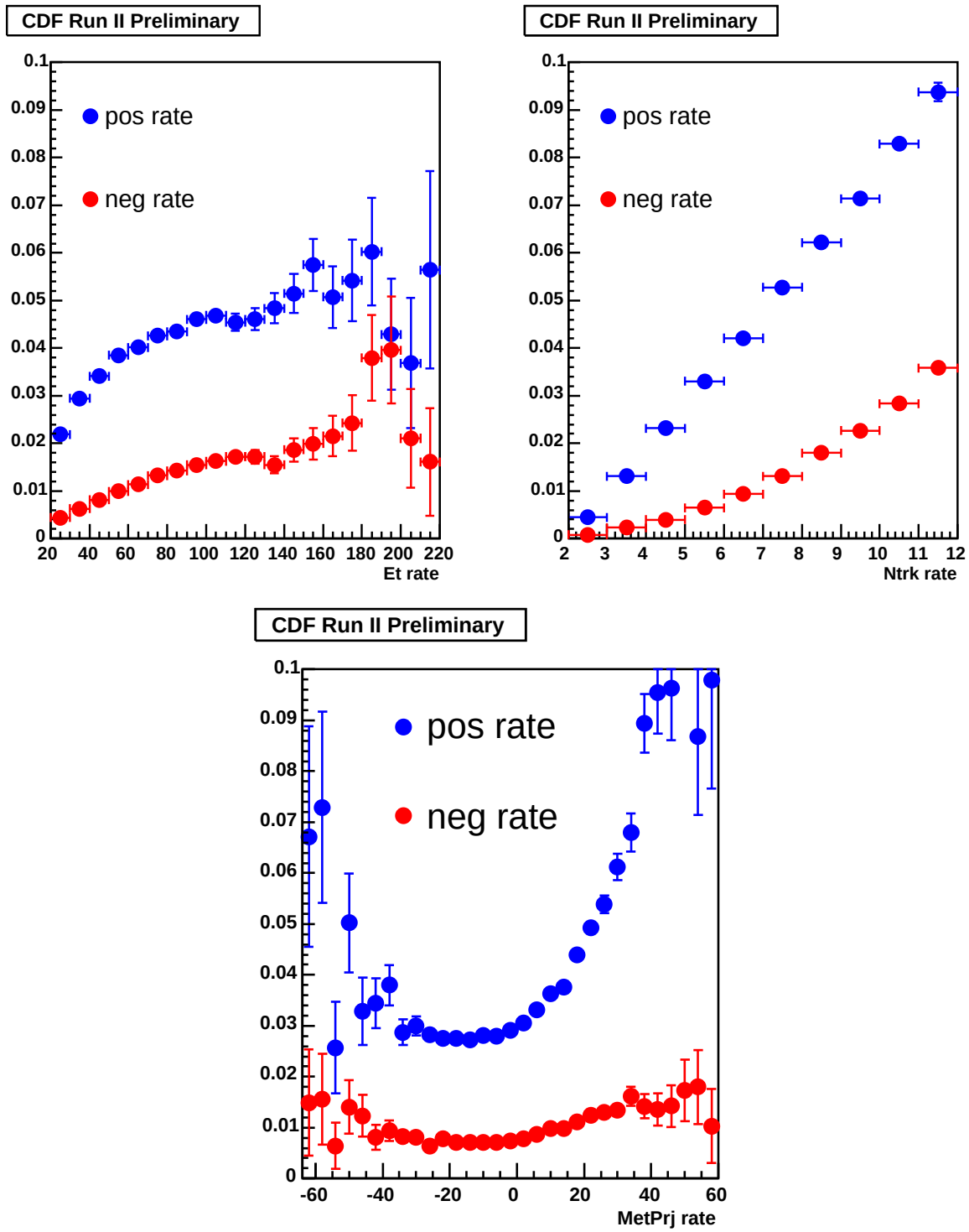


Figure 5.1: Positive and negative b-tagging rates as a function of E_T, N_{trk} and $\not\!\!E_T{}^{prj}$ for the class of 3-jet events in the multijet sample. See the text for details.

the positive and negative[1] tagging rates are shown in the figure.

[1] The SECVTX tagged jets are classified as positive or negative tags on the basis of the secondary vertex

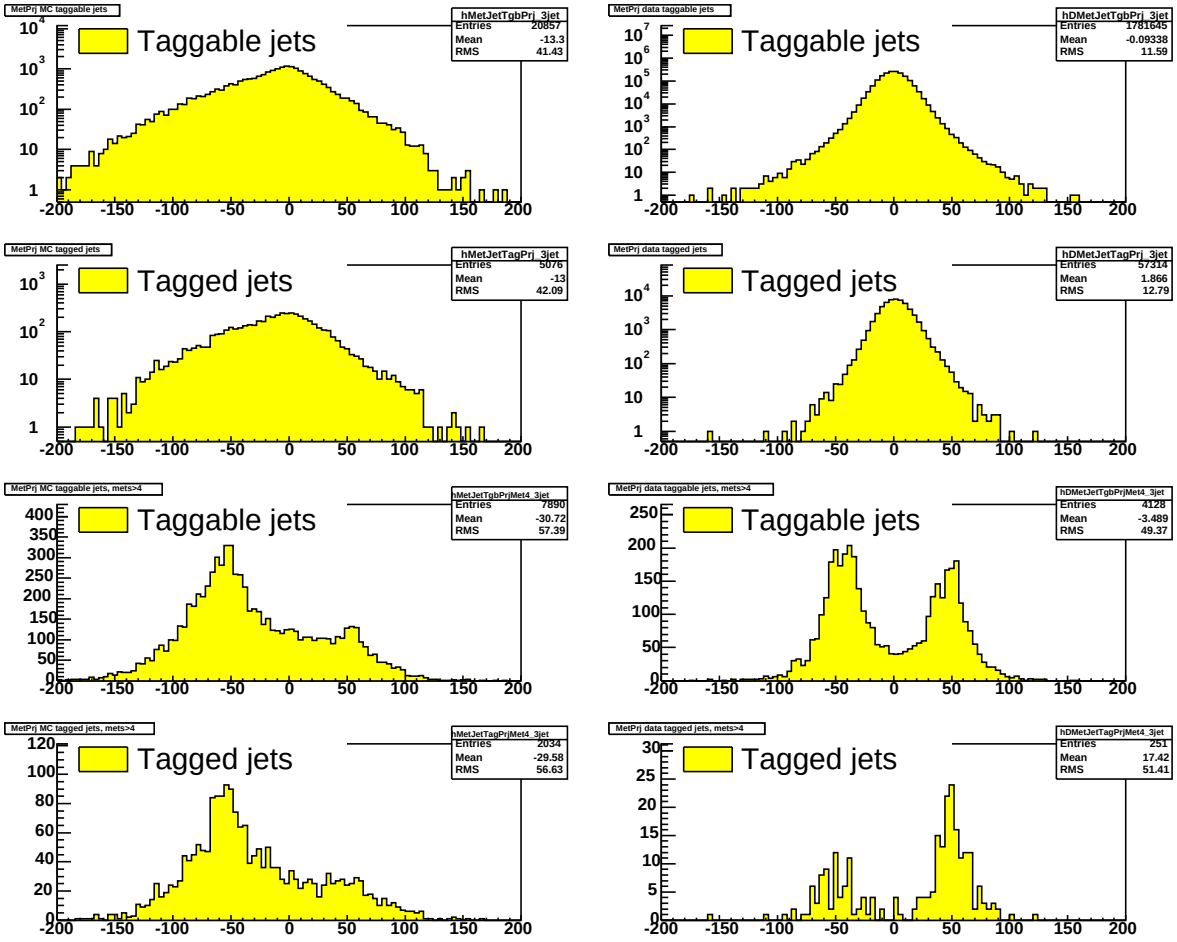


Figure 5.2: $\not\!\!E_T{}^{prj}$ distribution for inclusive Monte Carlo $t\bar{t}$ and data 3-jet events. Top row: left (right) plot log-scale $\not\!\!E_T{}^{prj}$ for taggable jets in $t\bar{t}$ (data). Second row: log-scale missing transverse energy projection for positive tagged jets for both $t\bar{t}$ (left) and data (right). Third row: left (right) plot $\not\!\!E_T{}^{prj}$ for taggable jets in $t\bar{t}$ (data) events satisfying the additional requirement of $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$. Bottom row: left (right) plot $\not\!\!E_T{}^{prj}$ for positive tagged jets in $t\bar{t}$ (data) events satisfying the additional requirement of $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$. See the text for an explanation of the features of these distributions.

Apart from the jet E_T and N_{trk} whose correlation with the tagging probability was expected due implementation details of the b-tagging algorithm (see Sec. 3.3.7), the missing E_T projection along the jet direction is powerful in the parameterization of the b-tag rate, given its consistent correlation with the heavy flavor component of the sample[85, 88], and in particular, given its skill to distinguish the $\not\!\!E_T$ origins in relation to geometrical properties of the events.

While E_T and N_{trk} parameterizations describe well the fake tag contribution to the sample, the $\not\!\!E_T^{prj}$ is found to be able to describe the genuine QCD heavy flavor production

L_{2d} sign (see Sec. 3.3.7).

component. This $\not\!\!E_T^{prj}$ feature is particularly needed since we are going to select the final sample by means of a tight cut on the $\not\!\!E_T^{sig}$ that in turn enrich the sample of QCD b-decays that needs to be accounted for in the background estimate.

In general b-quarks can yield a considerable amount of missing transverse energy due to their semi-leptonic decays (Sec. 4.3.1). In that case the $\not\!\!E_T$ is expected to be aligned with the jet direction. On the contrary, genuine $\not\!\!E_T$ produced in W boson decays stands more likely away from jets, depending on the process-allowed regions of the phases space.

These missing E_T projection features are depicted in Fig. 5.2. The upper left plot of Fig. 5.2 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 plots of the figure, the missing transverse energy projection is drawn for SECVTX positive tagged jets, for both the samples. Before any requirement on the event $\not\!\!E_T$ significance we can note that the $\not\!\!E_T{}^{prj}$ for data is well centered around zero. On the other hand, $t\bar{t}$ events show some kinematical related preference for negative projections (see the enhanced negative tail of the distribution) although the distribution extends itself through high positive values too.

By requiring the events to have large missing E_T significance ($\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$), we reject events whose missing E_T is mainly due to residual energy mis-measurement effects, and in turn concentrate our attention on physics-induced $\not\!\!E_T$. In this context, a sizable change in the distributions previously described is observed, as shown by the third and bottom row of plots in Fig. 5.2. The $\not\!\!E_T^{sig} \geq 4\ GeV^{1/2}$ requirement depletes the region around $\not\!\!E_T{}^{prj}$ equal to zero, and in the case of positive tagged jets, makes evident the different behavior of Monte Carlo (top-like) and data (QCD-like) events as far as this variable is considered: if for $t\bar{t}$ the $\not\!\!E_T$ is more likely oriented in an opposite direction with respect to jets, the opposite behavior is observed for positive tagged jets in data.

$\not\!\!E_T$ projection distributions in QCD samples

The observed behavior, as far as the data $\not\!\!E_T^{prj}$ distribution is concerned, can be explained in terms of $b\bar{b}$ background in our dataset with the help of additional QCD Monte Carlo samples. In particular, $b\bar{b}+4$ light flavor jets ($b\bar{b}+4P$) and 6 light flavor jets ($6\ P$) ALPGEN+HERWIG Monte Carlo samples are considered (Sec. 4.1). The $\not\!\!E_T{}^{prj}$ distributions for both $b\bar{b}+4P$ and 6 P samples are depicted in Fig. 5.3. The missing transverse energy projection along the positive tagged jet, in high $\not\!\!E_T$ significance events, is more likely positive for $b\bar{b}$ background events (Fig. 5.3, bottom left panel) but symmetric with respect to zero as far as six light flavor jet events are concerned (bottom right panel of the figure). Indeed, light flavor jet events are expected to contain tags originated only by misidentifications of the SECVTX tagging algorithm.

Weaker b-tagging rate dependencies

For completeness, Fig. 5.4 shows positive and negative b-tagging rate dependencies on other jet- and event-variables. The missing transverse energy significance, $\not\!\!E_T^{sig}$, the event luminosity, the event aplanarity, the number of good quality vertices, and the jet $|\eta|$ correlations were analyzed.

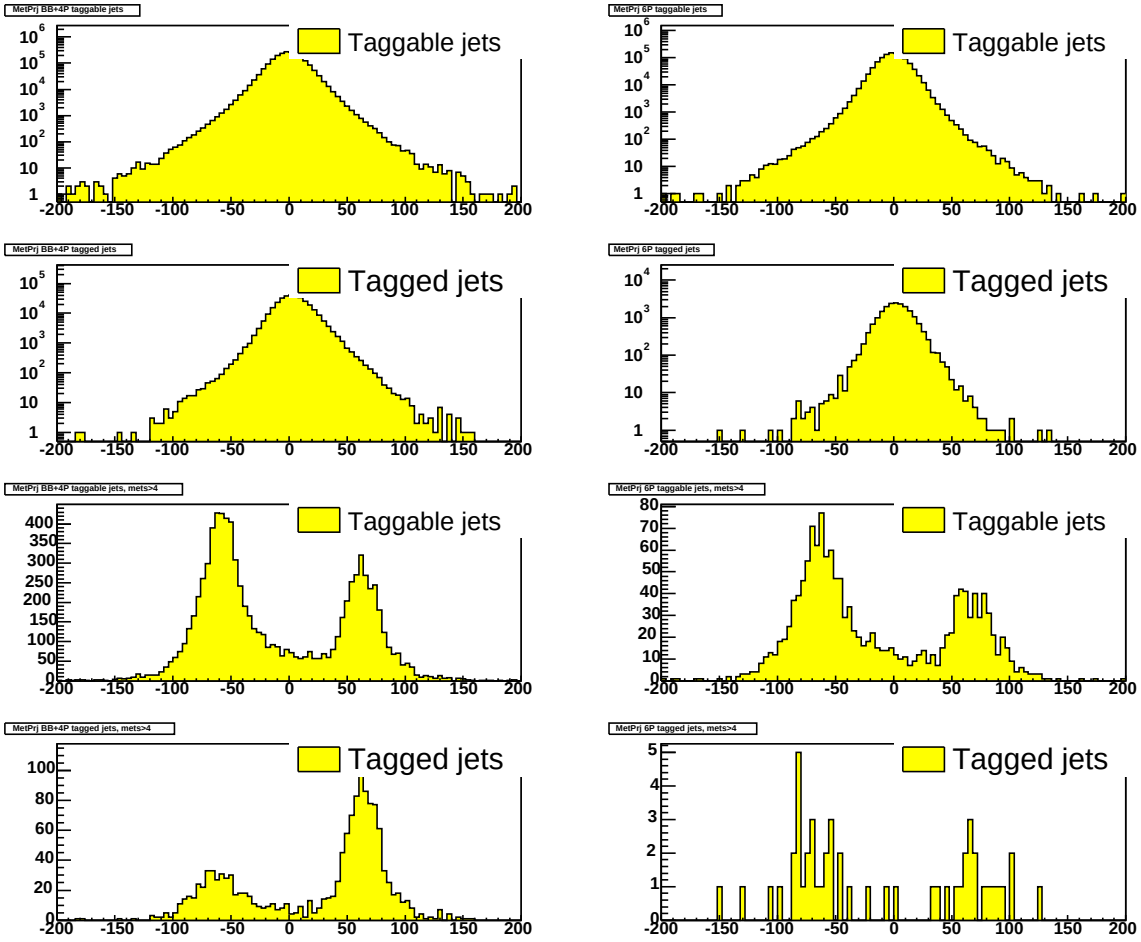


Figure 5.3: $\not\!\!E_T{}^{prj}$ distribution for $b\bar{b}+4P$ and $6P$ ALPGEN+HERWIG Monte Carlo events. Top row: left (right) plot log-scale $\not\!\!E_T{}^{prj}$ for taggable jets in $b\bar{b}+4P$ ($6P$). Second row: left (right) plot log-scale $\not\!\!E_T{}^{prj}$ for positive tagged jets in $b\bar{b}+4P$ ($6P$) . Third row: left (right) plot $\not\!\!E_T{}^{prj}$ for taggable jets in $b\bar{b}$+$4P$ ($6P$) satisfying the additional requirement of $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$. Bottom row: left (right) plot $\not\!\!E_T{}^{prj}$ for positive tagged jets in $b\bar{b}+4P$ ($6P$) satisfying the additional requirement of $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$. See the text for details.

In general, most of the dependencies observed on these variables 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, and the number of good quality vertices. On the other hand, as far as the $\not\!\!E_T^{sig}$ dependence is concerned, it is already accounted for by the $\not\!\!E_T$ projection parameterization. Moreover, jet η is found to be strongly correlated with the number of tracks in the jet (N_{trk}), as shown in Fig. 5.5. Thus, given that the higher the track multiplicity the most central the jet is, 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. The reliability of these choice will be demonstrated in the following sections.

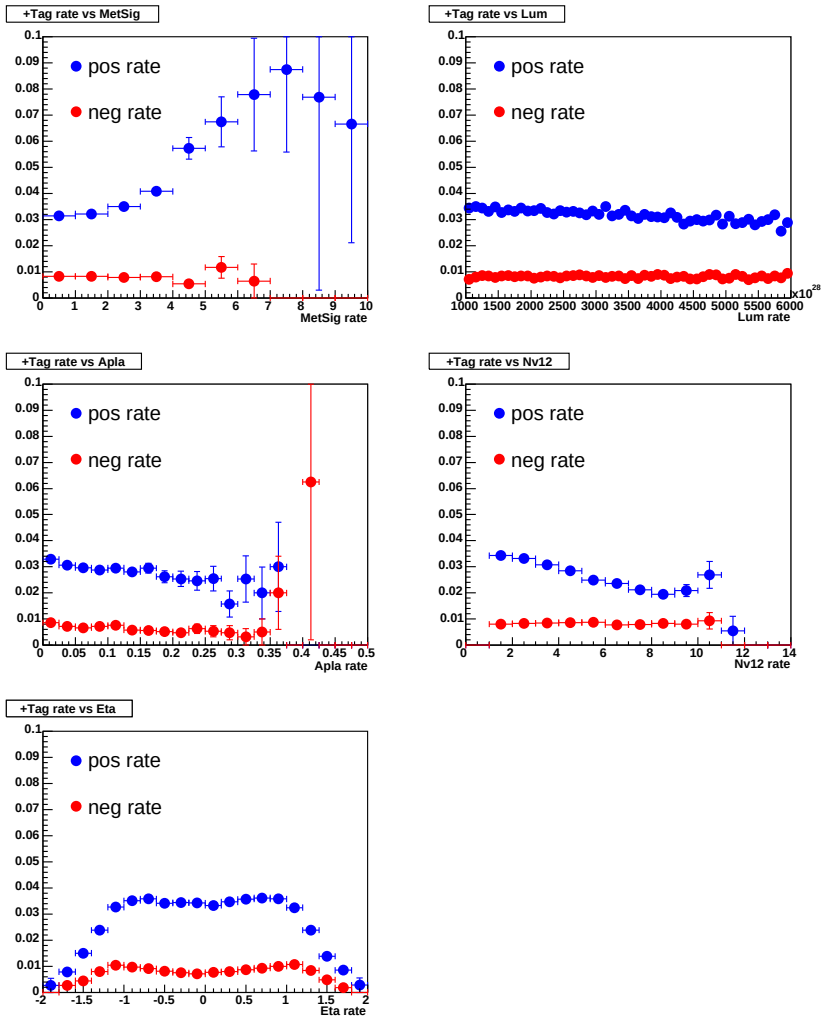


Figure 5.4: Positive and negative b-tagging rates as a function of $\not\!\!E_T^{sig}$ (top left panel), luminosity (top right panel), aplanarity (middle left panel), number of good quality vertices (middle right panel), and jet $|\eta|$ (bottom left panel) in the 3-jet data sample.

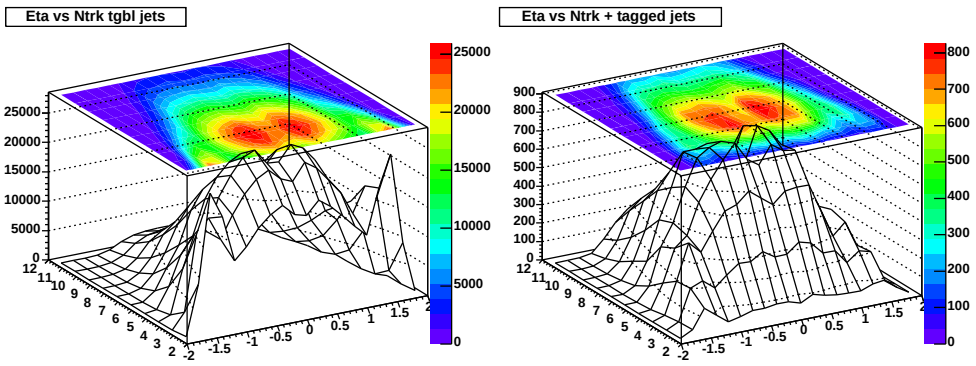


Figure 5.5: Jet η-N_{trk} correlation for taggable jets (left plot) and positive tagged jets (right plot) in the multijet sample.

5.1.2 b-tagging matrix construction

We can define a so-called b-tagging matrix, exploiting the per-jet b-tagging probability dependencies shown in Fig. 5.1, by defining a $3-d$ b-tagging probability

$$\mathcal{P}(E_T^k, N_{trk}^k, \not\!\!E_{Tprj}^{\ k}) \tag{5.3}$$

for each $k-th$ jet in the event. This probability, whose parameterization is defined using 3-jet data events with a negligible $t\bar{t}$ contamination, will allow to calculate the number of background b-tags expected in a given data sample. In particular, the number of expected background b-tags for the $i-th$ event in a given sample, is defined as:

$$N_{tags}^i = \sum_{k=1}^{n} \mathcal{P}(E_T^k, N_{trk}^k, \not\!\!E_{Tprj}^{\ k}), \tag{5.4}$$

where the index k runs over the number of taggable jets in the event. In this way, the total number of tagged jets expected for a given category of events will be the simple sum of the expected tags per each event.

The matrix binning, chosen in order to best suit the tagging rate dependencies shown in Fig. 5.1 and in order to minimize the number of low statistics or undefined matrix bins, is set as follows:

- 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.

5.1.3 Preliminary b-tagging matrix checks

In the previous section the background b-tagging rate parameterizations we adopted in order to construct the b-tagging matrix have been described. Moreover, the matrix binning choice has been justified in order to avoid low statistics or undefined bins. Before applying this parametrization to estimate the number of background b-tagged jets in a given data sample selected by means of kinematical requirements, we need to demonstrate the its capabilities in predicting the right amount of b-tags as well as kinematical distributions for b-tagged events in samples of data before kinematical selection, where the $t\bar{t}$ signal contamination is small.

b-tagging rate extrapolation at high jet multiplicities

The first check consists in extrapolating the b-tagging rate dependencies towards high jet multiplicities, 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 multijet data sample, before any kinematical requirement except the clean-up cuts. Results are shown in Tab.5.2 and Fig. 5.6. The top plot of Fig. 5.6 shows the number of observed and matrix-predicted b-tagged jets as a function of the jet multiplicity of the

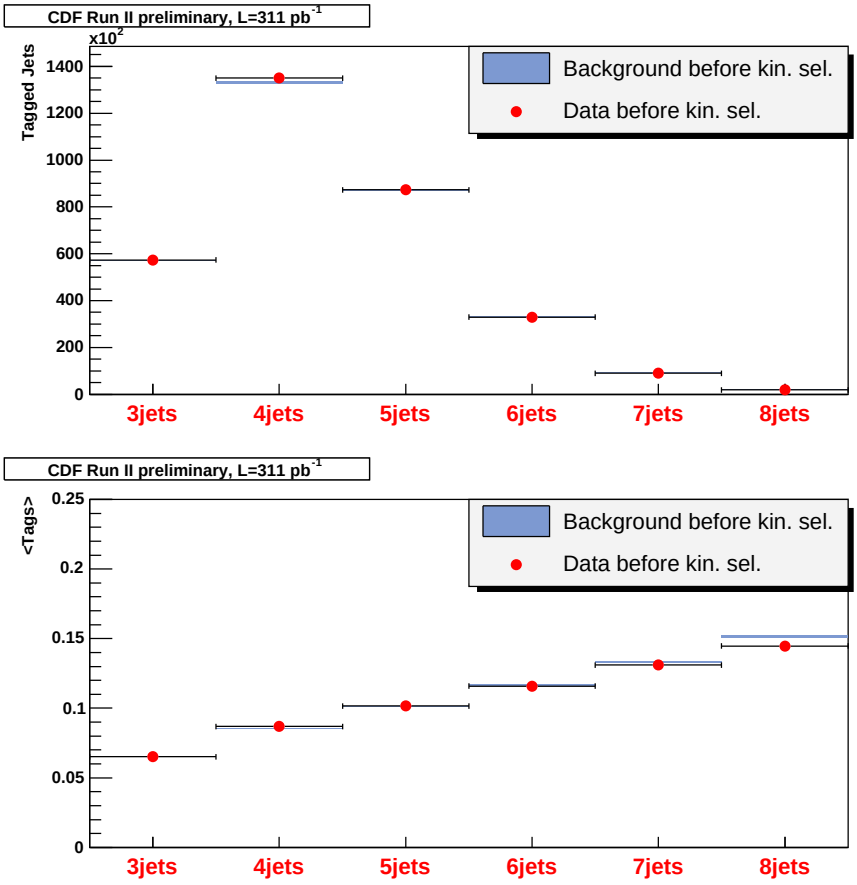


Figure 5.6: Tagging matrix check before any kinematical selection. Top plot: observed and predicted positive b-tags as a function of the jet multiplicity. Bottom plot: average number of observed and predicted b-tags as a function of the jet multiplicity.

N jet	3	4	5	6	7	8
Nevts	879,187	1,553,525	859,543	284,062	68,628	13,237
N tgbl jets	1,781,645	4,197,888	2,877,148	1,134,089	318,410	69,966
Obs + tags	57,314	135,056	87,332	32,914	8,992	1,914
Exp + tags	$57,314 \pm 233$	$133,275 \pm 546$	$87,156 \pm 370$	$33,184 \pm 149$	$9,147 \pm 43$	$2,000 \pm 10$
				Total observed + tags:		323,522
				Total expected + tags:		$322,076 \pm 717$

Table 5.2: Tagging matrix check before any kinematical selection. For each jet multiplicity bin, the number of observed and predicted positive b-tags is shown. Uncertainties are statistical only.

events. On the other hand, the bottom plot of the same figure shows the average number of observed and predicted tags per event, $< tag >$, defined as: $< tag >= N_{tag}/N_{evt}$. The uncertainties both in Tab.5.2 and Fig. 5.6 are statistical only and are derived from the b-tag rate parameterization on 3-jet data events.

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.

Kinematical distributions of matrix-predicted background

Once the extrapolation reliability at higher jet multiplicity is established, we can use the matrix definition in order to construct kinematical distributions and compare them with the observed data distributions for events with at least $N_{jet}(E_T^{L5} \geq 15\ GeV; |\eta| \leq 2.0) \geq 4$ and at least one b-tagged jet before any other kinematical requirements except the clean-up selection.

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

Fig. 5.7 shows the observed and matrix-predicted distribution for jet variables such as E_T (top left panel), η (top right panel), ϕ (middle left panel), $\Delta\phi(\not\!\!E_T, jet)$ (middle right panel) and $\not\!\!E_T^{prj}$ (bottom left panel). The insets on 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 flat for all the jet variables here considered. Exceptions are 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 parametrization.

Besides the jet-related distributions, event spectra can still be obtained from b-tagging matrix. The comparison between observed and matrix-predicted distributions is provided in Fig. 5.8 and 5.9.

Fig. 5.8 provides the comparison between observed and predicted distributions for $H_T = \not\!\!E_T + \sum E_T$, number of good quality vertices, z coordinate of the primary vertex, run number, luminosity and aplanarity. Furthermore, Fig. 5.8 shows these b-tagging matrix checks on event sphericity, centrality, $\sum E_T$, $\sum E_T^3 = \sum E_T - E_T^{jet1} - E_T^{jet2}$, $\not\!\!E_T^{sig}$ and $\not\!\!E_T$. Also in the case of event variables the ratio between observed and expected distributions behaves well, confirming not only the reliability of the tagging matrix in predicting the number of b-tagged jets in a background-dominated sample, but also its effectiveness in describing the kinematical distribution of tagged data.

Further investigation on the reliability of the b-tagging matrix predictions will be provided after having established an optimized kinematical selection. This will be the subject of Sec.5.2.2.

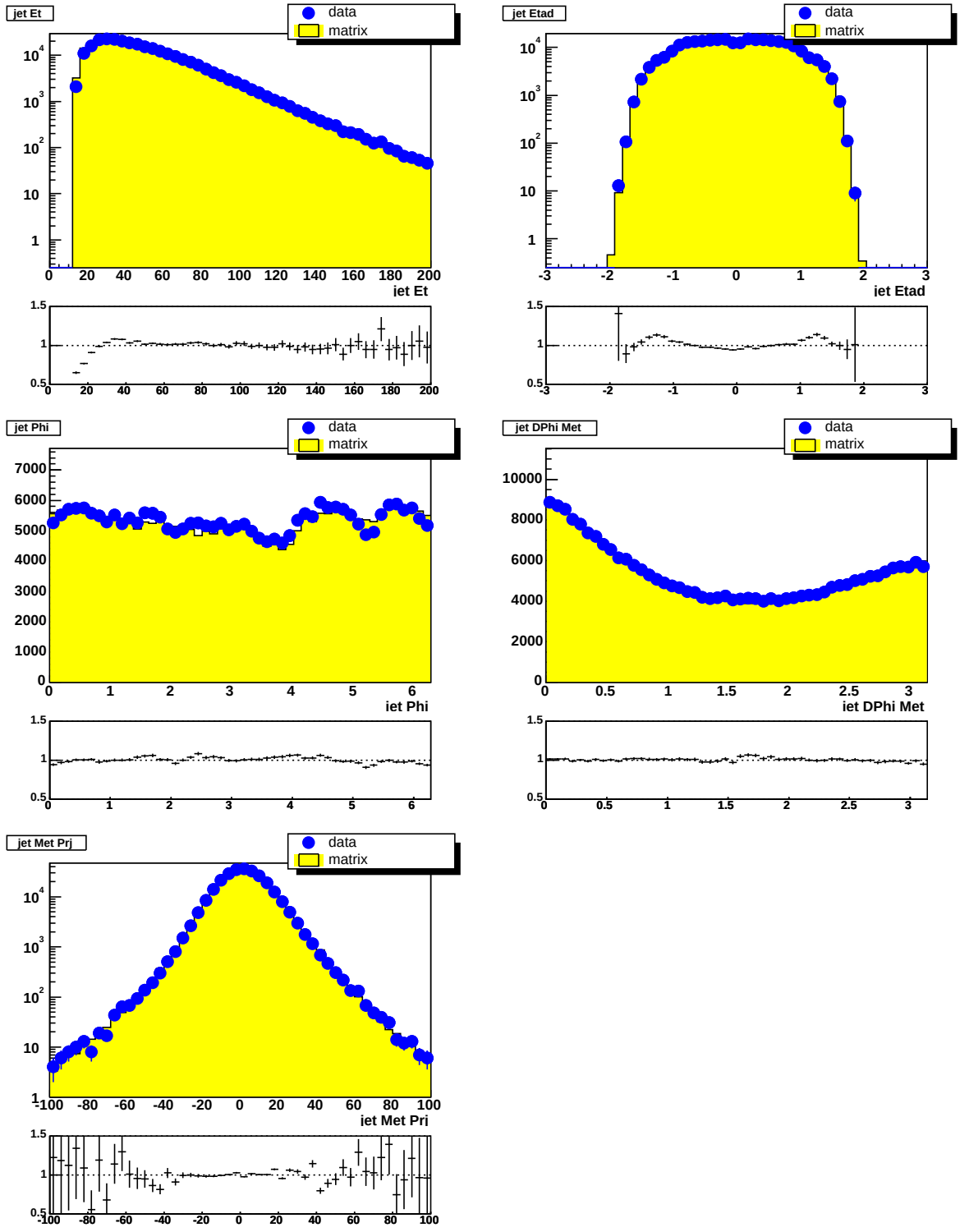


Figure 5.7: Checks of tagging matrix-based jet variables distributions in data events with at least four $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ jets. Jet E_T (top left panel), η (top right panel), ϕ (middle left panel), $\Delta\phi(\not\!\!E_T, jet)$ (middle right panel) and $\not\!\!E_T^{prj}$ (bottom left panel). The insets in the bottom of each panel display the bin-by-bin ratio of observed to matrix-calculated distributions.

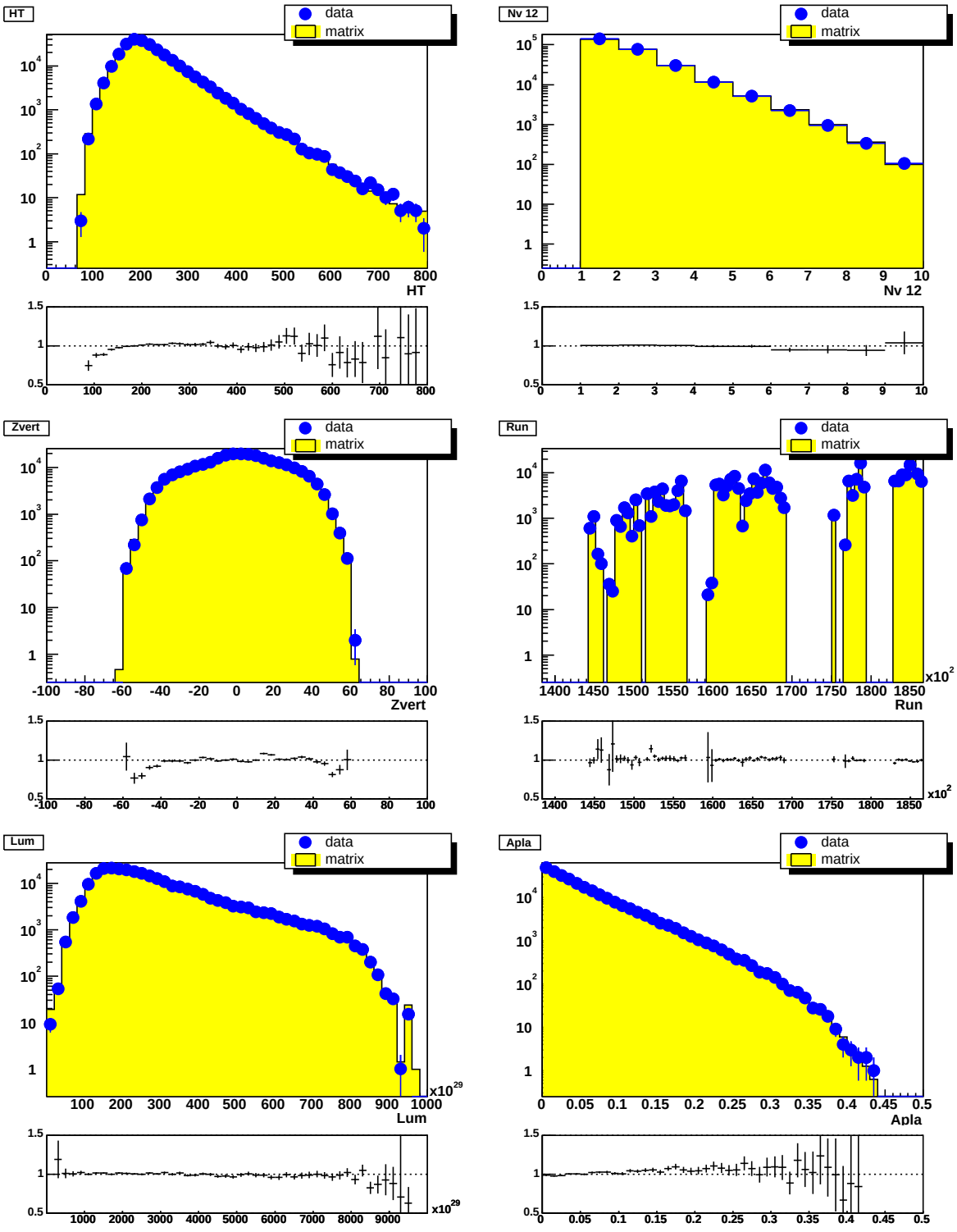


Figure 5.8: Checks of tagging matrix based event variables distributions in data events with at least four $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ jets. H_T (top left panel), number of good quality vertices (top right panel), z coordinate of the primary vertex (middle left panel), run number (middle right panel), luminosity (bottom left panel) and aplanarity (bottom right panel). The insets in the bottom of each panel display the bin-by-bin ratio of observed to matrix-calculated distributions.

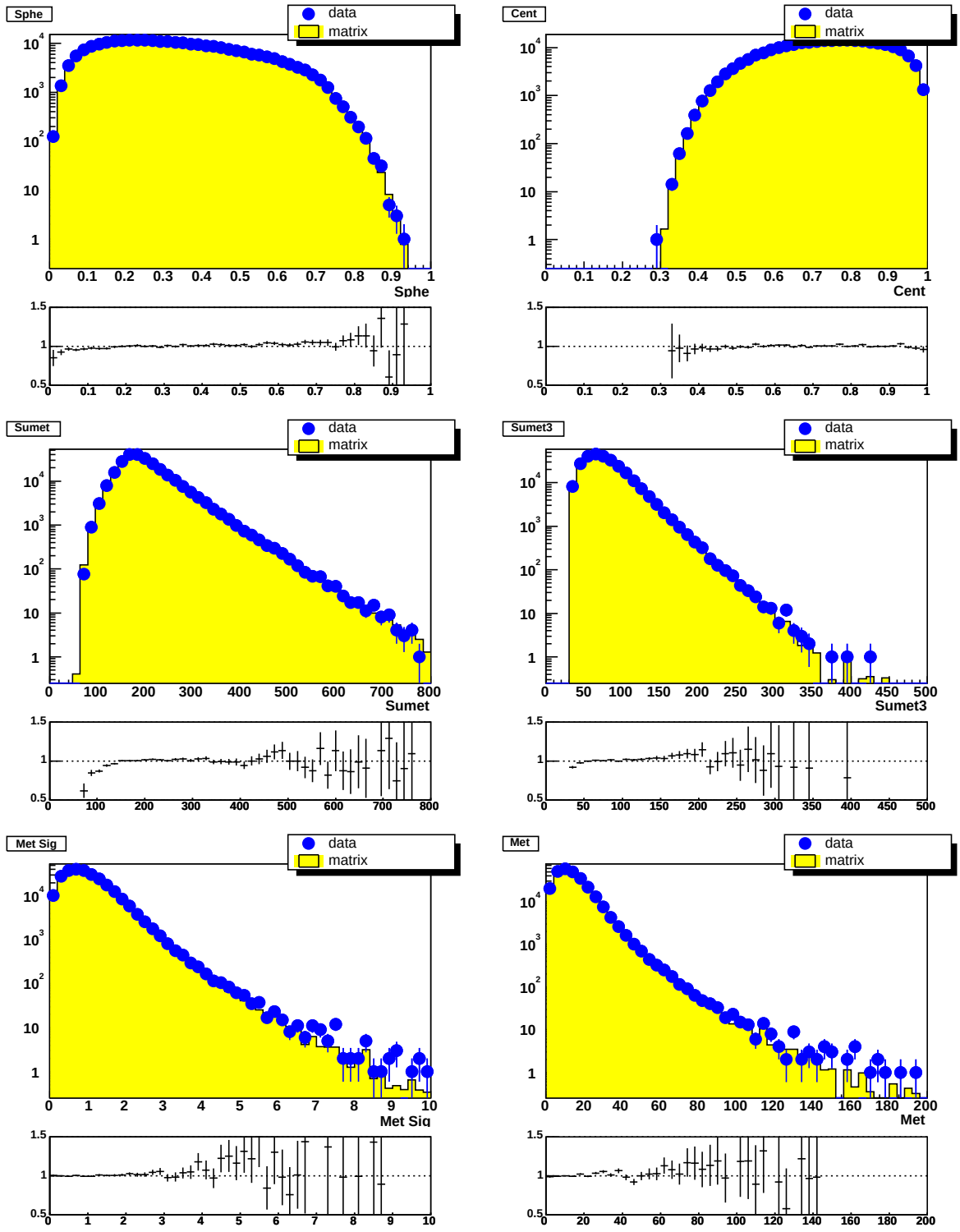


Figure 5.9: Checks of tagging matrix based event variables distributions in data events with at least four $E_T^{L5} \geq 15\ GeV$ and $|\eta| \leq 2.0$ jets. Sphericity (top left panel), centrality (top right panel), $\sum E_T$ (middle left panel), $\sum E_T^3$ (middle right panel), $\not\!\!E_T^{sig}$ $\not\!\!E_T^{prj}$ (bottom left panel) and $\not\!\!E_T$ (bottom right panel). The insets in the bottom of each panel display the bin-by-bin ratio of observed to matrix-calculated distributions.

5.2 Event selection

In this section the optimization procedure we adopted to set the best values of the analysis cuts will be described.

As already highlighted in Sec. 4.3, significant $\not\!\!E_T$ is in general a good indicator for establishing the presence of neutrinos in the event. Given that $\not\!\!E_T$ resolution is observed to degrade as a function of the $\sum E_T$, in the form $\not\!\!E_T{}^{res} = const \cdot \sqrt{\sum E_T}$, the missing E_T significance, defined as $\not\!\!E_T/\sqrt{\sum E_T}$ is used as a selection variable.

On the other hand, the minimum $\Delta\phi$ difference between the $\not\!\!E_T$ and jet direction was demonstrated to provide an extra handle in order to increase the signal to background ratio (Sec.4.3.2). Indeed, while for signal events little correlation between the direction of a given jet and the $\not\!\!E_T$ is expected, the latter quantity being determined by the presence of the neutrino from W boson decay, for background processes the missing E_T is mainly yielded by jet energy mis-measurement and then is aligned with the jet direction.

As described in Sec. 4.3, aside $\not\!\!E_T$ related quantities, b-jet identification provided by the SECVTX algorithm constitutes an effective handle to further 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 kinematical selection, the requirement of at least one positive SECVTX tagged jet.

Additionally to the previous kinematical and b-tagging requirements, a cut on event aplanarity is considered for the event selection.

The optimization procedure we seek is aimed at minimizing the statistical uncertainty on the cross section measurement. In this context, the rejection power of kinematical variables ($\not\!\!E_T^{sig}$, $\min\Delta\phi(\not\!\!E_T, jets)$ and aplanarity) will be considered 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.

5.2.1 Analysis cuts optimization

After the clean up cuts described in Sec. 4.5, 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 contain an intrinsic bias.

The optimization procedure for the event selection definition is performed after the $N_{jets} \geq 4$ requirement and scans different set of cuts on $[\not\!\!E_T/\sqrt{\Sigma E_T},\ A,\ \min\Delta\phi(\not\!\!E_T, jet)]$. Among all possible cut configurations it chooses the one promising the minimum relative statistical error on a cross section measurement.

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

$$\sigma(p\bar{p} \to t\bar{t}) \times BR(t\bar{t} \to \not\!\!E_T + jets) = \frac{N_{obs} - N_{exp}}{\epsilon_{kin} \cdot \epsilon_{tag}^{ave} \cdot \mathcal{L}}, \tag{5.5}$$

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 and kinematical 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, $\mathcal{L}$ is the luminosity of the dataset used.

Using in input to Equation 5.5 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 parametrization in the selected sample, we can estimate the expected cross section value and its relative statistical uncertainty for each cut set. The only missing piece is N_{obs}. We can not 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 kinematical selection, 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 after the application of the given cut set, respectively. Using these values, the statistical uncertainty affecting the measurement can be computed before looking at the "post-tagging" data sample, allowing in this way to choose among multiple set of cuts the one that minimizes the relative error on the cross section measurement.

Operatively, the procedure scans $1,000$ different sets of cuts, obtained by varying each of the variable cuts of the tern ($\not\!\!E_T/\sqrt{\Sigma E_T}$, A, $\Delta\phi_{min}(\not\!\!E_T, jet)$) in different ranges: ten cut values are used for $\not\!\!E_T/\sqrt{\Sigma E_T}$ in the range $2.0 \div 6.5$ $GeV^{1/2}$ with steps of 0.5; A is scanned from 0.00 to 0.09 in steps of 0.01; and finally, $\min\Delta\phi(\not\!\!E_T, jet)$ thresholds are chosen in the range $0.2 \div 1.1$ radiants in steps of 0.1. For each of these selections the following quantities are calculated, as displayed in Tab. 5.3:

- MC^{evt}, and MJ^{evt}: number of inclusive Monte Carlo $t\bar{t}$ and multijet 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 parametrization after the kinematical selection defined by the tern ($\not\!\!E_T/\sqrt{\Sigma E_T}$, A, $\Delta\phi_{min}(\not\!\!E_T, jet)$).
- tag $S/\sqrt{N}$: 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;
- $\frac{\sigma_{xsec}}{xsec}$: relative error on the cross section measurement.

The list of the analyzed set of cuts is then ordered by increasing relative statistical error in the cross section measurement. Results are reported for the top five selections in Tab. 5.3. The cross section uncertainty is calculated using only the statistical errors of the involved quantities.

cut set				MC^{evt}	MJ^{evt}	N_{MC}	N_{exp}	$S/\sqrt{N}$	$\frac{\sigma_{xsec}}{xsec}$
1	4.00	0.00	0.40	93	597	73	67 ± 3	8.91	17.47%
2	4.00	0.00	0.50	82	461	65	46 ± 3	9.55	17.48%
3	4.00	0.01	0.40	90	549	71	63 ± 4	8.93	17.60%
4	4.00	0.01	0.50	80	426	63	43 ± 3	9.60	17.60%
5	4.00	0.02	0.40	85	494	68	58 ± 3	8.91	17.84%
...	...	...	...	...	...	...	...	...	...

Table 5.3: Different possible kinematical selections ordered by increasing relative error on the cross section measurement, and their figures of merit ($S/\sqrt{N}$ and $\sigma_{xsec}/xsec$), for different choices of the cuts for $\not\!\!E_T/\sqrt{\Sigma E_T}$, A, $\Delta\phi_{min}(\not\!\!E_T, jet)$. For each cut set the number of events in the pre-tagging sample and the number of expected b-tagged jets is shown for Monte Carlo and data samples, respectively. The amount of tags in the data is obtained from the tagging rate parametrization as described in Sec. 5.1.1

Best cut values

The final result of this procedure sets as the best event selection cuts the following:

- $N_{jet}(E_T^{L5} \geq 15\ GeV, |\eta| \leq 2.0) \geq 4$;
- $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$;
- $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4\ rad$;

promising a relative statistical cross section uncertainty of 17.5%. The pre-tagging combined kinematical efficiency of trigger, event clean-up, and selection cuts on $t\bar{t}$ inclusive events is measured to be $\epsilon_{kin} = 4.878 \pm 0.021\%$, where the uncertainty is statistical only. The average number of tags per $t\bar{t}$ event under these selections is found to be $\epsilon_{tag}^{ave} = 0.7889 \pm 0.0018$, and is determined by dividing the number of b-tagged jets in the kinetically selected sample ($N_{tag} = 39,326$) by the number of inclusive $t\bar{t}$ events surviving the selection ($N_{evt} = 49,848$). The ϵ_{kin} and ϵ_{tag}^{ave} values will be used for the cross section measurement as it will be described in Chapter 7.

A cut on aplanarity is found to be unnecessary by the optimization procedure. Fig. 5.10 shows the $\not\!\!E_T/\sqrt{\Sigma E_T}$, $\min \Delta\phi(\not\!\!E_T, jet)$ and aplanarity distributions as extracted from inclusive Monte Carlo $t\bar{t}$ events with at least one positive b-tagged jets and from tagging rate parametrization on multijet data with $N_{jet} \geq 4$. The $\min \Delta\phi(\not\!\!E_T, jet)$ distribution is provided after the missing E_T significance selection, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$; on the other hand the aplanarity distribution is provided after both the $\not\!\!E_T^{sig}$ and $\min \Delta\phi(\not\!\!E_T, jets)$ cuts specified by the best kinematical selection. In this case, background and signal distributions are very similar, preventing any net improvement in the selection.

The impact of the optimal kinematical selection on both inclusive $t\bar{t}$ and multijet data events is shown in Tab. 5.4. With this set of cuts we end up in the pre-tagging sample with 93 events expected from $t\bar{t}$ inclusive events against 597 observed events in the multijet data sample, providing an estimated signal over background ratio of $S/N \approx 1/5$ that

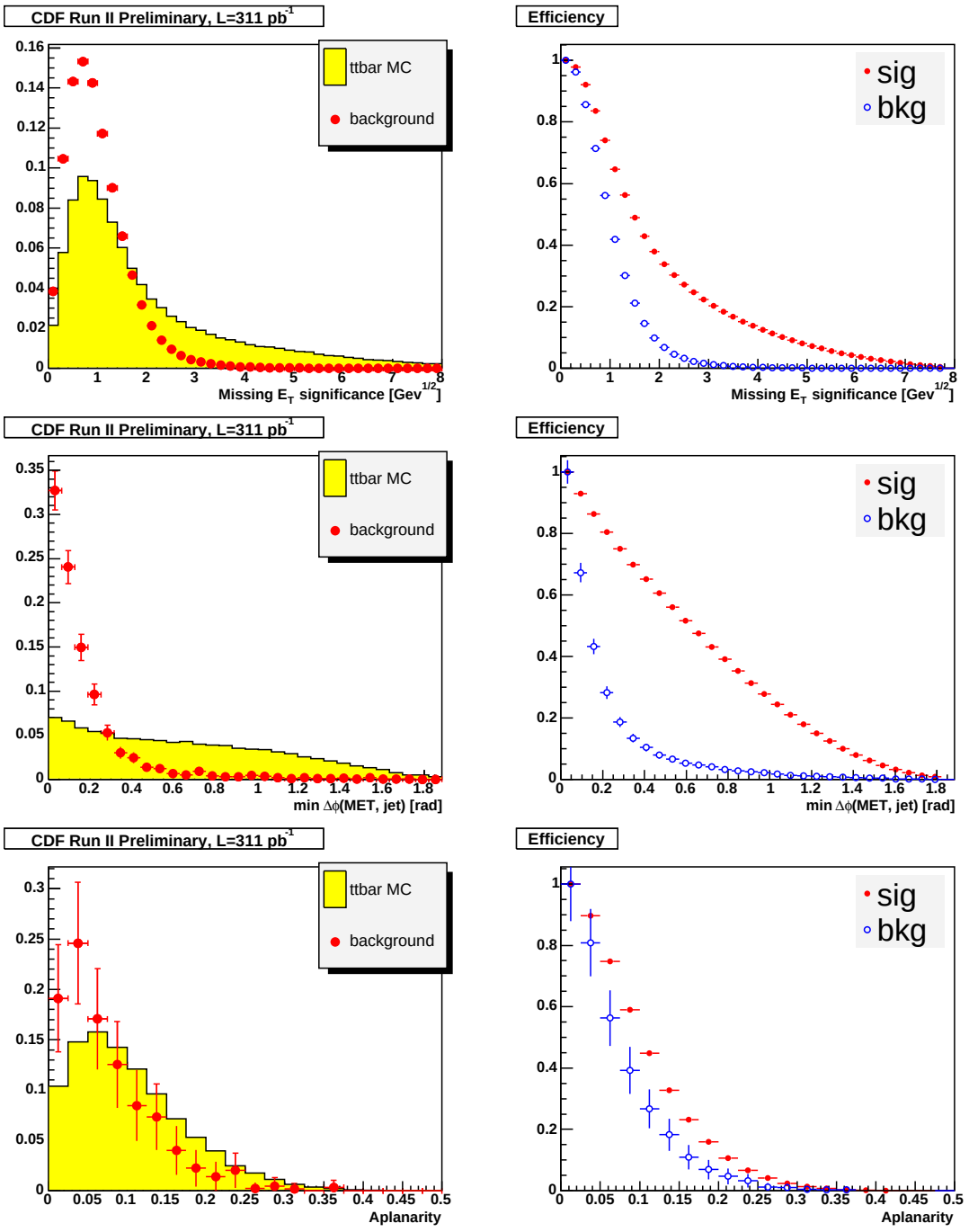


Figure 5.10: $\not\!\!E_T/\sqrt{\sum E_T}$ (top panel), $\min\Delta\phi(\not\!\!E_T, jet)$ (middle panel) and aplanarity (bottom panel) distributions for inclusive $t\bar{t}$ events (filled histograms) and tagging-matrix predicted distributions as extracted from data events. The right plots show the relative efficiencies of a cut $x > x_{cut}$ on the corresponding variable x for both signal and background. See the text for details.

will be further improved by the additional requirement for at least one SECVTX positive tagged jet.

N evt	$MC_{Incl.}$	MJ
Total	1,021,924	4,249,644
Prereq	558,528	3,897,755
$N_{jet} \geq 4$	549,138	2,781,788
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	78,145	3,996
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	49,848	597
in 311 pb^{-1}:	92.63	597
S/N Inclusive: $N_{mc}^{incl}/(N_{obs}^{data} - N_{mc}^{incl})$		$= 0.18$

Table 5.4: Effect of the kinematical selection $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$ on $t\bar{t}$ inclusive events as well as on multijet data sample. The row labeled as Prereq contains the clean up cuts: trigger simulation (for Monte Carlo events only), number of tight leptons = 0, $|Z_{jvert}| < 60\ cm$, $|Z_{jvert} - Z_{pvert}| < 5.0\ cm$ and $n_{good\ quality}^{vertices} \geq 1$.

N evt $t\bar{t}$	$\tau + jets$	ϵ_{cut} [%]	$e + jets$	ϵ_{cut} [%]	$\mu + jets$	ϵ_{cut} [%]
BR	149,323	–	149,225	–	149,159	–
Trigger+ Good Run	82,200	55.05	98,879	66.26	57,086	38.27
Lepton Veto	78,084	94.99	58,863	59.53	42,008	73.59
Vertex Req.	74,904	95.93	54,889	93.25	39,993	95.20
$N_{jet} \geq 4$	72,708	97.07	53,039	96.63	38,047	95.13
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	29,830	41.03	20,286	38.25	19,422	51.05
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	19,079	63.96	13,797	68.01	13,905	71.59
Total eff wrt to BR		12.77		9.25		9.32

Table 5.5: Effect of the clean up and kinematical selections for $\tau/e/\mu + jets$ exclusive $t\bar{t}$ events. For each cut the efficiency with respect to the previous selection is provided for each $t\bar{t}$ lepton+jets decay channel.

Event selection acceptances for $e/\mu/\tau$ + jets $t\bar{t}$ events

The beauty of this selection is that it permits to achieve a good S/N regardless of any lepton identification procedure. The $\not\!\!E_T^{sig}$-based selection here described allows the $t\bar{t}$ isolation by means of 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 $\tau + jets$, $e + jets$ and $\mu + jets$ decay modes.

Tab. 5.5 shows the impact of the kinematical selection and of the clean up cuts on exclusive $\tau + jets$, $e + jets$ and $\mu + jets$ $t\bar{t}$ Monte Carlo events respectively. Despite the well-identified high-P_T lepton veto imposed for e and μ, the final acceptance provided by the kinematical selection is quite similar for all the $l + jets$ decays. The efficiencies calculated with respect to the number of events for each decay mode (provided by the first row of Tab. 5.5) of all requirements is found to be $12.77 \pm 0.09\%$, $9.25 \pm 0.07\%$, and $9.32 \pm 0.07\%$ for $\tau + jets$, $e + jets$ and $\mu + jets$ channels, respectively. These features can

be explained by comparing the effect of each of the selection cuts applied separately for $\tau/e/\mu + jets$ events:

- The clean-up selection contains the tight lepton veto, and trigger as well as vertex requirements. The efficiency of all the clean-up cuts together with respect to the initial number of events is different for the three lepton plus jets decay modes. This is mainly due to the lepton veto and trigger simulation effects: the lepton veto decreases the selection efficiency for both e and μ plus jets events. In the case of $\mu + jets$ $t\bar{t}$ decays, extra signal acceptance is lost due to trigger requirements (the muons does not make any jet). This is well described in the first four rows of the table where the effect of each of these requirements is displayed separately.

- The $N_{jet} \geq 4$ requirement does not differentiate the $t\bar{t}$ decay modes after the clean-up selections. The efficiency of the N_{jet} requirement with respect to the previous cut is indeed found to be: $97.07 \pm 0.06\%$, $96.63 \pm 0.07\%$, and $95.13 \pm 0.07\%$ for $\tau/e/\mu + jets$ events, respectively.

- 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 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 $\min \Delta\phi(\not\!\!E_T, jets)$ selection slightly suppresses the $\tau + jets$ decays, given that the neutrino from the τ decay is in general aligned to the τ (jet) direction[89]. In the case of $\tau + jets$ decays, the $\not\!\!E_T$ results from the sum of the missing energies produced by the neutrinos from W boson and τ decays.

At the end of the game, despite the described peculiarities, the 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 kinematical selection according to the 311 pb^{-1} data luminosity, we expect about 35/25/25 events from $\tau/e/\mu + jets$ decays, respectively, out of the 93 expected in total. The remaining signal acceptance arises from di-lepton and all-hadronic $t\bar{t}$ signatures.

5.2.2 Positive tagging matrix predictions in control samples

As anticipated in Sec. 5.1.3, the kinematical selection definition allows the definition of additional data control samples in which to further check the b-tagging rate parametrization for the background.

Once the kinematical selection is defined, its cuts can be used in order to construct control samples close to the signal region where to compare the number of observed positive tags to the number of predicted tags derived from the tagging rate parametrization applied to data. The construction of these additional samples allows to demonstrate the

tagging matrix capabilities to account for background b-tagged jets even in data samples in which the relative contribution of different background processes is changed by means of the selection cuts.

Control samples, as depicted in Fig. 5.11(a), are defined as follows:

1.) multijet data before any kinematical selection[2];

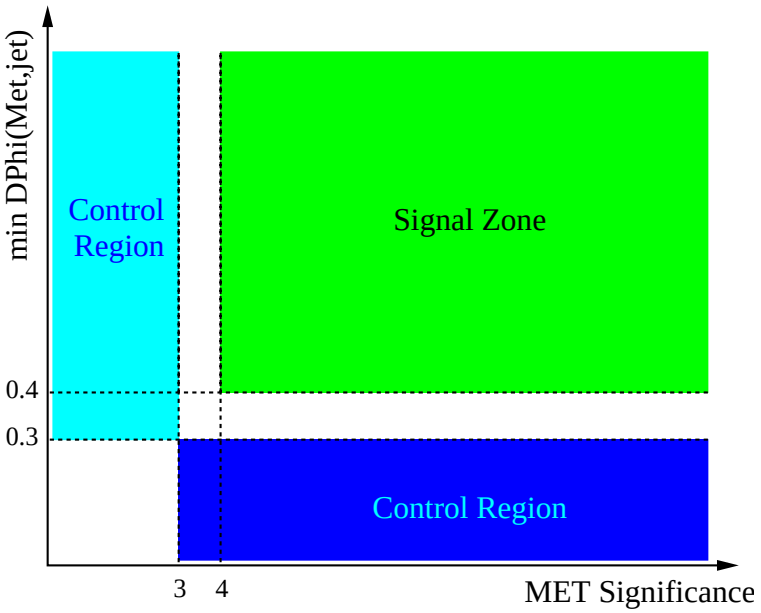
2.) multijet data with $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jet) \leq 0.3$;

3.) multijet data with $\not\!\!E_T/\sqrt{\Sigma E_T} \leq 3\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jet) \geq 0.3$.

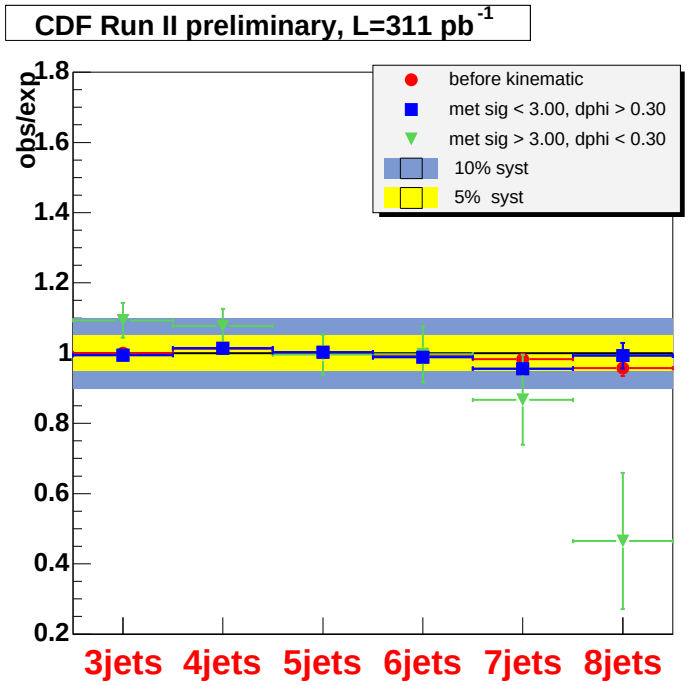
Fig. 5.11(b) shows the performance of the tagging matrix in the control samples previously defined in terms of the ratio of observed to matrix-predicted positive SECVTX tags. In the figure, light and dark shaded areas indicate the 5% and 10% discrepancy bands, respectively.

For the data sample before the kinematical selection, the agreement is good as already mentioned in Sec. 5.1.3, any discrepancy being limited at the level of few percent. The same situation is observed as far as the low missing E_T significance sample is concerned. On the other hand a decreasing trend is observed in the high $\not\!\!E_T^{sig}$ control sample especially at high jet multiplicity (7- and 8-jet events). Some investigations on this behavior will be described in the following.

[2] This sample was already used for the preliminary tagging rate parameterization check in Sec. 5.1.3



(a)



(b)

Figure 5.11: Signal and control regions definition in the $\not\!\!E_T/\sqrt{\Sigma E_T}$, $\min \Delta\phi(\not\!\!E_T, jet)$ plane (a). Observed to matrix predicted positive tagged jets ratio in the control samples (b): 1) multijet data before any kinematical selection, 2) multijet data $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jet) \leq 0.3$; 3) multijet data $\not\!\!E_T/\sqrt{\Sigma E_T} \leq 3\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jet) \geq 0.3$. Light and dark shaded areas indicate the 5% and 10% discrepancy bands.

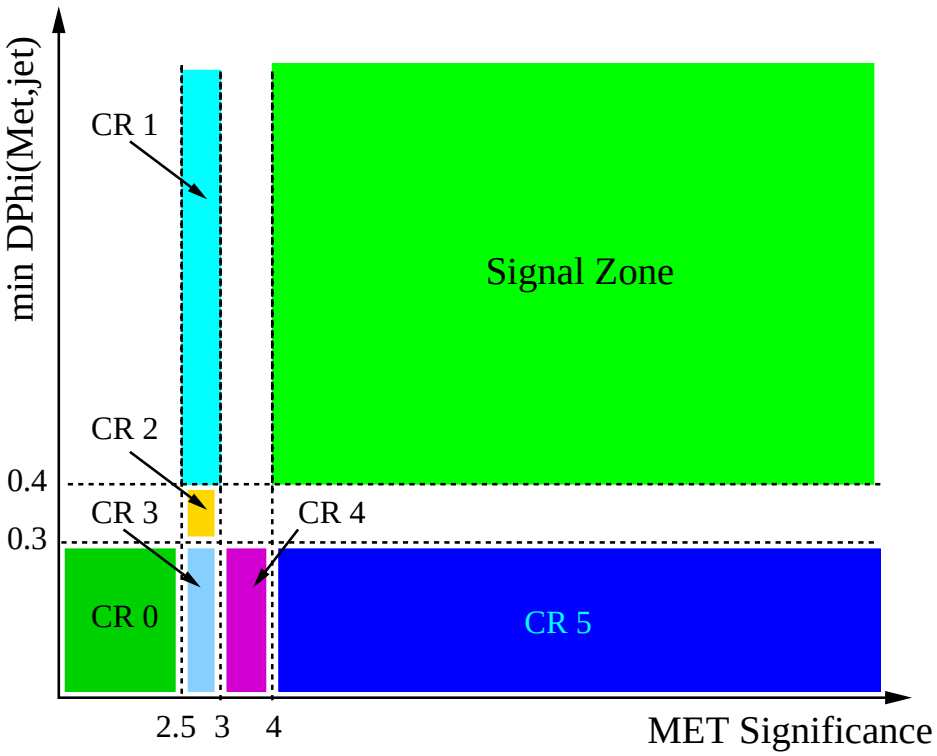


Figure 5.12: Additional control region definitions.

Further checks for the high $\not\!\!E_T^{sig}$ control sample

Fig. 5.11(b) depicts a disagreement between observed and expected b-tags, in the high $\not\!\!E_T^{sig}$ sample, mainly concentrated at high jet multiplicities. Even if the high $\not\!\!E_T^{sig}$ sample statistics is really poor in the 8-jet bin, 6 and 13 being the number of observed and predicted tags, respectively, we further investigated the tagging matrix prediction in several other regions. The purpose is to highlight any possible pathology affecting our parameterization when applied to high missing E_T significance events or to relate the observed trend to simple statistical fluctuations in the tagged data.

The additional control samples, depicted in Fig. 5.12, are defined as follows:

- region-0: $\not\!\!E_T^{sig} < 2.5\ GeV^{1/2}$; $\min \Delta\phi(\not\!\!E_T, jets) < 0.3$;
- region-1: $2.5\ GeV^{1/2} \leq \not\!\!E_T^{sig} \leq 3.0\ GeV^{1/2}$; $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$;
- region-2: $2.5\ GeV^{1/2} \leq \not\!\!E_T^{sig} \leq 3.0\ GeV^{1/2}$; $0.3 \leq \min \Delta\phi(\not\!\!E_T, jets) < 0.4$;
- region-3: $2.5\ GeV^{1/2} \leq \not\!\!E_T^{sig} \leq 3.0\ GeV^{1/2}$; $\min \Delta\phi(\not\!\!E_T, jets) < 0.3$;
- region-4: $3.0\ GeV^{1/2} \leq \not\!\!E_T^{sig} < 4.0\ GeV^{1/2}$; $\min \Delta\phi(\not\!\!E_T, jets) < 0.3$;
- region-5: $\not\!\!E_T^{sig} \geq 4.0\ GeV^{1/2}$; $\min \Delta\phi(\not\!\!E_T, jets) < 0.3$;

The ratio between observed and expected b-tags is provided for each control region in Fig. 5.13.

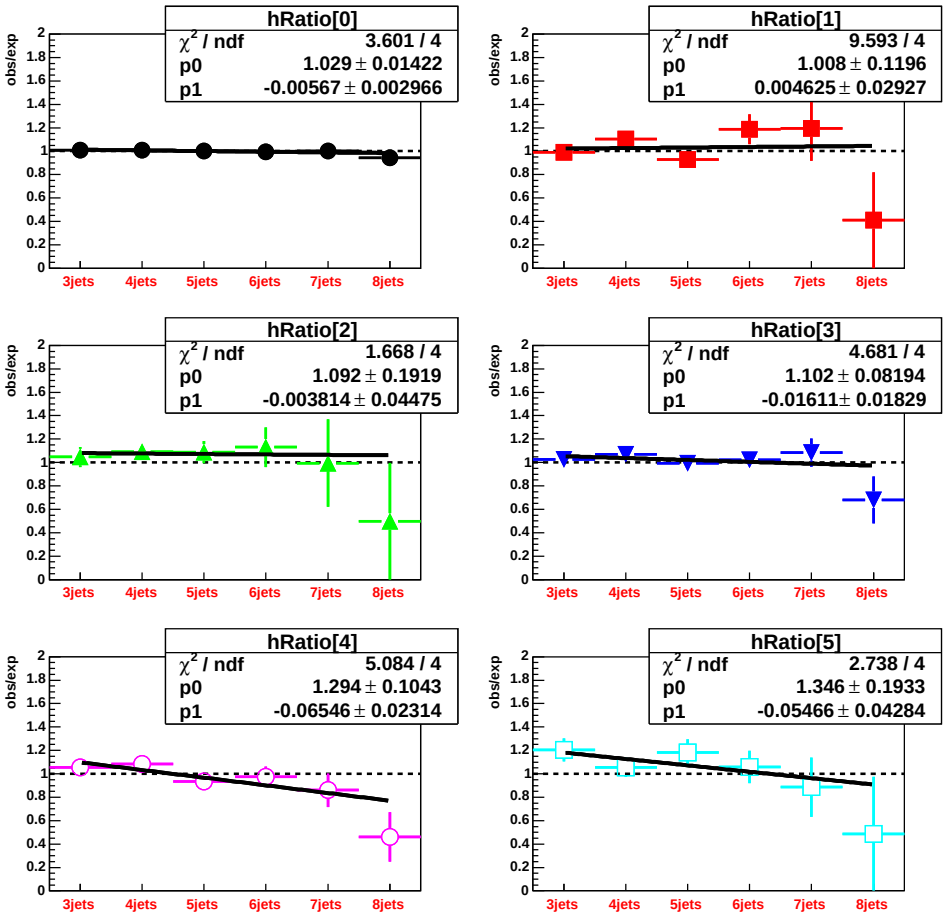


Figure 5.13: Observed to expected number of b-tags ratio in the additional control samples.

In region-0, 1, 2, 3 no decreasing trends in term of obs/exp ratio is observed as a function of jet multiplicity, as demonstrated by the polynomial fit slope parameters and their associated errors. On the other hand, the fit to the obs/exp ratio as a function of N_{jet} in control region-4 returns a significant slope, not compatible with zero. The same tendency is apparent in region-5 although parameter errors are large there.

The behavior of the observed over expected b-tags ratio in these additional control samples suggests that the observed decreasing trend in the control data sample $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jet) \leq 0.3$ (which results from the sum of control Region-4 and 5), is mainly due to a statistical fluctuation in Region-4, rather than real pathologies connected to the b-tagging rate parameterization.

To be totally fair, the structure of the obs/exp ratio as a function of the jet multiplicity, in the high $\not\!\!E_T^{prj}$ control sample, could in some sense indicate a degraded effectiveness of the missing E_T projection parameterization on very high jet multiplicity events. In high jet multiplicity events, the $\not\!\!E_T^{prj}$ is more likely to have large values, since there will always be a jet close in $\Delta\phi$ to the $\not\!\!E_T$. Indeed the more jets an event has, the higher the $\not\!\!E_T$ projection among at least one of the jet directions. Now, since the background b-tagging rate is observed to increase non-linearly as a function of the $\not\!\!E_T^{prj}$ (Fig. 5.1), an

N jets	3	4	5	6	7	8
bkg b-tags	32.68 ± 1.14	37.53 ± 1.76	21.44 ± 1.74	7.27 ± 0.56	0.75 ± 0.08	0.45 ± 0.98

Table 5.6: Number of expected b-tags in the signal region as a function of the event jet multiplicity. Uncertainties are statistical only.

artificially high value of the $\not\!\!E_T$ projection, due to the large number of jets, can results on an overestimated number of background b-tags. The effect is particularly amplified on high $\not\!\!E_T^{sig}$ events. Anyway, even by considering the observed trend as due to this effect, we can estimate its actual impact on the background prediction in the signal region by folding it with the number of expected b-tags (Tab. 5.6).

In particular the real impact of the observed trend in the high $\not\!\!E_T^{sig}$ sample on the background prediction used for the cross section measurement can be estimated as:

$$\frac{\int N_{exp} dN_{jet} - \int N_{exp} F(N_{jet}) dN_{jet}}{\int N_{exp} dN_{jet}} = 3.8\%, \tag{5.6}$$

where N_{exp} is the number of expected b-tags in the signal region for a given jet multiplicity and $F(N_{jet})$ is a functional form describing the observed obs/exp trend. This calculation allows to estimate a background prediction reliability at the level of few percent.

Moreover, the plots in Fig. 5.13 demonstrate that in the region close to the signal zone, in particular region-1 and region-5, the matrix prediction is well under control, especially in the 4-, 5- and 6-jet bins where we expect the signal to be present. For this reason, we can conclude that, even assuming a loss of effectiveness of the $\not\!\!E_T^{prj}$ in events with more than six jets, the overall background prediction is strongly reliable for the jet multiplicities typical of the $t\bar{t}$ production.

Background prediction systematic uncertainty assignment

From the studies performed in Sec. 5.1.3 and those provided above, a systematic uncertainty on the background prediction can be derived.

Considering the obs/exp b-tag ratio as a function of the jet multiplicity in the control sample of data before kinematical selection, with $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 3\, GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jet) \leq 0.3$ requirements and finally with $\not\!\!E_T/\sqrt{\Sigma E_T} \leq 3\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jet) \geq 0.3$, folded with the actual number of events populating each jet multiplicity, the overall discrepancy between observed and matrix predicted number of b-tags can be quoted very conservatively at 10%. 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.

Chapter 6

Characterization of the final data sample

6.1 The final sample

The optimized kinematical selection described in Sec. 5.2.1 allows to isolate a pre-tagging data sample in which the $t\bar{t} \to \not\!\!E_T + jets$ signal is estimated to contribute with signal to background ratio $S/N \sim 1/5$. The background contamination in the sample can be further limited by requiring explicitly the events to contain at least one positive SECVTX tagged jet[1]. 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. As shown in Tab. 6.1, the application of the additional requirement of at least one b-tagged jet, on the kinematically selected sample, increases the signal to background ratio to $S/N \gtrsim 1$, 56 out of 106 being the number of expected inclusive $t\bar{t}$ events in the selected multijet data sample.

The amount of signal b-tagged jets for each $t\bar{t}$ decay modes (assuming the theoretical production cross section value of 6.1 pb for $M_{top} = 178\ GeV/c^2$), the number of background b-jets estimated using the tagging rate parametrization, and the actual number of b-jets observed in the selected data sample are shown in Tab. 6.2, as a function of the jet multiplicity of the events. The main $t\bar{t}$ decay channel contributors to the expected signal are the $\tau+jets$, $e+jets$, and $\mu+jets$ ones, where as far as $e/\mu+jets$ modes are concerned the lepton is failing the standard identification requirements. The $\tau + jets$ contribution is more than one third of the total $t\bar{t}$ expectation.

Fig. 6.1 displays graphically the results of Tab. 6.2: the predicted amount of background b-tags after the complete kinematical selection is shown by the gray histogram together with the expected $t\bar{t}$ contribution for the exclusive $\tau + jets$ and inclusive channels (green and red lines); the observed positive tags in the data are shown by dots. A good agreement between observed and predicted background tags is observed in the 3-jet bin, where the tagging matrix is computed before the kinematical selection, while on the contrary for 4 to 6 jet bins the addition of Monte Carlo inclusive contribution is required

[1] The kinematical selection optimization indeed maximizes in some sense the number of signal b-tags against the number of matrix-predicted background b-tags.

N evt	$MC_{Incl.}$	MJ
Total	1,021,924	4,249,644
Prereq	558,528	3,897,755
$N_{jet} \geq 4$	549,138	2,781,788
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	78,145	3,996
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	49,848	597
≥ 1 tag	30,410	106
in 311 pb^{-1}:	56.5	106
S/N Inclusive: $N_{mc}^{incl}/(N_{obs}^{data} - N_{mc}^{incl})$		1.14

Table 6.1: Effect of the kinematical selection $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$, $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$ and ≥ 1 SecVtx positive tag on Monte Carlo $t\bar{t}$ inclusive events as well as on multijet data sample. The row named as Prereq contains the clean up cuts: trigger (Monte Carlo events only), number of tight leptons = 0, $|Z_{jvert}| < 60\ cm$, $|Z_{jvert} - Z_{pvert}| < 5.0\ cm$ and number of good quality primary vertices ≥ 1.

Number of jets	3	4	5	6	7	8
$t\bar{t} \to ee$	0.08	0.41	0.18	0.03	0.01	–
$t\bar{t} \to e\mu$	0.06	0.29	0.11	0.04	0.01	–
$t\bar{t} \to \mu\mu$	0.01	0.05	0.01	0.01	–	–
$t\bar{t} \to e\tau$	0.11	0.93	0.38	0.12	0.03	–
$t\bar{t} \to \mu\tau$	0.05	0.29	0.15	0.05	0.01	–
$t\bar{t} \to \tau\tau$	0.06	0.58	0.26	0.04	0.01	–
$t\bar{t} \to e + jets$	0.68	6.61	8.70	3.34	0.70	0.21
$t\bar{t} \to \mu + jets$	1.07	11.92	6.56	2.03	0.38	0.06
$t\bar{t} \to \tau + jets$	1.00	10.98	11.71	4.37	0.98	0.18
$t\bar{t} \to jets$	0.01	0.09	0.14	0.15	0.04	0.03
$t\bar{t} \to X$	3.13	32.15	28.14	10.18	2.17	0.48
Background b-jets	32.68 ± 3.46	37.53 ± 4.14	21.44 ± 2.76	7.27 ± 0.91	0.75 ± 0.11	0.45 ± 0.98
Observed b-jets	31.00	53.00	55.00	15.00	2.00	2.00

Table 6.2: Number of b-tagged jets expected from $t\bar{t}$ production using $\sigma(t\bar{t}) = 6.1$ pb, predicted by the tagging rate parametrization, and observed in the selected sample, per each jet multiplicity of the events. A 10% systematic uncertainty on the background estimate is included.

in order to explain the data behavior.

Since the 597 data events selected before the tagging requirement are expected to contain a non-negligible $t\bar{t}$ component, the tagging rate parametrization procedure overestimates the background. In both Tab. 6.2 and Fig. 6.1 no attempt is made to correct for this effect. This issue will be treated in detail in Sec.7.2.1, where an iterative procedure, using the measured value of the average number of SECVTX tags per $t\bar{t}$ event, will be employed.

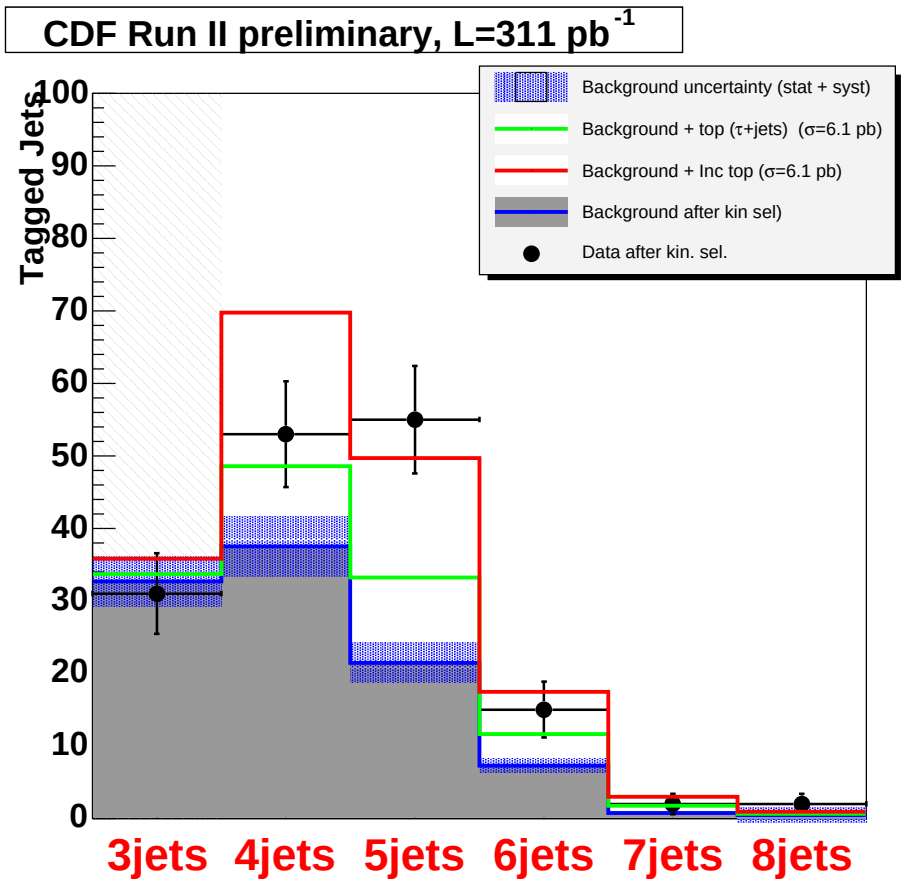


Figure 6.1: Tagging matrix background predictions after kinematical selection $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$. The predicted number of positive b-tags as a function of the jet multiplicity is shown for background, together with the expected contribution from $t\bar{t} \to \tau + jets$ and inclusive Monte Carlo events. Points refers to the observed positive SecVtx tagged jets in the data.

6.2 Two-component fits to kinematical selected and b-tagged data

The excess in terms of number of observed b-tags over the tagging matrix predictions can be ascribed to top pair production: the number of positive tagged jets expected from inclusive decaying Monte Carlo $t\bar{t}$ events surviving the kinematical selection (Tab. 6.2) is well consistent with the excess in the number of b-tagged jets observed in the data over the background prediction.

In order to further establish the $t\bar{t}$ presence in the selected data we performed binned likelihood two-component fits to some data kinematical distributions.

The templates used to fit the data are derived on one hand from the $t\bar{t}$ inclusive Monte Carlo sample, and on the other hand calculated from the positive tagging matrix application to the selected data (*i.e.* weighting each jet in the event with its positive tagging probability returned by the tagging matrix, as in Sec. 5.1.3), the latter being

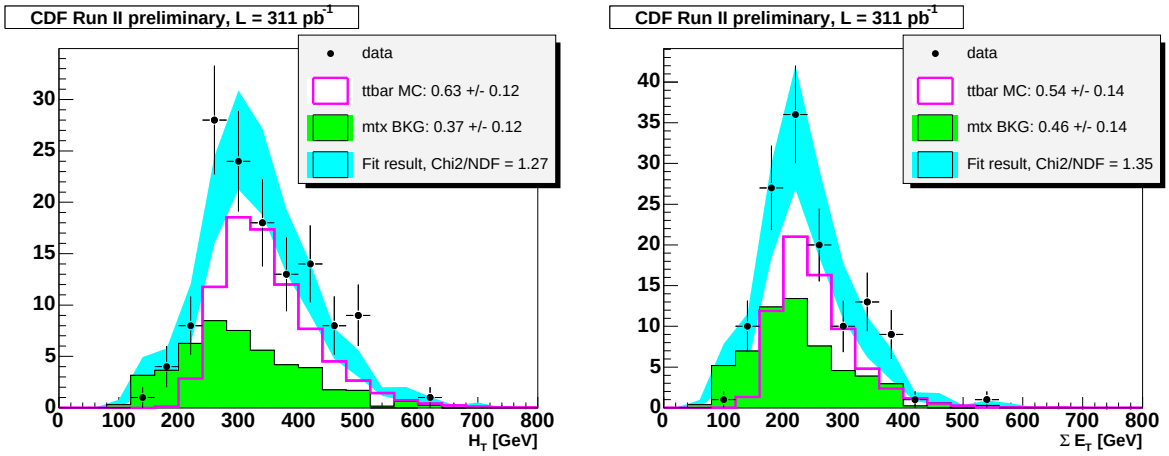


Figure 6.2: H_T and $\sum E_T$ distributions for data after kinematical selection and with at least one positive SecVtX tag. The distributions are fitted to the sum of Monte Carlo $t\bar{t}$ and background templates, the latter being derived from the tagging matrix application to data.

intended as background template. Both templates are obtained after having applied the optimized kinematical selection; finally, as far as data and Monte Carlo distributions are concerned, there is an entry per each tagged jet in the event.

Fits are performed by means of a binned likelihood technique, in which the likelihood function reads:

$$L = -2 \sum_{i=1}^{N\ bins} \left(N_i \log[f_b B_i + (1-f_b)S_i] - [f_b B_i + (1-f_b)S_i]\right), \qquad (6.1)$$

where N_i is the data content in the i-bin; f_b is the background fraction in the data and needs to be determined by the fit; finally, B_i and S_i are the bin contents of background and signal templates respectively.

Fig. 6.2 and Fig.6.3 show the result of the fits to $H_T = \sum E_T + \not\!\!E_T$, $\sum E_T$, $\sum E_T^3$, $\not\!\!E_T$, jet E_T and $\Delta\phi$ between the b-tagged jet and the $\not\!\!E_T$ direction. In the figures, data distributions are drawn with points, the fit results and the associated errors are shown by the light shaded areas. The filled histograms represent the background template; the $t\bar{t}$ contribution is represented by the open histograms and both templates are normalized to the values returned by the fits.

The $t\bar{t}$ component in the selected data, extracted by means of the fits, is in agreement with the overall prediction of the tagging matrix in the N_{jet} spectrum (Fig. 6.1 and Tab. 6.2). Fig. 6.4 shows the tagging matrix absolute prediction together with the fit results. The shaded lines represent the 10% and the 20% discrepancy regions around the tagging matrix overall prediction integrated with N_{jet}. The conclusion is that the $t\bar{t}$ fraction in our final sample is consistently found to be $\sim 50\%$. Note that many of the determinations are correlated to each other in Fig. 6.4. We will return on this issue in the following section.

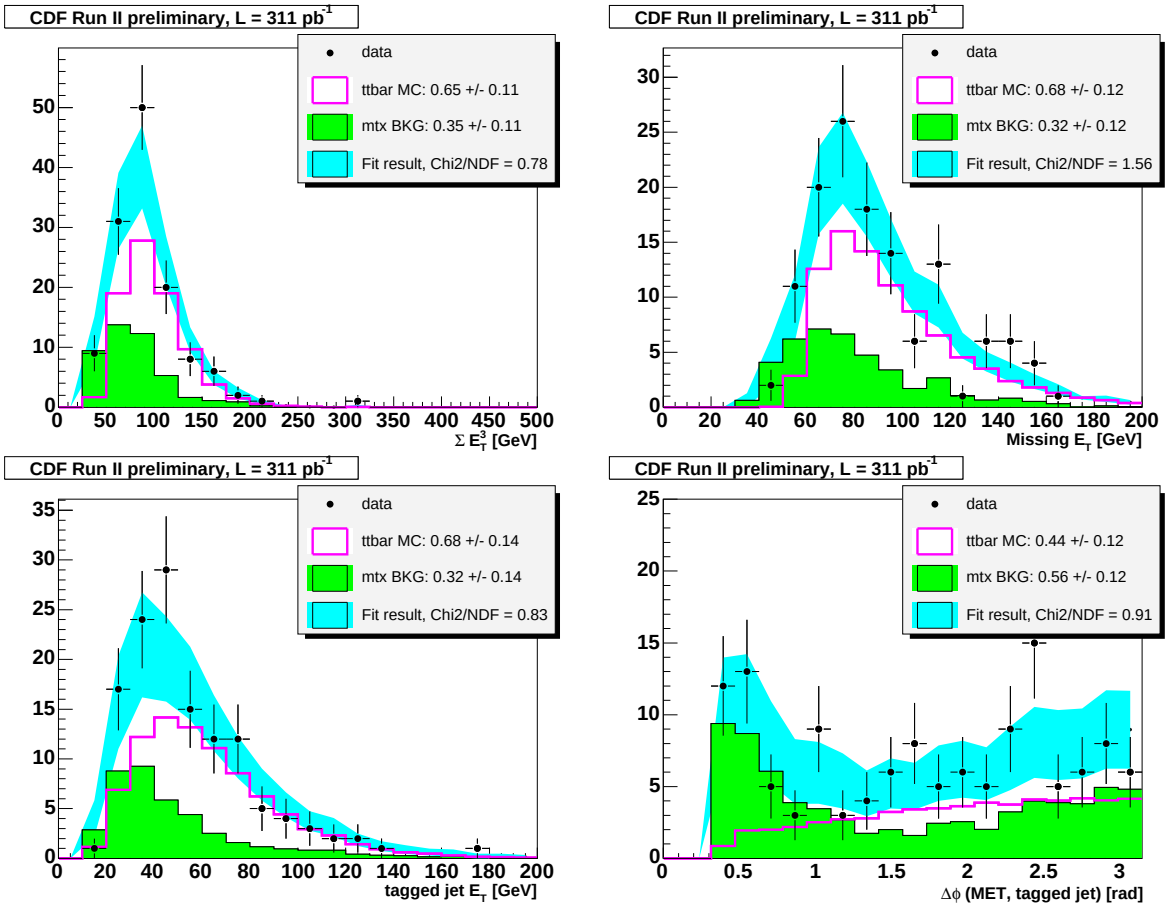


Figure 6.3: $\sum E_T^3$, $\not\!\!E_T$, jet E_T and $\Delta\phi(\not\!\!E_T,\ tagged\ jet)$ distributions for data after kinematical selection and with at least one positive SecVtX tag. The distributions are fitted to the sum of Monte Carlo $t\bar{t}$ and background templates, the latter being derived from the tagging matrix application to data.

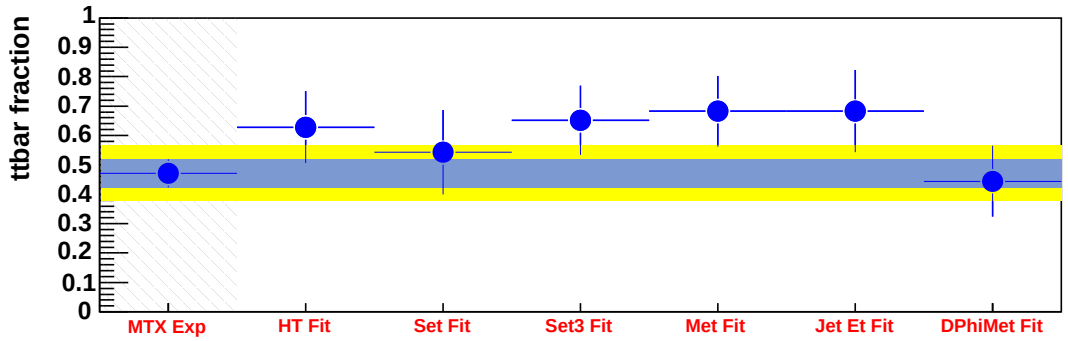


Figure 6.4: $t\bar{t}$ component in the selected data as returned by the tagging matrix (first bin on the left) approach and by the fits to kinematical variables. Note that many of the determinations are correlated to each other.

Fitted variable	F_{bkg}	$F_{t\bar{t}}$	$\sigma(F_{bkg})^{PE}$	$pull^{PE}$ mean	$pull^{PE}$ sigma
H_T	0.372 ± 0.122	0.628 ± 0.122	0.103 ± 0.012	-0.09 ± 0.03	0.93 ± 0.03
ΣE_T	0.457 ± 0.143	0.542 ± 0.143	0.119 ± 0.012	-0.03 ± 0.03	0.94 ± 0.03
ΣE_T^3	0.349 ± 0.118	0.651 ± 0.118	0.108 ± 0.011	-0.11 ± 0.03	0.97 ± 0.03
$\not\!\!E_T$	0.317 ± 0.120	0.682 ± 0.120	0.101 ± 0.012	-0.03 ± 0.03	0.95 ± 0.03
jet E_T	0.317 ± 0.140	0.682 ± 0.140	0.127 ± 0.008	0.03 ± 0.03	0.86 ± 0.03
jet min $\Delta\phi(\not\!\!E_T, jet)$	0.555 ± 0.121	0.444 ± 0.121	0.126 ± 0.008	0.05 ± 0.03	0.95 ± 0.03

Table 6.3: Fit results and the corresponding pseudoexperiment outputs. For event and jet variables.

6.2.1 Pseudo-experiments as a test of the fitting procedure

In the previous section, we performed binned likelihood fits to several kinematical distributions extracted from the final data sample, obtained by requiring the kinematical selection cuts and at least one positive tagged jet.

The fitting technique has been checked using pseudo-experiments in several ways, as it will be clear in the following.

As a first step, we generated 1,000 pseudo-experiments with the same statistics as the data, using the background fraction returned from the fits as an input to randomly construct pseudo-data distributions from the original signal and background templates weighted according to the input background fraction. The pseudo-data distributions are then fit using the same technique presented in the previous section. Fig. 6.5 shows the results of this check in the case of the $\not\!\!E_T$ fit. The top left (right) panel shows the pseudo-experiment fits output for the background fraction (and its error). The bottom left panel shows the pull distribution, where the pull is defined as the difference between fitted and input background fraction values divided by the error returned by the fit: $\frac{f_b^{fit} - f_b^{input}}{\sigma(f_b^{fit})}$. Since the pull is centered at zero with a unity σ we are confident that our fits return unbiased results and well measured errors.

Tab. 6.3 reports the fit results over the six variables considered in the previous section together with the expected errors on the background fraction, the pull means and sigmas obtained from the corresponding pseudo-experiments.

Stability of the fits with input background fraction

The second check to establish the stability of the fitting procedure is performed by generating pseudo-experiments using different input background fractions to create the pseudo-data distribution.

Fig. 6.6(a) shows the result of this check as far as the $\not\!\!E_T$ fit is concerned. The top panel shows the correlation between input and output background fraction, the former used to generate pseudo-data distributions, the latter being determined by the binned likelihood fit and averaged over 500 pseudo-experiments. The bottom panel of Fig. 6.6(a), shows the averaged pull mean and σ over the 500 pseudo-experiments performed per each input background fraction in the range $[10, 90]\%$. The pulls are centered at zero with a unity σ over the background fraction range $[20, 80]\%$, demonstrating that the fitting technique

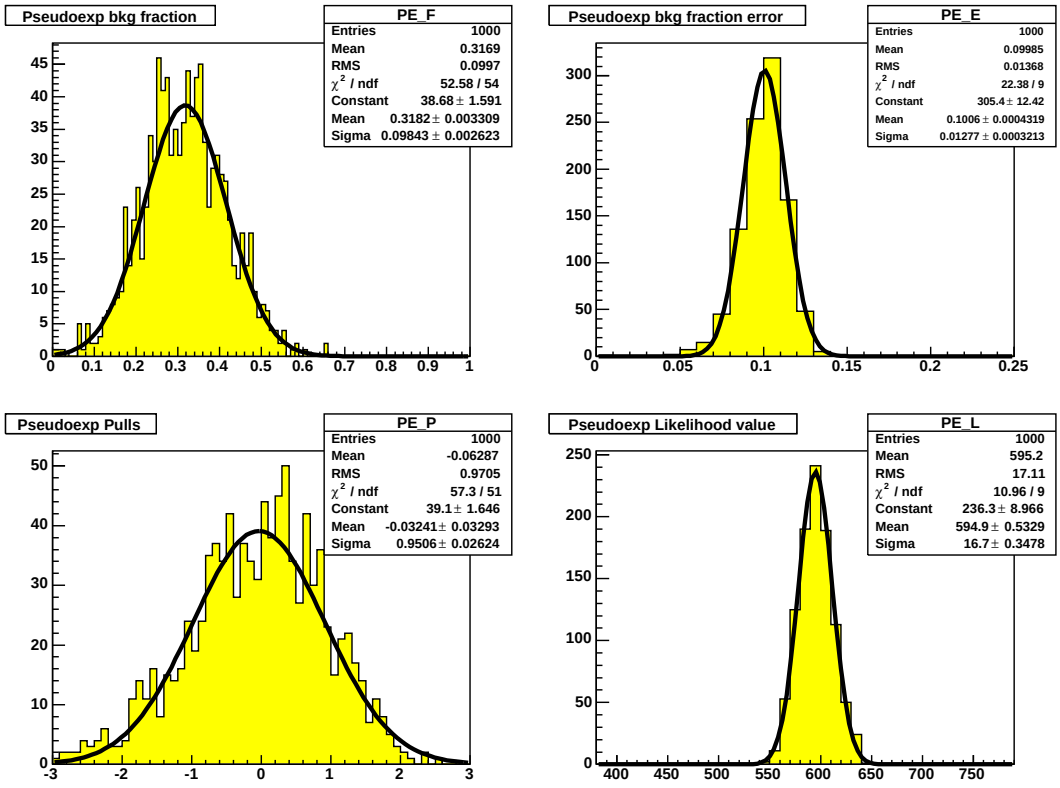
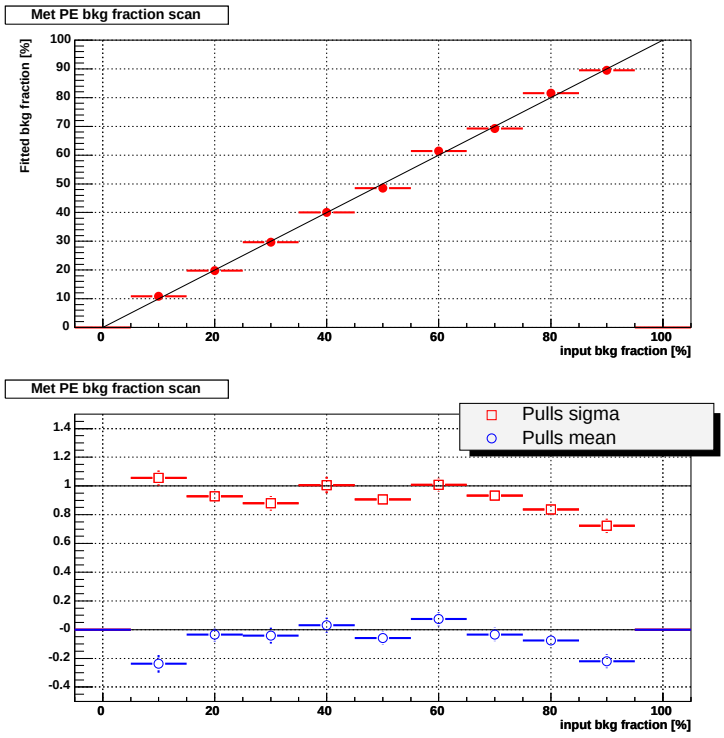
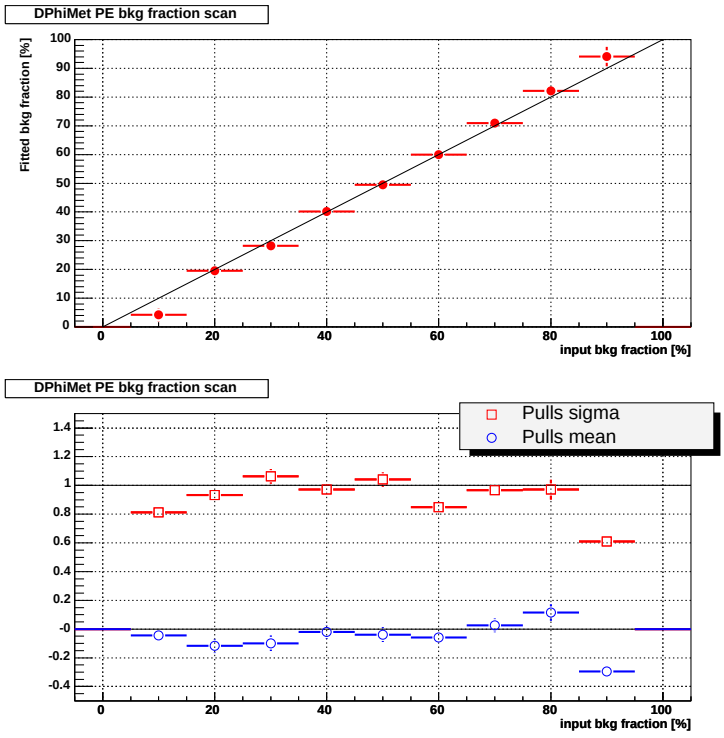


Figure 6.5: Pseudo-experiment results for the $\not\!\!E_T$ distribution.

is stable and not influenced by the actual background fraction in the data to be fit. The results of this check extended to the $\Delta\phi(\not\!\!E_T, \mathit{tagged\ jet})$ fit is also reported in Fig. 6.6(b). Also in this case the fitting technique is found to be stable for all the background fractions in the range $[20, 80]\%$.



(a)



(b)

Figure 6.6: $\not\!\!E_T$(a) and $\Delta\phi(\not\!\!E_T, tagged\ jet)$ (b) fit stability checks. Top panel: pseudo-experiment input versus fitted background fraction correlation. Bottom panel: pseudo-experiments pulls and pull-sigmas as a function of the input background fraction.

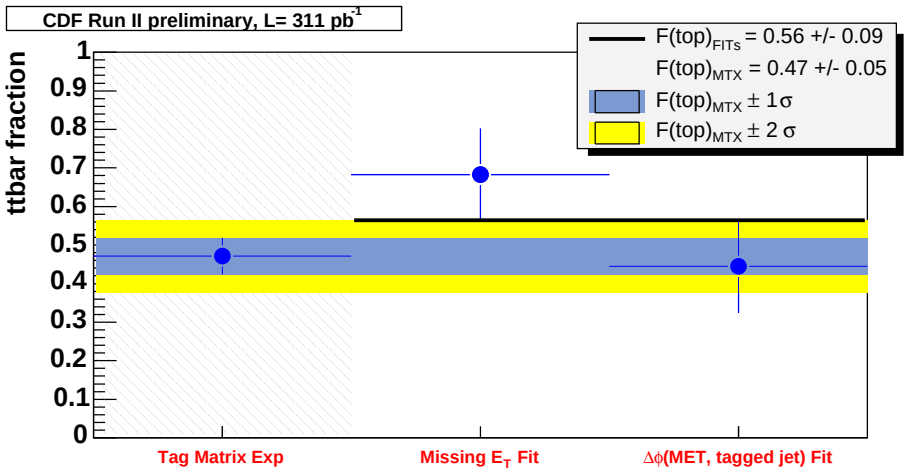


Figure 6.7: $t\bar{t}$ component in the selected data as returned by the tagging matrix approach (first bin on the left) and by the fits to $\not\!\!E_T$ and min $\Delta\phi(\not\!\!E_T, tagged\ jet)$.

6.2.2 Two-component fit results

Fits to $H_T = \sum E_T + \not\!\!E_T$, $\sum E_T$, $\sum E_T^3$, $\not\!\!E_T$, jet E_T and $\Delta\phi$ between the b-tagged jet and the $\not\!\!E_T$ direction distributions provided in Sec. 6.2 can not be directly used to derive an overall estimate of the $t\bar{t}$ fraction in the selected data sample. Indeed, as already noted, many of the variables used to perform the fits are correlated with each other.

In order to reduce the correlation between the fit results, we chose to consider one event-variable and one jet-variable to be used to quote an overall determination of background fraction, in particular to use the results of the $\not\!\!E_T$ and min $\Delta\phi(\not\!\!E_T, tagged\ jet)$ fits. This choice is justified by comparing the average error on the fitted background fraction that the pseudo-experiments return. Tab. 6.3 shows indeed that $\not\!\!E_T$ and min $\Delta\phi(\not\!\!E_T, tagged\ jet)$ fits return on average the minimum uncertainty on the fitted parameter for the event- and jet- variables respectively.

Combining together the determinations obtained by fitting $\not\!\!E_T$ and min $\Delta\phi(\not\!\!E_T, tagged\ jet)$ data distributions we obtain a fit-based determination of the $t\bar{t}$ fraction in the sample of 0.59 ± 0.09, to be compared to the result returned from the tag counting method of 0.47 ± 0.05[2] (*i.e.* comparing the amount of observed tags to the one predicted by the tagging matrix application in the sample after kinematical selection). The determinations, from tag counting and fitting, are in good agreement. The results are displayed in Fig. 6.7.

6.3 Hints on the background sample composition

In Sec. 6.2 we fitted the kinematical distributions of b-tagged data to the sum of $t\bar{t}$ and tagging matrix extracted templates. The background templates were the direct outcome of the positive tagging matrix application to data sample. It is interesting at this point

[2]Note the fit and the counting method results provided here are not corrected for the $t\bar{t}$ presence in the pre-tagging sample.

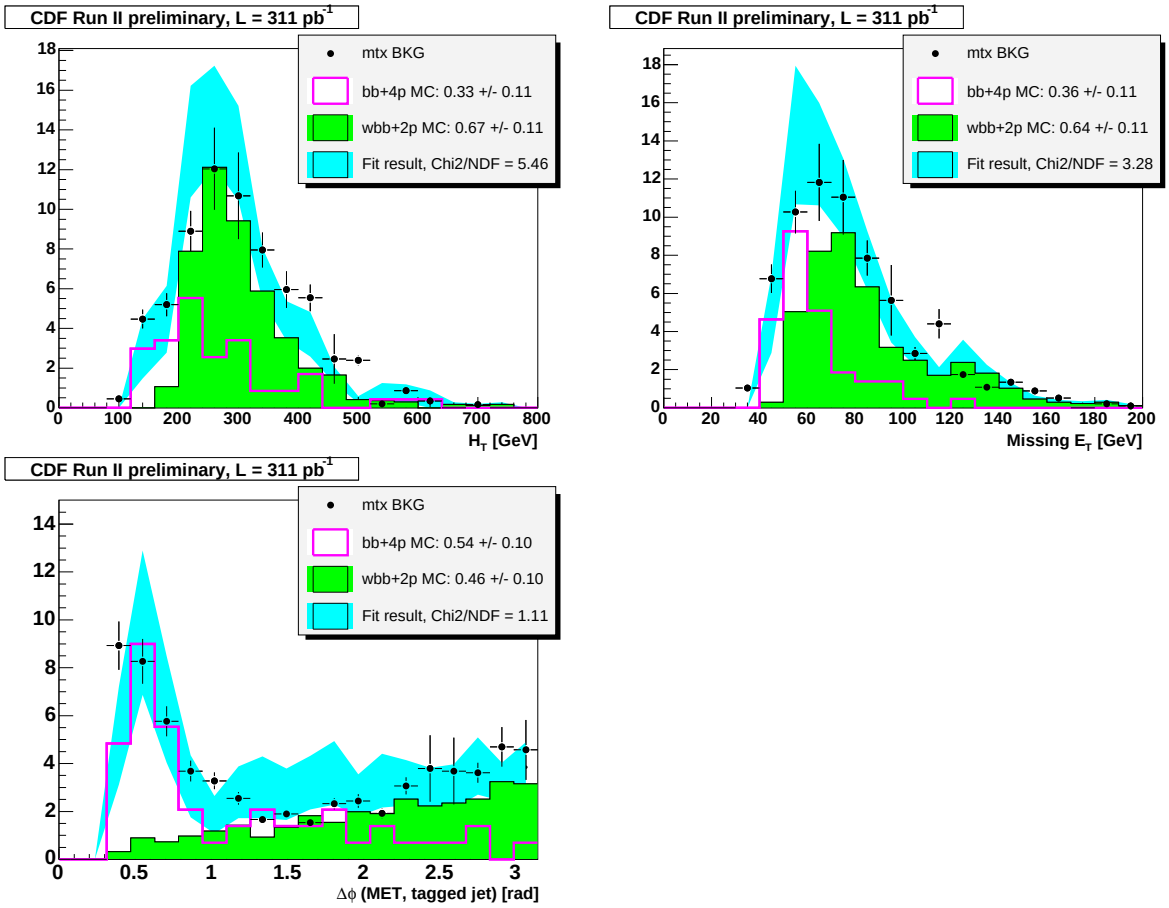


Figure 6.8: H_T, $\not\!\!E_T$ and jet $\Delta\phi(\not\!\!E_T, tagged\ jet)$ distributions for background after kinematical selection and with at least one positive SecVtX tag. The distributions are fitted to the sum of $Wb\bar{b}+2P$ and $b\bar{b}+4P$ Monte Carlo templates.

to characterize the tagging matrix distributions in terms of the background processes we expect to populate the signal region.

In particular, given the tight cut on the $\not\!\!E_T/\sqrt{\sum E_T}$ imposed by our kinematical selection and the ≥ 1 positive b-tag requirement, we expect the background to be composed mainly of $b\bar{b}+jets$ and $Wb\bar{b}+jets$ processes.

We attempted to fit the background kinematical distribution returned by the tagging matrix to the sum of $b\bar{b}$+$jets$ ($b\bar{b}$+$4P$) and $Wb\bar{b}$+$jets$ ($Wb\bar{b}$+$2P$) templates extracted from ALPGEN+HERWIG Monte Carlo events. In this case, not all the kinematical variables already shown for the data fits maintain their discrimination power. In particular the $b\bar{b}+4P$ and $Wb\bar{b}+2P$ templates for $\sum E_T$, $\sum E_T^3$ and jet E_T are very similar, preventing the fit to return significant results. Anyway, following the same approach described in the previous sections we end up with the results shown in Fig. 6.8. There H_T, $\not\!\!E_T$ and $\Delta\phi(\not\!\!E_T, tagged\ jet)$ distribution calculated with the positive tagging matrix application to the selected data are fitted to the sum of ALPGEN+HERWIG Monte Carlo templates for $b\bar{b}+4P$ and $Wb\bar{b}+2P$, respectively.

Each fit determination is consistent with the others as shown in Fig. 6.9 and Tab. 6.4.

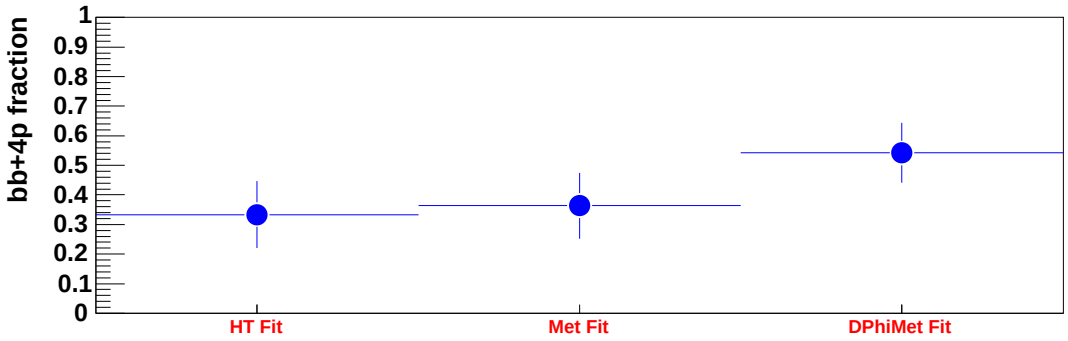


Figure 6.9: Background $b\bar{b} + 4P$ component after kinematical selection as returned fits to the background kinematical variables to the sum of $Wb\bar{b} + 2P$ and $b\bar{b} + 4P$ ALPGEN+HERWIG Monte Carlo templates.

Fitted variable	$F_{b\bar{b}}$
H_T	0.33 ± 0.11
$\not\!\!E_T$	0.36 ± 0.11
jet $\min \Delta\phi(\not\!\!E_T, jet)$	0.54 ± 0.10

Table 6.4: Background fit results. For event and jet variables.

These results prove our assumption on the background composition and return an estimate of the $b\bar{b}$ background component in the range $[33, 54]\%$. Even if these results will not be used for the $t\bar{t}$ cross section determination, they further check the ability of the b-tagging rate parametrization in describing the overall background to the $t\bar{t}$ signature we are looking at.

Chapter 7

Cross section measurement and systematic uncertainties

In the previous Chapter we estimated the $t\bar{t}$ component in the selected sample to be $\sim 50\%$ of the total, the overall background to the signal signature being estimated with the tagging rate parametrization applied to data. In this Chapter we provide the top pairs production cross section measurement in the selected data sample.

We have all the ingredients to proceed directly for a measurement, except the systematic uncertainties determination. The measurement we are going to 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:

$$\begin{aligned} \mathcal{L} \;=\; & e^{-\frac{(L-\bar{L})^2}{2\sigma_L^2}} \cdot e^{-\frac{(\epsilon_{kin}-\bar{\epsilon}_{kin})^2}{2\sigma^2_{\bar{\epsilon}_{kin}}}} \cdot e^{-\frac{(\epsilon^{ave}_{tag}-\bar{\epsilon}^{ave}_{tag})^2}{2\sigma^2_{\epsilon^{ave}_{tag}}}} \cdot e^{-\frac{(N_{exp}-\bar{N}_{exp})^2}{2\sigma^2_{N_{exp}}}} \cdot \\ & \cdot \frac{\left(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N_{exp}\right)^{N_{obs}}}{N_{obs}!} \cdot e^{-(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N_{exp})}, \end{aligned} \tag{7.1}$$

where L is the integrated luminosity of the multijet data sample, ϵ_{kin} is the combined trigger and kinematical 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; and finally, N_{obs} is the number of b-jets observed in the data. The cross section central value is given by the likelihood maximization as:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{exp}}{\epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L}. \tag{7.2}$$

In Chapter 5 we determined the overall kinematical efficiency and the average number of b-tagged jets per $t\bar{t}$ event to be $\epsilon_{kin} = 4.878 \pm 0.021\%$ and $\epsilon^{ave}_{tag} = 0.7889 \pm 0.0018$, respectively. Using the tagging rate parametrization, in Chapter 6, Tab. 6.2, the background amount in terms of b-tagged jets was calculated to be $67.4 \pm 2.7 \pm 6.7$, where

the first uncertainty is statistical only, while the latter is systematic and is calculated by comparing observed to matrix-predicted b-tags in several data control samples. On the other hand, the number of observed b-tagged jets in the data sample selected with the kinematical selection is found to be 127. Finally, the integrated luminosity of the considered data sample is $L = 311 \pm 18\ pb^{-1}$. We therefore have all the input values we need to perform the likelihood maximization.

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 of this Chapter, before quoting the final result, the sources of systematic uncertainty will be described and quantified.

7.1 Systematics

The base Monte Carlo sample adopted for this work consists of $1,021,924$ PYTHIA generated $t\bar{t}$ inclusive events. The corresponding integrated luminosity of this sample, assuming $\sigma_{t\bar{t}} = 6.1\ pb$ for $M_{top} = 178\ GeV$, is calculated to be 167.5 fb^{-1}. On the other hand, a $t\bar{t}$ inclusive decay sample generated with HERWIG is used to compute the systematic uncertainty due to generator dependence. The sample consists of $1,133,065$ events, corresponding to an integrated luminosity of 185.7 fb^{-1}.

Other inclusive $t\bar{t}$ samples are used to estimate the kinematical efficiency for various M_{top} values. These samples consist of $200,000$ events generated with HERWIG with M_{top} in the range $[130, 230]\ GeV$ in steps of $10\ GeV$. Additionally, $1M$ event $t\bar{t}$ samples with varied initial and final state radiation settings were analyzed.

All these samples are processed through the CDF detector and trigger simulation, as described in Sec. 4.1.

7.1.1 Background prediction systematic

The systematic uncertainty on the background prediction is calculated, as already explained in Sec. 5.2.2, by comparing the number of b-tags provided by applying the tagging matrix to the actual number of positive SECVTX tags in control samples depleted of signal contamination, obtained from the TOP_MULTI_JET triggered dataset. As a result of these checks a 10% systematic uncertainty to the number of background b-tags returned by the tagging matrix application to data is conservatively assigned.

7.1.2 Luminosity systematic

The integrated luminosity calculation is based on the instantaneous luminosity measurement provided by the CLC detector (Sec. 3.5.2). Two components of uncertainty play 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 4.4% and 4.0% respectively, giving a total uncertainty of 6% on the integrated luminosity calculated for any given CDF dataset[93].

7.1.3 Trigger systematic

In order to evaluate trigger systematics a sample of collider data 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;

is used. The corresponding integrated luminosity of this data sample, collected before February 2004, is evaluated to be $196 \pm 12\ pb^{-1}$. This dataset, named as "Single Tower-10", is used to extract the efficiency of the TOP_MULTI_JET trigger on a data-driven basis. In particular, we can evaluate the TOP_MULTI_JET trigger efficiency by applying its L2 requirements, directly on "Single Tower-10" triggered data.

The systematic uncertainty affecting the trigger efficiency measurement on Monte Carlo $t\bar{t}$ events, which is extracted directly by using the TRIGSIM++ output, is evaluated by comparing "Single Tower-10" data and Monte Carlo trigger turn-on curves.

Turn-on curves are derived as a function of the 4^{th} jet L5-corrected E_T in the event both for inclusive $t\bar{t}$ Monte Carlo and "Tower-10" events. A trigger efficiency systematic of $(14.8 \pm 0.2)\%$ is derived from the mis-match between Monte Carlo and data turn-on curves. Details and recent investigation on this issue are provided in Appendix C[1].

7.1.4 PYTHIA/HERWIG-related systematics

The kinematical selection optimization was derived by using 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 $1.1M$ $t\bar{t}$ events generated with HERWIG.

Tab.7.1 shows the effect of each cut of the kinematical selection for both $\tau + jets$ and 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 is taken as:

$$syst_{gen} = \frac{\Delta\epsilon}{\epsilon} = \frac{\epsilon_{HERWIG} - \epsilon_{PYTHIA}}{\epsilon_{PYTHIA}} = 8.2\%, \tag{7.3}$$

where $\epsilon_{PYTHIA} = (4.878 \pm 0.021)\%$ and $\epsilon_{HERWIG} = (5.277 \pm 0.021)\%$ are the kinematical efficiency on $t\bar{t}$ inclusive Monte Carlo events generated with PYTHIA and HERWIG, respectively.

In Tab. 7.2 the relative difference on the kinematical efficiency, $\Delta\epsilon/\epsilon$, is reported for each of the applied cuts. As far as the $\tau + jets$ exclusive signal is concerned, the main source of discrepancy between the two generators lies in the trigger simulation. An additional source of discrepancy, as far as the exclusive channel is concerned, originates from

[1]Recent studies demonstrated the possibility to quote a reduced trigger systematic at the level of few percent. The description of these recent investigations and their impact on the final measurement are provided in Appendix C.

	PYTHIA			
N evt	$MC_{(\tau+jets)}$		$MC_{Incl.}$	
Good Run	1,021,924	–	1,021,924	–
$BR(t\bar{t} \to \tau + jets)$	149,323	14.61 %	–	–
Trigger	82,200	55.05 %	647,365	63.35 %
Nlep = 0	78,084	94.99 %	583,697	90.16 %
$\lvert Z_{vert}\rvert < 60\ cm$	75,035	96.09 %	559,342	95.82 %
$\lvert Z_{pvert} - Z_{jvert}\rvert < 5\ cm$	74,912	99.83 %	558,494	99.84 %
$N_{vert}(Q = 12) \geq 1$	74,904	99.99 %	558,528	99.99 %
$N_{jet} \geq 4$	72,708	97.07 %	549,138	98.32 %
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	29,830	41.00 %	78,145	14.23 %
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	19,079	63.95 %	49,848	63.80 %
efficiency %	1.867 ± 0.014		4.878 ± 0.021	
	HERWIG			
N evt	$MC_{(\tau+jets)}$		$MC_{Incl.}$	
Good Run	1,133,065	–	1,133,065	–
$BR(t\bar{t} \to \tau + jets)$	167,374	14.77 %	–	–
Trigger	98,088	58.60 %	738,957	65.22 %
Nlep = 0	93,364	95.18 %	661,410	89.51 %
$\lvert Z_{vert}\rvert < 60\ cm$	89,857	96.24 %	633,557	95.79 %
$\lvert Z_{pvert} - Z_{jvert}\rvert < 5\ cm$	89,675	99.80 %	632,579	99.84 %
$N_{vert}(Q = 12) \geq 1$	89,653	99.97 %	632,554	99.99 %
$N_{jet} \geq 4$	87,170	97.23 %	621,670	98.28 %
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	35,716	40.97 %	93,669	15.07%
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	22,837	63.94 %	59,787	63.83%
efficiency %	2.015 ± 0.013		5.277 ± 0.021	

Table 7.1: Effect of the kinematical selection $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$ on $t\bar{t} \to \tau + jets$ and inclusive $t\bar{t}$ events generated with PYTHIA and HERWIG.

the use of different W branching ratios by the two generators: PYTHIA uses the measured $BR(W \to l\nu) = 0.108$ value, while HERWIG assumes the theoretical estimate of $BR(W \to l\nu) = 0.111$. In the case of the inclusive channel the trigger simulation related discrepancy has less impact compared to the $\Delta\epsilon$ returned for the missing transverse energy significance requirement. These effects can be understood in terms of different fragmentation models between the generator yielding different $\not\!\!E_T$, and consequently different $\sum E_T$ values.

7.1.5 PDF-related systematics

The parton distribution functions (PDFs) chosen for this analysis correspond to the CTEQ parametrization outlined in [66]. There are uncertainties associated with this parametrization, as well as other parameterizations of the PDFs that could slightly change the kine-

	PYTHIA vs HERWIG	
$\frac{\epsilon_{HERWIG}-\epsilon_{PYTHIA}}{\epsilon_{PYTHIA}}$	$MC_{(\tau+jets)}$	$MC_{Incl.}$
Good Run	–	–
$BR(t\bar{t} \to \tau + jets)$	1.1 %	–
Trigger	6.4 %	2.9 %
Nlep = 0	0.2 %	-0.7 %
$\lvert Z_{vert}\rvert < 60\ cm$	0.1 %	0.0 %
$\lvert Z_{pvert} - Z_{jvert}\rvert < 5\ cm$	0.0 %	0.0 %
$N_{vert}(Q = 12) \geq 1$	0.0 %	0.0 %
$N_{jet} \geq 4$	0.2 %	0.1 %
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	0.0 %	5.9 %
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	0.0 %	0.0 %
total	8.0 %	8.2 %

Table 7.2: $\Delta\epsilon/\epsilon$ between PYTHIA and HERWIG generator for each cut of the prerequisites and of the kinematical selection.

matics and thus the acceptance for signal events.

In order to account for these effects, we re-weighted 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[91].

As a result of this approach we set a 1.6% systematic uncertainty associated with our choice of PDFs.

7.1.6 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}, $\not\!\!E_T/\sqrt{\Sigma E_T}$ and $\min \Delta\phi(\not\!\!E_T, jet)$ requirements.

We evaluated this effect using different inclusive Monte Carlo $t\bar{t}$ samples generated with different tuning as far as initial (ISR) and final state (FSR) radiation are concerned: less/more ISR, and less/more FSR.

The impact of different ISR/FSR radiation settings on the kinematical selection is presented in Tab. 7.3. The systematic uncertainty to be related to initial state radiation effect is taken as

$$syst_{ISR} = \frac{|\epsilon_{+ISR} - \epsilon_{-ISR}|}{2\epsilon_{PYTHIA}} = 1.5\%. \tag{7.4}$$

On the other hand, final state radiation effect is computed as

N evt MC_{incl}	base	less ISR	more ISR
Total	1,021,924	827,974	826,436
Prereq	558,528	453,588	460,550
$N_{jet} \geq 4$	549,138	445,950	453,027
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	78,145	63,258	65,389
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	49,848	40,396	41,470
efficiency %	4.878 ± 0.021	4.879 ± 0.024	5.018 ± 0.024
	less FSR	more FSR-1	more FSR-2
Total	829,839	415,043	406,199
Prereq	458,737	228,406	223,483
$N_{jet} \geq 4$	450,979	224,488	219,722
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	64,836	31,993	31,242
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	41,522	20,406	19,882
efficiency %	5.004 ± 0.024	4.916 ± 0.033	4.895 ± 0.033

Table 7.3: Effect of ISR/FSR radiation variation on kinematical selection $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$.

$$syst_{ISR} = \frac{|\epsilon_{+FSR} - \epsilon_{-FSR}|}{2\epsilon_{PYTHIA}} = 1.0\%. \tag{7.5}$$

Summing in quadrature the two effects, we estimated a total systematics to be assigned to initial and final state radiation effects of $syst_{ISR/FSR} = 2\%$.

7.1.7 Systematics due to the jet energy response

All the jet quantities used for this analysis are corrected using Level 5 jet corrections. As already mentioned in Sec. 4.3.1, Level 5 jet corrections provide the best discrimination power, as far as the cut on missing transverse energy significance is concerned, compared to other levels of available corrections.

In this section we report on the systematic uncertainty related to the jet energy response. In Sec. 3.3.5 the total systematic uncertainty on the corrected jet E_T was found to vary in the range [3,8]%, 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 our cross section measurement, we varied the corrected jet energies within $\pm 1\sigma$ of their corresponding systematic uncertainty. Therefore, signal acceptance is recalculated after these variations. The results are provided in Tab.7.4. We assign a systematic uncertainty as follows:

$$syst_{jetcorr} = \frac{|\epsilon_{jetcorr,+1\sigma} - \epsilon_{jetcorr,-1\sigma}|}{2\epsilon_{kin}} = 1.5\%. \tag{7.6}$$

N evt MC_{incl}	standard jet corrs	$+1\sigma$ jet systs	-1σ jet systs
Total	1,021,924	1,021,924	1,021,924
Prereq	558,528	558,528	558,528
$N_{jet} \geq 4$	549,138	550,287	547,381
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	78,145	80,007	76,253
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	49,848	50,497	49,022
efficiency %	4.878 ± 0.021	4.941 ± 0.021	4.797 ± 0.021

Table 7.4: Effect of the jet energy correction within their uncertainty on the kinematical selection $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$ on $t\bar{t}$ inclusive events.

The effect of $\pm 1\sigma$ variations in the jet energy response is small: we ask the events to have at least $4(E_T^{L5} \geq 15\ GeV, |\eta| < 2.0)$ jets while on the other hand the trigger already requires for $4(E_T^{raw} \geq 10\ GeV)$ jets. It is therefore unlikely for an event firing the trigger to have less than 4 offline jets, even varying the jet energy response.

We can check this explanation by computing the jet energy response systematics effect on the category of events with at least 5 $(E_T^{L5} \geq 15\ GeV, |\eta| < 2.0)$ jets. In this case we are going to reduce the trigger bias as far as jet counting is concerned. The result of this check is indeed a larger systematics effect: the number of inclusive $t\bar{t}$ events surviving the kinematical selection, with $N_{jet} \geq 5$ is 27,471, 29,548, and 25,331 out of 1,021,924 for the standard, and $\pm 1\sigma$ jet correction respectively. In this case $syst_{jetcorr} = \frac{|\epsilon_{jetcorr,+1\sigma} - \epsilon_{jetcorr,-1\sigma}|}{2\epsilon_{kin}}$ is found to be 7.6%.

7.1.8 *b*-tagging scale factor systematics

The SECVTX efficiency scale factor used in this analysis, as explained in Sec. 4.3.3, to count the number of *b*-tags on Monte Carlo events is $SF = 0.909 \pm 0.060$. The average number of *b*-tags per $t\bar{t}$ event, ϵ_{tag}^{ave}, enters directly in the cross section measurement, and for this reason we computed the systematics effect related to its knowledge.

In particular, to account for the scale factor uncertainty we varied it within $\pm 1\sigma$ from its central value of 0.909 and we determined the difference in terms of average number of *b*-tags per event we get on Monte Carlo events with respect to the standard value[2].

The results are summarized in Tab. 7.5. As a results we assign:

$$syst_{\epsilon_{tag}} = \frac{|\epsilon_{tag,+1\sigma} - \epsilon_{tag,-1\sigma}|}{2\epsilon_{tag}^{ave}} = 5.8\%. \tag{7.7}$$

7.1.9 Summary of the systematics

The summary of all the sources of systematic uncertainty to the cross section evaluation is listed in Tab. 7.6.

[2]The SECVTX scale factor has the same central value for both *b*- and *c*-quarks, but for the latter the scale factor uncertainty is doubled: $SF_b = 0.909 \pm 0.060$, $SF_c = 0.909 \pm 0.120$.

N evt MC_{incl}	standard SF	$+1\sigma$ SecVtx SF	-1σ SecVtX SF
Total	1,021,924	1,021,924	1,021,924
After Kin Sel	49,848	49,848	49,848
≥ 1 tag	30,410	31,591	29,021
Tot tags	39,326	41,435	36,876
ϵ_{tag}^{ave}	0.7889 ± 0.0018	0.8312 ± 0.0017	0.7398 ± 0.0020

Table 7.5: Effect of the SECVTX scale factor variation within its uncertainty. The last row reports the average number of b-tags per $t\bar{t}$ event in the case of standard, and $\pm 1\sigma$ scale factor settings.

Systematics uncertainty sources		
Source	Method	Uncertainty
ϵ_{kin} systematics		
Trigger simulation	turn-on curves	14.8 %
Generator dependence	$\frac{\lvert\epsilon_{PYTHIA}-\epsilon_{HERWIG}\rvert}{\epsilon_{PYTHIA}}$	8.2 %
PDFs	MC reweighting	1.6 %
ISR/FSR	samples comparison	2.0 %
Jet Energy Scale	$\frac{\lvert\epsilon_{jetcorr,+1\sigma}-\epsilon_{jetcorr,-1\sigma}\rvert}{2\epsilon_{kin}}$	1.5 %
ϵ_{tag} systematics		
SecVtX scale factor	$\frac{\lvert\epsilon_{tag,+1\sigma}-\epsilon_{tag,-1\sigma}\rvert}{2\epsilon_{tag}}$	5.8 %
Tagging matrix systematics		
Data control samples	N_{obs}/N_{exp}	10.0 %
Luminosity systematics		
Luminosity measurement	–	6.0 %

Table 7.6: Summary of the systematic uncertainty sources.

7.2 Towards the cross section measurement

We have studied the background contribution as well as the top component in the sample selected by our kinematical selection. On the other hand we 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 now ready to proceed describing the cross section measurement.

7.2.1 Pre-tagging $t\bar{t}$ subtraction

After the application of the kinematical selection we are left with 597 data events for $4 \leq N_{jet} \leq 8$, out of which 106 have at least one positive SECVTX tag (see Tab. 7.7), for a total of 127 b-tagged jets.

As already mentioned in Chapter 6, since the 597 data events selected before the tagging requirement are expected to contain a non-negligible $t\bar{t}$ component, the tagging rate parametrization procedure overestimates the background. The expected number of b-tags provided by the positive tagging matrix parametrization is $N_{exp} = 67.4 \pm 2.7(stat) \pm 6.7(syst) = 67.4 \pm 7.2$. This expectation does not refer to background events only, since it receives a contribution from $t\bar{t}$ events in the pre-tagging sample. The consequence of this is that we need to remove the $t\bar{t}$ contribution 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 as follows[92]:

$$N'_{exp} = N^{fix}_{exp}\frac{N_{evt} - N^{t\bar{t}}_{evt}}{N_{evt}} = N^{fix}_{exp}\frac{N_{evt} - \frac{N_{obs}-N_{exp}}{\epsilon^{ave}_{tag}}}{N_{evt}}, \tag{7.8}$$

where N^{fix}_{exp} is equal to 67.4 and it is fixed, $N_{evt} = 597$ is the number of events in the pre-tagging kinematically selected data sample, and $N^{t\bar{t}}_{evt}$ is the $t\bar{t}$ contamination in the pre-tagging sample estimated as $\frac{N_{obs}-N_{exp}}{\epsilon^{ave}_{tag}}$.

The procedure stops when the difference $|N'_{exp} - N_{exp}| \leq 1\%$. The results of this approach is that 10.0 out of 67.4 tags are attributed to the $t\bar{t}$ presence in the pre-tagging sample yielding a top-free background determination of $N'_{exp} = 57.4 \pm 8.1$, where the uncertainty is evaluated propagating the total N_{exp} and ϵ^{ave}_{tag} statistical uncertainties.

7.2.2 Cross Section measurement

The excess in the number of tags defined as $N_{obs} - N'_{exp}$ is associated to the $t\bar{t}$ production and used for the cross section measurement.

In Fig. 7.1 the number of expected and observed b-tags are shown as a function of the jet multiplicity together with the expectation from inclusive Monte Carlo $t\bar{t}$ events, assuming a $t\bar{t}$ production cross section of 6.1 pb. The background expectations are corrected iteratively for the $t\bar{t}$ presence in the pre-tagging sample.

As already mentioned in the introduction of this Chapter, the cross section is measured

N evt	$MC_{(\tau+jets)}$	$MC_{Incl.}$	MJ
Total	149,323	1,021,924	4,249,644
Prereq	74,904	558,528	3,897,755
$N_{jet} \geq 4$	72,708	549,138	2,781,788
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	29,830	78,145	3,996
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	19,079	49,848	597
≥ 1 tag	11,666	30,410	106
in 311 pb^{-1}:	21.7	56.5	106
S/N $\tau + jets$: $N^{\tau}_{mc}/(N^{data}_{obs} - N^{incl}_{mc})$			0.44
S/N Inclusive: $N^{incl}_{mc}/(N^{data}_{obs} - N^{incl}_{mc})$			1.14

Table 7.7: Effect of the kinematical selection $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$, $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$ and ≥ 1 SECVTX positive tag on $t\bar{t} \to \tau + jets$ and $t\bar{t}$ inclusive events as well as on multijet data sample.

Variable	Symbol	Input Value	Output Value
Integrated Luminosity (pb^{-1})	$\mathcal{L}$	311 ± 18	311 ± 18
Observed Tags	N_{obs}	127	–
Expected Tags	N'_{exp}	57.4 ± 8.1	57.0 ± 8.1
Kin. efficiency (%)	ϵ_{kin}	4.878 ± 0.838	4.813 ± 0.854
Ave. b-tags per $t\bar{t}$ event	ϵ^{ave}_{tag}	0.7889 ± 0.0458	0.7874 ± 0.0458

Table 7.8: Input and output values of the likelihood maximization.

by maximizing $\log \mathcal{L}$, where the likelihood function is defined as follows:

$$\mathcal{L} = e^{-\frac{(L-\bar{L})^2}{2\sigma_L^2}} \cdot e^{-\frac{(\epsilon_{kin}-\bar{\epsilon}_{kin})^2}{2\sigma^2_{\epsilon_{kin}}}} \cdot e^{-\frac{(\epsilon^{ave}_{tag}-\bar{\epsilon}^{ave}_{tag})^2}{2\sigma^2_{\epsilon^{ave}_{tag}}}} \cdot e^{-\frac{(N'_{exp}-\bar{N}'_{exp})^2}{2\sigma^2_{N'_{exp}}}} \cdot \cdot \frac{(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N'_{exp})^{N_{obs}}}{N_{obs}!} \cdot e^{-(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L + N'_{exp})}. \quad (7.9)$$

The central value is given by the likelihood maximization, that is:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N'_{exp}}{\epsilon_{kin} \cdot \epsilon^{ave}_{tag} \cdot L}. \quad (7.10)$$

The input and output parameters of the likelihood maximization are quoted in Tab.7.8. The measured cross section value is found to be:

$$\sigma_{t\bar{t}} = 5.9 \pm 1.2(stat)\ ^{+1.4}_{-1.0}(syst)\ pb \quad (7.11)$$

$$= 5.9\ ^{+1.8}_{-1.6}\ pb. \quad (7.12)$$

Separating the contribution to the uncertainty from the luminosity measurement we can re-write the result as follows:

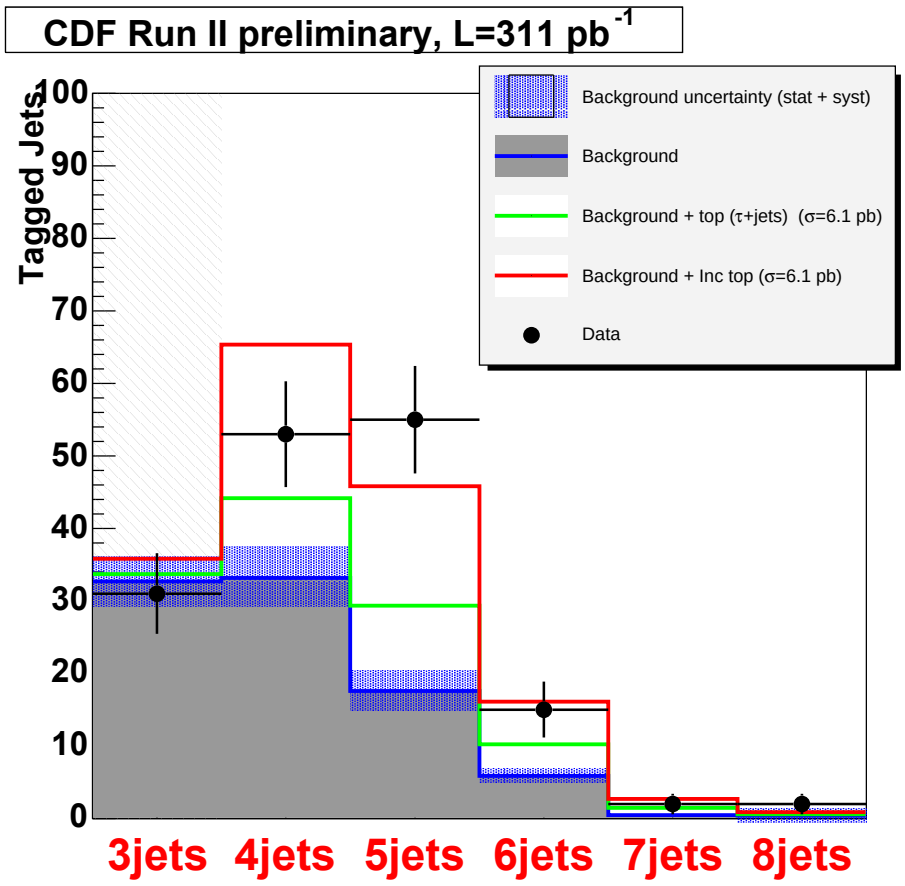


Figure 7.1: Number of tagged jet versus jet multiplicity. Data (points), iteratively corrected background (shaded histogram) and $t\bar{t}$ expectation (lines) ($\sigma_{t\bar{t}} = 6.1\ pb$) are shown after kinematical selection.

$$\sigma_{t\bar{t}} = 5.9 \pm 1.2(stat)\ ^{+1.3}_{-0.9}(syst)\ ^{+0.4}_{-0.3}(lum)\ \ pb. \tag{7.13}$$

Fig. 7.2 depicts the number of expected and observed b-tags as a function of the jet multiplicity together with the expectation from inclusive and $\tau + jets$ Monte Carlo $t\bar{t}$ events normalized to the measured cross section value of 5.9 pb.

7.2.3 Top mass cross section dependen e

The value of M_{top} used to generate $t\bar{t}$ events is of course fundamental to determine the efficiency of the kinematical selection, being directly related to the energy of the $t\bar{t}$ pair decay products.

For this reason, we measured the kinematical efficiency for inclusive HERWIG $t\bar{t}$ events generated with M_{top} in the range of $[130, 230]\ GeV/c^2$ with a step size of 10 GeV/c^2.

Tab. 7.9 shows the effect of the kinematical selection over 12 different HERWIG samples. Fig. 7.3 puts in a graphical way the results of the table (there the kinematical

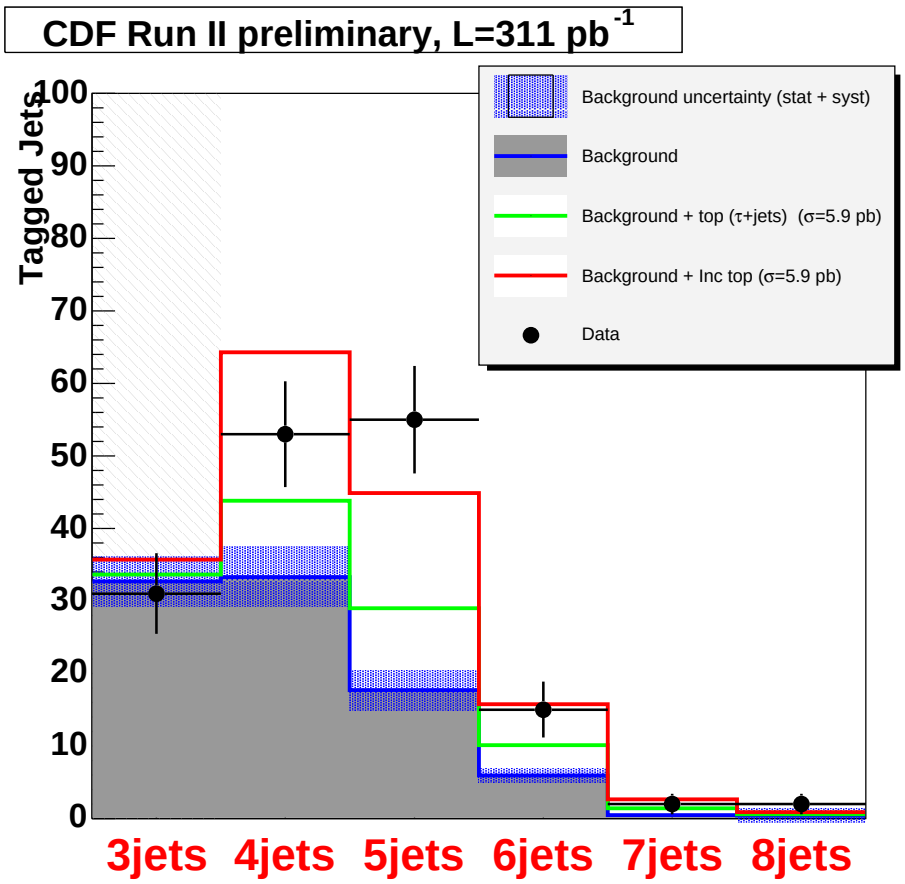


Figure 7.2: Number of tagged jet versus jet multiplicity. Data (points), iteratively corrected background (shaded histogram) and $t\bar{t}$ expectation (lines) ($\sigma_{t\bar{t}} = 5.9\ pb$) are shown after kinematical selection.

efficiency determination using PYTHIA $M_{top} = 178\ GeV/c^2$ is also shown for comparison).

From this information we can evaluate the impact on the cross section measurement of different top quark masses. In particular, assuming a flat average number of tags per $t\bar{t}$ event, ϵ_{tag}^{ave}, over different masses we can use the kinematical efficiencies listed in Tab. 7.9 to compute the corresponding cross section values[3]. The results are shown in Fig. 7.4.

The cross section changes by ± 0.05 pb for each ∓ 1 GeV/c^2 change in the top mass from the initial value of 178 GeV/c^2. For instance, we measure $\sigma_{t\bar{t}} = 6.1 \pm 1.2(stat.)^{+1.4}_{-1.0}(syst.)$ pb for $M_{top} = 175$ GeV/c^2. The change is due to the varying signal selection efficiency with top quark mass.

[3]The cross section determinations using HERWIG kinematical efficiencies are then rescaled by a factor 1.082 in order to quote the PYTHIA-based values.

N evt MC_{incl}	$M_{top} = 178\ GeV$	$M_{top} = 130\ GeV$	$M_{top} = 140\ GeV$	$M_{top} = 150\ GeV$
Total	1,133,065	205,958	196,419	201,984
Prereq	632,554	82,744	88,972	99,515
$N_{jet} \geq 4$	621,670	80,995	87,299	97,690
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	93,669	7,919	9,530	11,841
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	59,787	5,211	6,337	7,792
kin efficiency %	5.277 ± 0.021	2.530 ± 0.034	3.226 ± 0.038	3.857 ± 0.043
N evt MC_{incl}	$M_{top} = 160\ GeV$	$M_{top} = 170\ GeV$	$M_{top} = 180\ GeV$	$M_{top} = 190\ GeV$
Total	201,570	202,792	199,546	199,107
Prereq	105,417	110,704	111,902	113,979
$N_{jet} \geq 4$	103,488	108,755	109,990	112,121
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	13,708	15,667	16,676	18,004
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	9,051	10,119	10,557	11,066
efficiency %	4.490 ± 0.045	4.989 ± 0.048	5.290 ± 0.050	5.558 ± 0.051
N evt MC_{incl}	$M_{top} = 200\ GeV$	$M_{top} = 210\ GeV$	$M_{top} = 220\ GeV$	$M_{top} = 230\ GeV$
Total	197,980	193,844	196,370	192,931
Prereq	115,151	114,182	116,631	115,358
$N_{jet} \geq 4$	113,317	112,435	114,929	113,662
$\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4$	19,079	19,629	21,035	21,590
$\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$	11,613	11,746	12,380	12,233
efficiency %	5.866 ± 0.052	6.059 ± 0.054	6.304 ± 0.054	6.341 ± 0.055

Table 7.9: Effect of the kinematical selection $N_{jet}(E_T^{L5} \geq 15, |\eta| \leq 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$ and $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4$ on different HERWIG $t\bar{t}$ inclusive samples generated with M_{top} in the range $[130,\ 230]\ GeV/c^2$.

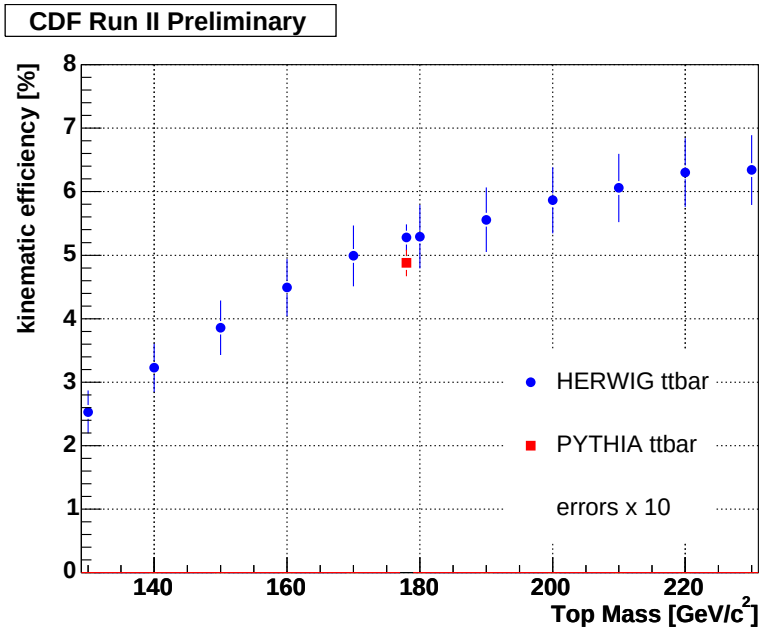


Figure 7.3: Kinematic efficiency as a function of M_{top}.

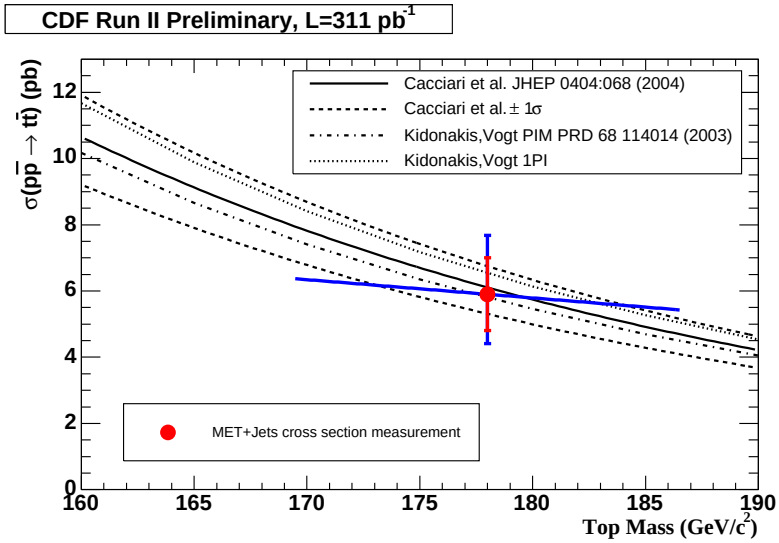


Figure 7.4: $t\bar{t}$ cross section dependence on M_{top}. The point with error bar corresponds to the measured cross section $\sigma_{t\bar{t}} = 5.9\ ^{+1.8}_{-1.6}\ pb$. The measurement is compared to the theory predictions.

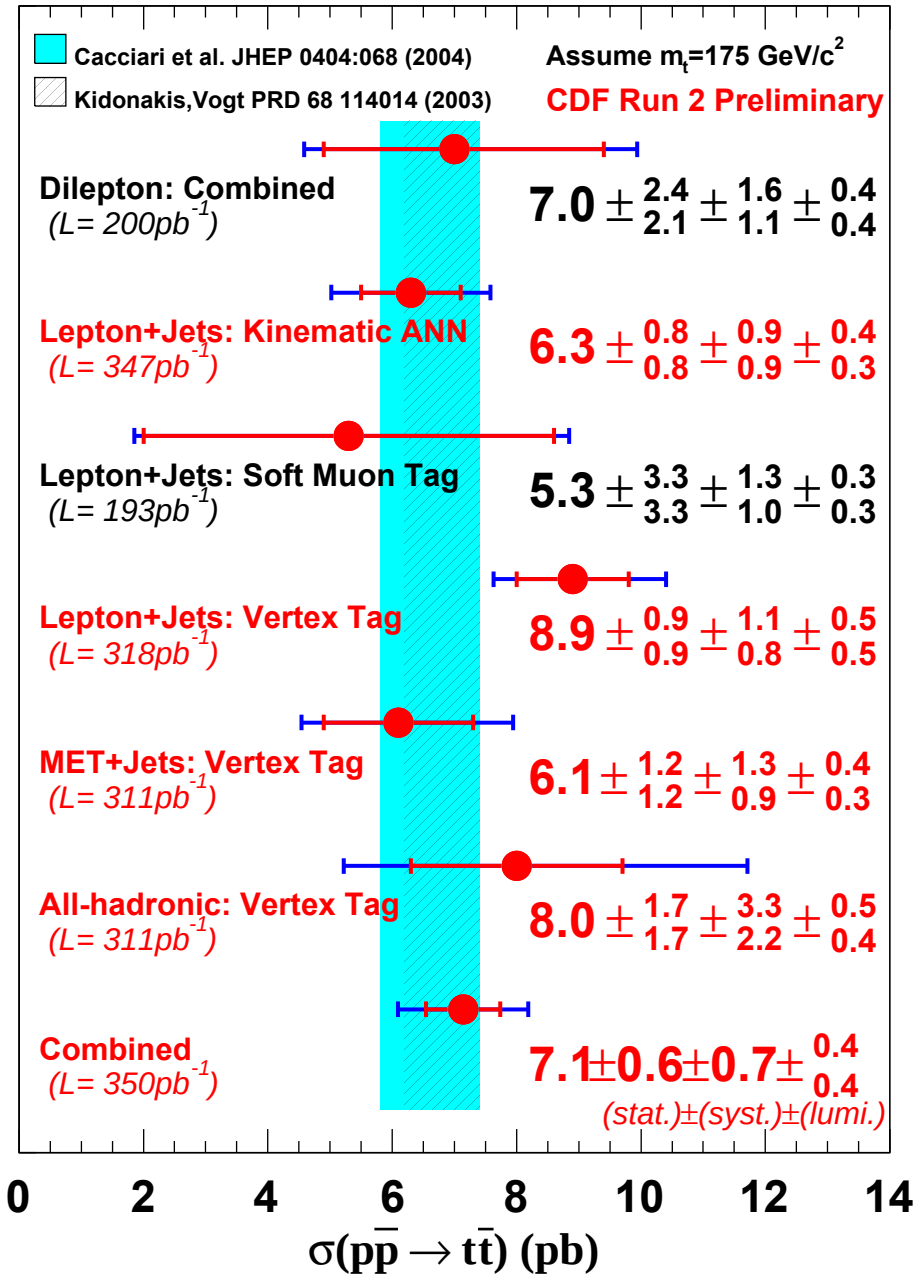


Figure 7.5: Latest CDF cross section results.

7.3 Cross section measurements combination

The latest CDF $t\bar{t}$ production cross section measurements are reported on Fig. 7.5 for a top quark mass of 175 GeV/c^2, together with the theoretical expectation. We used the kinematical efficiency dependence on M_{top}, described in Sec. 7.2.3, to rescale the measured cross section value to a $M_{top} = 175\ GeV/c^2$, obtaining:

$$\sigma_{t\bar{t}}(M_{top} = 175\ GeV/c^2) = 6.1 \pm 1.2(stat)\ ^{+1.3}_{-0.9}(syst)^{+0.4}_{-0.3}(lum)\ pb.$$

Our determination together with the CDF latest results depicted in in Fig. 7.5 are combined together to obtain[94]:

$$\sigma_{t\bar{t}}^{combined} = 7.1 \pm 0.6(stat) \pm 0.7(syst.) \pm 0.4(lum)\ pb,$$

which is the most precise measurement to date.

In the following details on the combination of the latest cross section determinations will be given.

Combination of the latest CDF results

CDF has produced several measurements of the top pair production cross section in the di-lepton, lepton+jets and all-hadronic channels. These results are summarized in Fig. 7.5. As 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 following measurements are combined:

- A combination of two di-lepton channel measurements obtained using the di-lepton and lepton+track methods, respectively. The di-lepton method requires two high-P_T identified leptons, large $\not\!\!E_T$, and at least two high-E_T jets. The lepton+track method increases the acceptance for signal events relaxing the identification requirements for the second lepton to an isolated track. For both measurements the dominant systematics originate from the estimate of background from Z/γ production with associated jets. The diboson background is estimated from Monte Carlo simulation with NNLO theoretical prediction. The background from false lepton identification is estimated using inclusive jet data samples.

- The lepton+jet channel measurement using Artificial Neural Networks (ANN). This measurement exploits the different kinematics and topology of $t\bar{t}$ and background from W+jets and QCD multi-jet processes. The number of $t\bar{t}$ events is extracted from a fit to the ANN output distribution in the $W+\geq 3$ jets sample. The dominant systematic uncertainty originates from the jet energy scale systematics and from the dependence of the leading-order Monte Carlo model for W+jets background.

- The traditional results in the lepton+jets channel with secondary vertex tagging (SECVTX tags). The reconstruction of secondary displaced vertices is used to identify b-jets and to suppress the dominant W+jets background. The requirement of at least one b-tag in $t\bar{t}$ events has an efficiency of about 55% for events passing the $W+\geq 3$ jets event selection requirements. The overall background originates from several processes: W+jets production in association with heavy flavors, W+jets production where a light flavor jet is mis-identified by the SECVTX tagging algorithm; QCD multijet, di-boson, and single-top production. The background contribution to the selected sample is estimated using data and Monte Carlo. The number of W+heavy flavor before b-tagging is estimated using the fraction of W+heavy flavor to W+jets estimated from ALPGEN Monte Carlo, scaled by the number of observed W+jets events in the pre-tagging data sample. The number of W+heavy

flavor in the tagged sample is then given by applying the b-tag efficiency estimated in W+heavy flavor Monte Carlo events. The mis-tag rate is parameterized from independent data samples. The dominant systematic uncertainty is from the b-tagging efficiency, closely followed by the uncertainty on the W+heavy flavor, mis-tag and QCD multijet backgrounds.

- The lepton+jet measurement using soft lepton tagging technique (SLT) for b-jet identification. This analysis uses the b-jet identification provided by the reconstruction of low-P_T leptons close to the jet axis. The requirement of at least one b-tag, using the soft muon method, has an efficiency of about 15% for events passing the $W+\geq 3$ jets event selection requirements. The dominant background arises from mis-tags of the b-jet identification algorithm, rather than W+heavy flavor processes. The fake tag rate is parameterized from independent data samples. The dominant systematic uncertainty here is from the fake SLT tag prediction and the SLT tag rate efficiency.

- The all-hadronic measurement. This analysis uses the same multijet trigger employed from the $\not\!\!E_T + jet$ analysis described in this document, and requires very energetic signatures for the event selection together with the requirement of at least one SECVTX b-tag. Due to the low signal over background ratio even after b-tagging, the dominant systematics arise from background estimate, which is performed using b-tag rates parameterized from independent data sample.

- Finally, the $\not\!\!E_T + jets$ analysis described in this Thesis.

The combination technique uses the BLUE algorithm[95], 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. Without going into detail, which can be found in [94], the input to the combination are listed in Tab. 7.10.

The total (statistical+systematic) correlation matrix for the 6 analyses is shown in Tab. 7.11. The combined value for the top pair production cross section, assuming a top quark mass of 175 GeV/c^2, is:

$$\sigma_{t\bar{t}}^{combined} = 7.14 \pm 1.05 \quad pb.$$

where the total error of 1.05 pb contains the following components:

- a statistical error of 0.60 pb;

- a systematic error contribution of:

  - acceptance: 0.55 pb;
  - luminosity: 0.41 pb;
  - SVX b-tag efficiency: 0.21 pb;
  - SLT b-tag efficiency: 0.02 pb;
  - $\not\!\!E_T + jets$ trigger efficiency: 0.18 pb;

Quantity		SVX	ANN	$\not\!\!E_T$+jets	All-had	SLT	DIL
Result	(pb)	8.9	6.3	6.1	8.0	5.3	7.0
Statistical uncertainty	(pb)	0.9	0.8	1.2	1.7	3.3	2.2
Acceptance	(%)	**7.7**	**6.4**	**8.7**	**22.4**	**7.8**	**15.0**
Luminosity	(%)	**5.8**	**5.8**	**5.8**	**5.8**	**5.8**	**5.8**
Trigger Efficiency	(%)	–	–	14.8	–	–	–
SVX b-tag	(%)	**6.0**	–	**5.8**	**7.8**	–	–
SLT b-tag	(%)	–	–	–	–	9.5	–
Bkg counting	(%)	6.9	–	10.0	42.0	13.0	11.0
Jet Energy Scale	(%)	–	7.7	–	–	–	–
Kinematic shapes	(%)	–	5.0	–	–	–	–
QCD background	(%)	–	1.3	–	–	–	–
Other	(%)	–	10.2	–	–	–	–
Total	(pb)	1.49	1.30	1.78	4.24	3.44	2.63

Table 7.10: Results, statistical and systematic uncertainties for the analyses used for the cross section combination. 100% correlation systematics are indicated in **bold font**. The uncertainties are quoted in % on the cross section. Numbers are from [94].

Correlation	SVX	ANN	$\not\!\!E_T$+jets	All-had	SLT	DIL
SVX	1.00	0.40(0.43)	0.27	0.23	0.16	0.21
ANN		1.00	0.18	0.16	0.14	0.19
$\not\!\!E_T$+jets			1.00	0.18	0.08	0.16
All-had				1.00	0.07	0.17
SLT					1.00	0.08
DIL						1.00

Table 7.11: Statistical+systematic correlation between the analyses used for the cross section combination. Statistical only correlation is shown in brackets. Diagonal symmetry is implicit. Numbers are from [94].

- background estimates: 0.24 *pb*;
- jet energy scale on kinematic shapes: 0.20 *pb*;
- other $t\bar{t}$ modeling of kinematic shapes: 0.13 *pb*;
- QCD multijet background: 0.03 *pb*;
- other systematics: 0.26 *pb*.

The SVX analysis carries a relative weight of 34%, the ANN of 40%, and the $\not\!\!E_T$+jets of 17%. Finally, the All-had, SLT and DIL analyses carry -0.5%, 4%, 6%, respectively. The combination is summarized in Tab. 7.12 where the pull of each measurement with respect the BLUE is also shown.

Analysis	Weight	Pull	Result (pb)	Total Uncertainty (pb)
SVX	0.34	+1.30	8.9	1.49 (1.36)
ANN	0.40	−0.63	6.3	1.30 (1.33)
$\not\!\!E_T$+jets	0.17	−0.54	6.1	1.78 (1.91)
All-had	−0.005	+0.21	8.0	4.24 (4.15)
SLT	0.04	−0.52	5.3	3.44 (3.51)
DIL	0.06	+0.05	7.0	2.63 (2.64)
Combined	1.00	–	7.14	1.05

Table 7.12: Weight of analysis results used for the cross section combination. The number in brackets is the total uncertainty when the acceptance and luminosity uncertainties are evaluated with respect to the combined value. Numbers are from [94].

Conclusions

The research presented in this Thesis is aimed at the isolation of the $t\bar{t} \to \not\!\!E_T + jets$ signal from a multijet triggered dataset. For the first time this decay channel has been extracted using high-P_T neutrino signatures and vetoing explicitly well identified high-P_T electrons or muons from W boson decay. Heavy flavor jets from top quark decay are identified with a secondary vertex b-tagging algorithm.

The background prediction is performed by parameterizing the b-tagging rate as a function of the jet E_T, N_{trk} and $\not\!\!E_T{}^{prj} = \not\!\!E_T \cos\Delta\phi(\not\!\!E_T, jet)$ in a data sample depleted of signal contamination. The performance of the tagging parametrization and in particular its capability to correctly predict the amount of background b-tags are checked in several control samples. The discrepancies between observed and predicted b-tags are found to be well below 10% level.

Using the b-tagging parametrization background predictions, we optimized a kinematical selection aimed at minimizing the relative statistical error on a cross section measurement. The best choice of cuts is found to be: $N_{jet}(E_T^{L5} \geq 15\ GeV, |\eta| < 2.0) \geq 4$, $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4.0\ GeV^{1/2}$ and $\min\Delta\phi(\not\!\!E_T, jet) \geq 0.4$. With these requirements we ended up in the pre-tagging sample with 93 $t\bar{t}$ events expected in 597 observed events in the multijet data sample, providing a signal to background ratio $S/N \sim 1/5$, before b-jet identification requirements.

The final sample, used for the cross section measurement, is finally defined by means of the additional requirement of at least one jet identified as originating from a b-quark; in this sample the signal to background ratio is estimated to be 1.14, 56.5 out of 106 being the number of inclusive $t\bar{t}$ events expected in 311 pb^{-1} of multijet data. The main $t\bar{t}$ decay channel contributors to the signal are identified as $\tau + jets$ and as $e/\mu + jets$ $t\bar{t}$ events, the latter failing the standard lepton identification cuts.

The inclusive $t\bar{t}$ contribution in the final sample is cross-checked by means of binned likelihood two-component fits to kinematical distributions of b-tagged data. The observed spectra are indeed fit to the sum of inclusive Monte Carlo $t\bar{t}$ and background templates, the latter being the result of the b-tagging rate parametrization application in the sample of data after kinematical selection. The kinematical fit-based determination of the $t\bar{t}$ fraction in the selected sample ($F_{top}^{Fit} = 0.59 \pm 0.09$) is in agreement with the result provided by the absolute b-tags counting method ($F_{top}^{matrix} = 0.47 \pm 0.05$).

A Poisson likelihood function in which the input parameters are subject to Gaussian constraints is used for a proper determination of the top pair production cross section. Assuming a top quark mass of 178 GeV/c^2, we measured:

$$\begin{aligned} \sigma_{t\bar{t}} &= 5.9 \pm 1.2(stat) \, ^{+1.4}_{-1.0}(syst) \; pb \\ &= 5.9 \, ^{+1.8}_{-1.6} \; pb, \end{aligned}$$

in agreement with Standard Model calculations as well as with previous determinations. The systematic uncertainties are dominated by trigger simulation acceptance and generator dependence effects.

The measured cross section value changes by ± 0.05 pb for each ∓ 1 GeV/c^2 change in the top mass from the initial value of 178 GeV/c^2. For instance, we measured $\sigma_{t\bar{t}} = 6.1 \pm 1.2(stat.)^{+1.4}_{-1.0}(syst.)$ pb for $M_{top} = 175$ GeV/c^2. The change is due to the varying signal selection efficiency with top quark mass.

The result presented in this work is finally combined with the latest CDF top pair production cross section determinations, in the di-lepton, single lepton, and all-hadronic channels, to obtain:

$$\sigma_{t\bar{t}}^{combined} = 7.1 \pm 0.6(stat) \pm 0.7(syst.) \pm 0.4(lum) \; pb,$$

which is the most precise measurement to date. The $\not\!\!E_T + jets$ channel is found to be very competitive with other previously established analyses, carrying a relative weight of 17% in the final cross section combination.

Appendix A

Standard τ-lepton identification

In this Appendix a short description of the standard τ-lepton identification procedure will be given. Attempts were made to optimize this work aiming at extracting the exclusive $t\bar{t} \to \tau$+jets decay. Unfortunately, as it will be described, the τ-identification efficiency is found to be $\approx 10\%$ on $t\bar{t} \to \tau$+jets events before any top-like kinematical or b-tagging requirement, preventing to isolate a sizable signal sample within the statistics of the data sample available at the time this work has been developed. The τ-lepton reconstruction using the information provided by the CDF II detector will be presented, together with the standard identification requirements and the identification efficiency evaluation on $t\bar{t} \to \tau$+jets signal.

A.1 Tau reconstruction

Tau lepton decay can be leptonic ($BR = 37\%$) or hadronic ($BR = 63\%$). The former appears as an isolated electron or muon, accompanied by undetected energy due to the two unseen neutrinos.

Tau-lepton based analyses are in general interested in hadronic lepton decays therefore in the reconstruction and identification of the hadronic system X_h of $\tau \to X_h \nu_\tau$ decay.

Hadronic tau decay appears as a very narrow jet with low track and π^0 multiplicity, consistent with low mass and with both hadronic and electromagnetic energy deposits. It is on these features that the τ-selection criteria are based (see for instance [96]).

The τ-lepton reconstruction procedure uses both calorimetric and tracking information. The τ-object building starts with a recognition of the calorimeter information looking for a seed tower with total transverse energy greater than a given threshold, $E_T^{tow\ seed}$. Then, adjacent shoulder towers with energies $E_T^{tow\ sh}$ are added to form the calorimetric τ-cluster. Due to the narrowness of the τ-jet, the total number of towers contributing to the cluster is generally small. Once the calorimeter recognition is done, the τ-reconstruction procedure looks for a seed track to be associated to the calorimetric τ-like cluster. The track must point at the center of the calorimetric cluster and is required to exceed a transverse momentum threshold, $P_T^{trk\ seed}$. If several such tracks are found, the one with the highest P_T is chosen. The direction identified by the seed track is then used as a reference for all the following steps of the procedure.

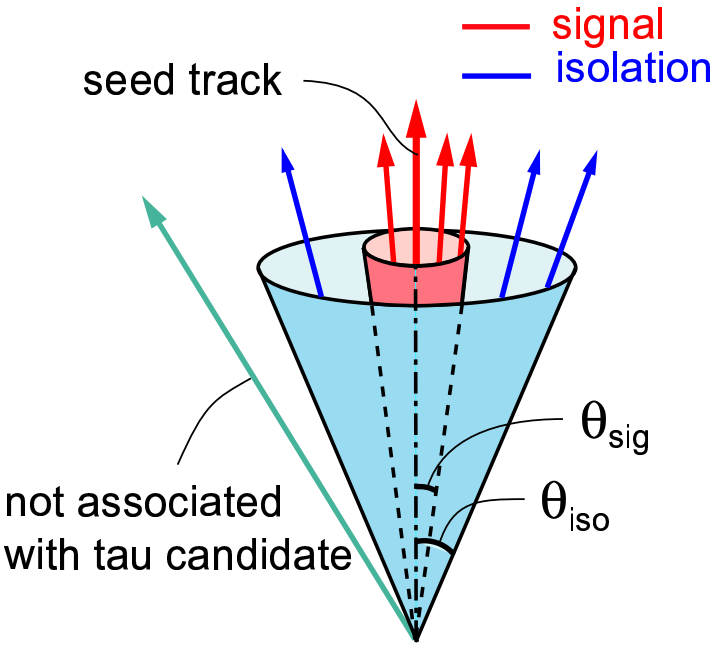


Figure A.1: Definition of isolation and signal annulus with respect to the τ seed track direction[96].

Other COT tracks are associated to the τ-object according to their closeness to the selected seed track. These shoulder tracks are required to have $P_T > P_T^{trk\ sh}$ and a small Δz separation from the seed track z intercept. Finally, shoulder tracks have to be reconstructed within 3-D angle θ_{iso} with respect to the reference direction. Another angle, $\theta_{sig} < \theta_{iso}$, is used to define the τ-lepton signal region: tracks within angle θ_{sig} are associated to the τ-object, while those with $\theta_{sig} < \theta < \theta_{iso}$ are used to veto τ candidates, being in general associated to false identifications.

The angles θ_{iso} and θ_{sig} defining respectively the isolation and signal annulus are shown in Fig. A.1.

The Run-I definition of the signal cone in which looking for tracks and π^0's was not energy dependent. The signal cone was defined to be the one within 10° from the seed track direction and the isolation annulus was that between 10° and 30° from the seed. More recent investigations have shown that the signal cone width is strongly correlated with the τ energy; therefore in Run II a new cone and annulus definition is used.

In particular the signal cone is defined to be the

$$\min\left(0.17;\ \frac{5.0\ rad/GeV}{E^{tau\ clu}}\right)\ rad. \tag{A.1}$$

On the other hand, the isolation annulus is defined to be the one between the τ's signal cone and 30°. This particular choice is found to provide a better background rejection at high energies.

The last step of the procedure consists in associating π^0 candidates to the τ's. The recipe is the same used for the assignment of the shoulder tracks. The π^0 candidates are

variable		cut
$E_T^{tow\ seed}$	$\geq$	6.0 GeV
$E_T^{tow\ sh}$	$\geq$	1.0 GeV
N^{tow}	$\leq$	6
θ_{sig}	$=$	$\min\left(0.17;\ \frac{5.0\ rad/GeV}{E^{tau\ clu}}\right)\ rad.$
θ_{iso}	$=$	0.52 rad
$P_T^{trk\ seed}$	$\geq$	6.0 GeV
$P_T^{trk\ sh}$	$\geq$	1.0 GeV
$\Delta z^{sh\ trk}$	$\leq$	10.0 cm

Table A.1: Offline reconstruction cuts for τ candidates.

reconstructed using CES and CEM information. In particular CES clusters of width 5 wires/strips are used for the position determination (z and ϕ), while the CEM returns the energy measurement.

The reconstruction of π^0's consists of two steps: the matching of wire to strip clusters and then the association of the matched CES cluster to the CEM energy deposit. In the case of multiple matches associated with the same CEM tower[1], the energy is shared according to the CES energy of each cluster. At this stage, the π^0 candidates are associated to the tau object[2].

The offline TauFinderModule performs the procedure described in this section with the selection listed in Tab. A.1.

A.2 Tau identification

Regardless of the crowded topology created by a $t\bar{t}$ decay, in [90] we showed how a standard τ-identification procedure preserves its ability at selecting good τ-lepton candidates. For this reason, we decided to apply for the τ-candidate selection, in $t\bar{t} \to \tau$+jets events, the same cuts used in the measurement of the $W \to \tau\nu$ cross section [97].

The τ selection cuts applied are listed in the following:

- tracks $|\eta| < 1.0$;
- seed track $P_T > 4.5\ GeV$;
- 9 $cm < |Z_{CES}| < 230\ cm$, where Z_{CES} is the z coordinate of the seed track extrapolated to the CES detector;
- Cosmic veto: $|d_0| < 0.2\ cm$, where d_0 is the impact parameter of the seed track;
- $E_T^\tau > 25\ GeV$, being E_T^τ the total transverse energy of the τ candidate;

[1] CEM energies are corrected for contribution of charged tracks (assumed to be pions).

[2] π^0 candidates can be both real π^0 or photons: for π^0 energies in the range of $5 - 6\ GeV$ the two photons from the neutral pion decay can not be resolved and appear as a single cluster. This is also the case of low energy pions when the energy asymmetry of the two photons is relatively large.

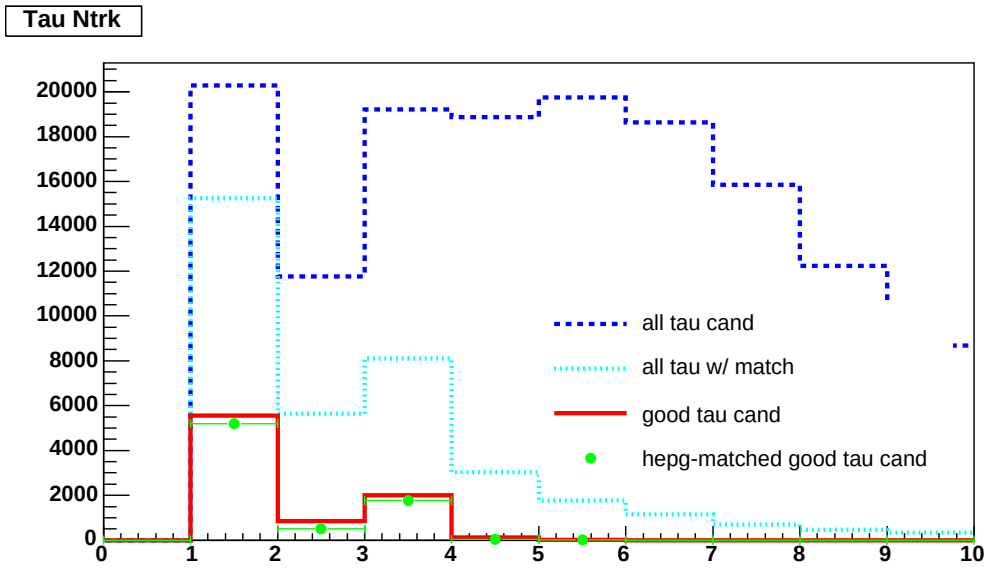


Figure A.2: Impact of the τ-identification cuts on the charged tracks multiplicity of the τ candidates in $t\bar{t} \rightarrow \tau$+jets events. The distributions for those candidates having a corresponding τ in the HEPG bank are also shown, before and after the selection cuts.

- $E_{0.4}/E_{\tau} < 0.1$: calorimetric isolation requirement defined as the sum of the electromagnetic and hadronic energy in the calorimetric tower not belonging to the τ-cluster inside a cone of radius 0.4 divided by the τ energy;

- $\sum P_T(30°)/P_T(\tau) < 0.1$: tracking based isolation cut defined as the transverse momentum sum of tracks inside a 30° cone from the seed track direction divided by the total transverse momentum of the τ candidate. The minimum P_T threshold for tracks to be considered in the sum is $P_T^{min} = 0.4\ GeV$;

- $EMFR < 1.0 - 0.15 * \frac{\sum P}{E}$, where $EMFR$ is the electromagnetic fraction of the calorimeter energy deposit associated to the candidate and $E/\sum P$ is the ratio between energy and momentum of the reconstructed τ;

- $M(tracks + \pi^0) < 1.8\ GeV$, where $M(tracks + \pi^0)$ is the invariant mass calculated using tracks and π^0 associated to the τ cluster;

- $N_{iso}^{tracks} = 0$, where tracks $P_T^{min} = 1\ GeV$. This cut requires the number of charged tracks reconstructed within the isolation annulus to be zero. The annulus is defined as the region between the signal cone, $\min\left(0.17;\ \frac{5.0\ rad/GeV}{E^{tau\ clu}}\right)\ rad$, and 30° cone around the seed track direction.

The impact of these requirements on the charged track multiplicity distribution of τ candidates in $t\bar{t} \rightarrow \tau$+jets Monte Carlo events is shown in Fig. A.2.

At this stage we can evaluate the purity of this procedure by calculating the ratio between the number of τ candidates passing the requirements which have a corresponding

particle in the HEPG bank, over the number of τ candidates passing the identification cuts. This is found to be around 93%, highlighting the effectiveness of this procedure.

The efficiency of the τ identification procedure here described is defined as the ratio between the number of candidates passing the selection over the number of $t\bar{t} \rightarrow \tau$+jets after the trigger simulation and prerequisites. This is found to be $\epsilon(\tau - id_{alone}) = (11.18 \pm 0.12)\%$.

In 311 pb^{-1} of TOP_MULTI_JET triggered data, after the kinematical selections, outlined in Sec. 5.2.1, we are left with 35 $t\bar{t} \rightarrow \tau$+jets events expected, before any b-tagging requirement. The application of the τ-identification on this sample will reduce the signal expectation to ~ 3 events, preventing any kind of significant measurement.

Appendix B

High-P_T lepton selection cuts

In the work described in this document we rejected high-P_T well reconstructed electrons or muons used by other $t\bar{t}$ pair production cross section analyses in the lepton+jet channel.

Analyses relying on the identification of leptons in general use dedicated lepton triggers, and use the requirements listed below for lepton selection.

- Electron selection (see also Section 3.3.8):
  - $E_T > 20\ GeV$;
  - $P_T > 10\ GeV/c$;
  - $|\eta| < 1.0$;
  - $E_{HAD}/E_{EM} < 0.055 + 0.00045 \times E$;
  - $E/p < 2.0$ if $P_T < 50 GeV/c$;
  - $|z_0| < 60\ cm$;
  - 3 COT Axial 5-hit superlayers;
  - 2 COT Stereo 5-hit superlayers;
  - isolation < 0.1;
  - $L_{shr} < 0.2$;
  - $|\Delta z| < 3\ cm$ for CES strip cluster Δz;
  - $-3.0 < Q\Delta x < 1.5\ cm$ for CES strip cluster Δx;
  - Strip $\chi^2 < 10$.
- Muon selection (see also Section 3.4.1):
  - $P_T > 20\ GeV/c$;
  - $|\eta| < 1.0$;
  - $E_{HAD} < 6.0 + \max[0, 0.0280(P - 100 GeV/c)] GeV$;
  - $E_{EM} < 2.0 + \max[0, 0.0115(P - 100 GeV/c)] GeV$;
  - $|z_0| < 60\ cm$;

- 3 COT Axial 5-hit superlayers;
- 2 COT Stereo 5-hit superlayers;
- isolation < 0.1;

For electrons the isolation is defined as $iso = E_T^{iso}/E_T^{cluster}$, where $E_T^{iso} = E_T^{0.4} - E_T^{cluster}$ is the transverse energy in a cone of radius $\Delta R = 0.4$ around the electron cluster excluding the electron cluster energy. On the other hand the isolation variable for muon identification is defined by: $iso = E_T^{iso}/P_T$ where $E_T^{iso} = E_T^{0.4} - E_T^{tower}$ is the difference between the energy in a cone of radius $\Delta R = 0.4$ around the muon track and E_T^{tower} is the amount of energy in the calorimeter tower associated with the muon track.

Appendix C

Trigger systematics

In this Appendix details on the evaluation of the systematic uncertainty on the trigger efficiency for $t\bar{t} \to \not\!\!E_T + jets$ cross section measurement analysis will be given. Two possible approaches will be described. Both methods rely on the comparison of trigger turn-on curves between Monte Carlo and data events. In general a trigger turn-on curve is calculated by dividing an offline variable distribution obtained after having applied the trigger cuts to the same variable distribution before any requirement.

The trigger efficiency on Monte Carlo events is calculated using the information provided by the trigger simulation. This procedure could be affected by systematic uncertainties arising from imperfect or not well modelled emulation of the trigger hardware as well as of the trigger decisions. For this reason in the following we will provide a way to evaluate the trigger acceptance systematics using collider data information.

We used a "Tower-10" data sample that is collected using the same L1 requirement as the TOP_MULTI_JET trigger and thus provides a base sample in which to determine the effect of the calorimetric clusters E_T and $\sum E_T$ requirement of the TOP_MULTI_JET trigger. In particular, Tower-10 data is triggered by applying 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.

The corresponding integrated luminosity of the sample is $196 \pm 12\ pb^{-1}$. This sample is used to derive data trigger turn-on curves.

On the other hand different sample on Monte Carlo events can be used to derive Monte Carlo hard-wired turn-on curves to be used for a comparison with the turn-on curves extracted from the data. We will start evaluating trigger turn-on curves calculated on $t\bar{t}$ inclusive events, which are used in our analysis. Afterward, trigger turn-on curve comparisons using different Monte Carlo samples, such as $b\bar{b}$ and multi-jet events will be discussed.

"Tower-10" data, as already noted, satisfies the same Level 1 requirements imposed by the TOP_MULTI_JET trigger and therefore can be used to extract the trigger turn-on curves as a function of some offline variables (*i.e.* offline raw jet E_T and offline raw $\sum E_T$), by applying by hand the L2 TOP_MULTI_JET requirements, on a data-driven basis. The

same approach can be taken for Monte Carlo events just after having asked for the L1 requirements to be satisfied.

Any systematic effect to the trigger acceptance, resulting from possible mis-matches between Monte Carlo and data trigger turn-on curves, depends on the kinematical selection used for the signal extraction as well as on the peculiar kinematics of the Monte Carlo sample used for the comparison.

In particular, the evaluated systematic uncertainty on the acceptance is directly connected to the applied kinematical cuts. To be more clear, some kinematical selection can move the bulk of the signal events away from the kinematical region which is affected by the trigger turn-on, thus reducing the effect of any turn-on mis-match. On the other hand, some discrepancies can be produced by kinematic properties of the particular Monte Carlo sample used, more than any imperfect simulation of the trigger cuts.

C.1 Tower-10 data vs $t\bar{t}$ Monte Carlo

We will illustrate here the work we have performed in order to extract the Monte Carlo turn-on curves for the sub sample of $t\bar{t}$ inclusive events satisfying our kinematical selection (Sec. 5.2.1):

- $N_{jet}(E_T^{L5} \geq 15\ GeV, |\eta| \leq 2.0) \geq 4$;
- $\not\!\!E_T/\sqrt{\Sigma E_T} \geq 4\ GeV^{1/2}$;
- $\min \Delta\phi(\not\!\!E_T, jets) \geq 0.4\ rad.$

The procedure we followed to estimate the trigger related systematic starts by selecting events with at least four jets and four L2 calorimetric clusters[1]. We then asked for a matching between clusters and jets and in particular, in order to preserve the energy hierarchy between clusters and jets, we required a matching between the 4^{th} cluster and the 4^{th} jet in the event, once both clusters and jets are ordered by decreasing transverse energy. Even though the jet-to-cluster matching is not required at trigger level, we found necessary to enforce it in order to deal with a clean sample of events in which the trigger behavior could be clearly reproduced. We will return to this issue in the following.

After the matching requirement, fourth jet E_T^{L5} distributions can be derived, both for inclusive Monte Carlo $t\bar{t}$ and Tower-10 data events, before and after having required the L2 TOP_MULTI_JET trigger selections to be satisfied. The ratio between the jet E_T distributions before and after the trigger requirements allows to build trigger turn-on curves as a function of the fourth L5-corrected jet E_T. These distributions are shown in Fig. C.1, where for Monte Carlo inclusive $t\bar{t}$ events, surviving the kinematical selection, the filled histogram in the top left panel shows the 4^{th} leading jet L5-corrected E_T distribution when it has a match with the 4^{th} L2 calorimetric cluster. On the other hand, the dotted distribution, shown in the same panel, displays the 4^{th} leading jet L5-corrected E_T spectra

[1]Clusters are of type jet, *i.e.* are reconstructed by the clustering algorithm using the settings: seed tower threshold = 3 GeV (for CEM+CHA and PEM+PHA) and shoulder tower threshold = 1 GeV (for CEM+CHA and PEM+PHA).

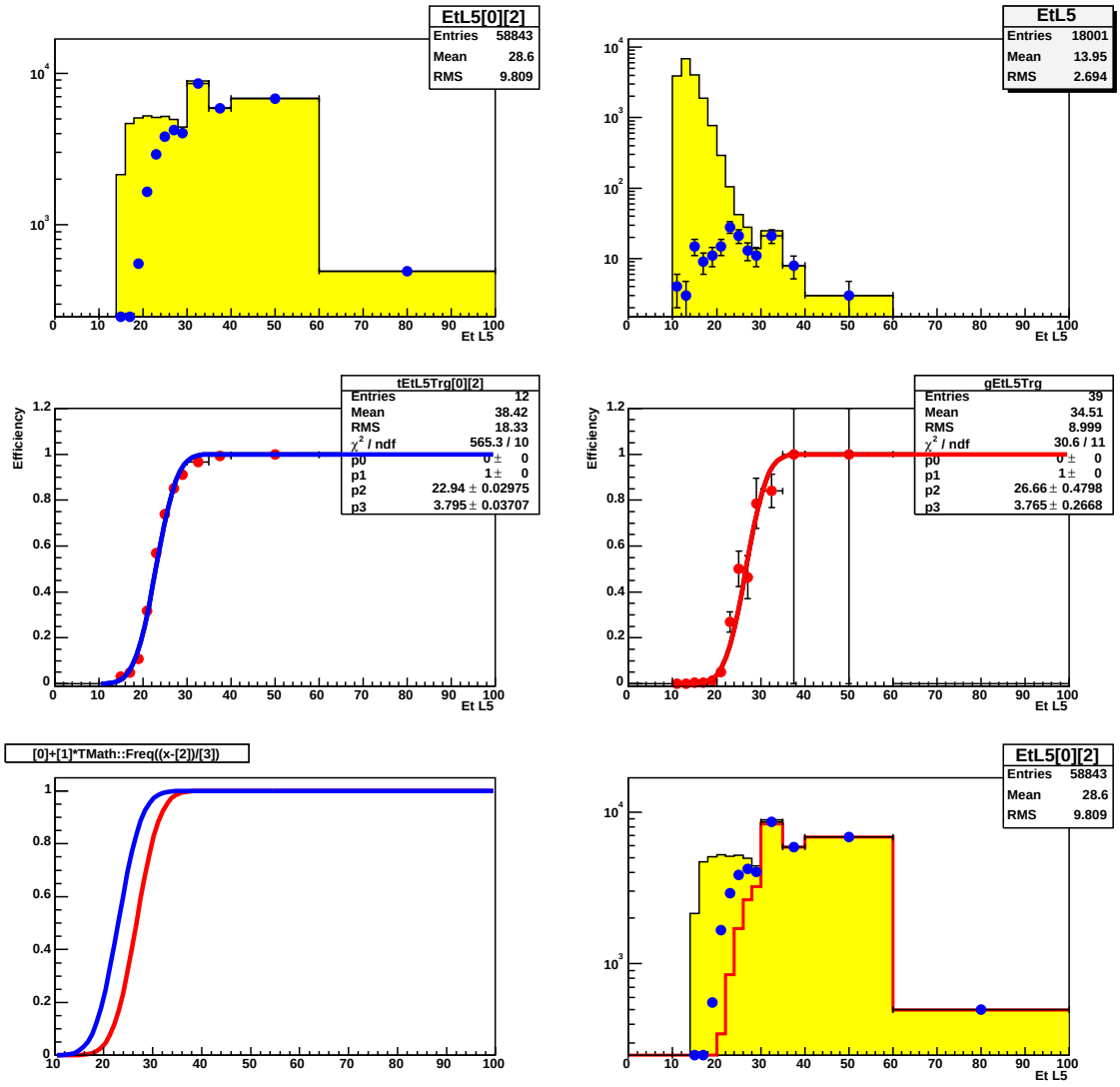


Figure C.1: Top left panel: Filled (dotted) histogram represents the 4^{th} leading jet E_T^{L5} distribution when a match with the 4^{th} L2 calorimetric cluster (with $E_T \geq 15\ GeV$ and $\sum E_T^{L2} \geq 125\ GeV$) is required on Monte Carlo inclusive $t\bar{t}$ events after our kinematical selection. The top right panel shows instead the corresponding distributions obtained from "Tower-10" data (no kinematical selection is applied in this case due to the poor data statistics). The middle left plot shows the simulated trigger turn-on curve for having at least 4 L2 clusters above 15 GeV for the subsample of events in which the 4^{th} leading jet has a match with the 4^{th} L2 cluster as a function of the offline jet E_T^{L5}. In the middle right plot the turn-on curve obtained from "Tower-10" data is shown. On the bottom left panel both the turn-on curves from Monte Carlo trigger simulation (blue) and from data (red) are over imposed. The bottom right plot shows the effect of the data turn-on curve when applied on the Monte Carlo 4^{th} jet E_T^{L5} spectra (red line). The filled and dotted histograms of the top left panel are shown here again for reference.

when the L2 cluster to which the jet is matched exceeds the 15 GeV threshold and the event satisfies the $\sum E_T^{L2} \geq 125\ GeV$ requirement. On the top right panel of Fig. C.1 the same kind of distributions are shown for "Tower-10" data events (no kinematical selection is applied in this case due to the lack of statistics).

Dividing the E_T^{L5} spectra for jets that have a match with a 15 GeV L2 calorimetric cluster by the jet E_T^{L5} distribution in which there is no energy threshold applied to clusters, we derived the turn-on curves shown in the middle row of Fig. C.1 for inclusive $t\bar{t}$ (left plot) and "Tower-10" data (right plot) events, respectively. For an easy comparison both turn-on curves from Monte Carlo and data respectively are superimposed in the bottom left panel of Fig. C.1. Finally, in the bottom right panel is shown (red line) the 4^{th} jet E_T^{L5} spectra after a matching with a 4^{th} 15 GeV L2 cluster one would obtain by using the turn-on curve extracted from data instead of the one built-in in the Monte Carlo trigger simulation together with the already known 4^{th} jet raw E_T distribution for Monte Carlo.

By comparing the 4^{th} cluster-matched jet E_T^{L5} spectra we see in the Monte Carlo to the one obtained by plugging in the turn-on curve extracted from data, we can evaluate the systematic effect to the trigger efficiency determination from Monte Carlo events alone. In particular the difference in terms of distribution integrals divided by the total number of cluster-matched jets in this case yields:

$$\Delta\epsilon^{trg} = \frac{\int \mathcal{D}^{4^{th}jet}_{clu15;set125}(E_T)dE_T - \int \mathcal{D}^{4^{th}jet}(E_T) \times F^{turn-on}_{gjs10d}(E_T)dE_T}{\int \mathcal{D}^{4^{th}jet}(E_T)dE_T} = (14.8 \pm 0.2)\%, \tag{C.1}$$

where $\mathcal{D}^{4^{th}jet}_{clu15;set125}(E_T)$ is the E_T^{L5} distribution of the 4^{th} jet in the event matched with the 4^{th} $E_T \geq 15\ GeV$ L2 cluster for inclusive $t\bar{t}$ events with the additional requirement $\sum E_T^{L2} \geq 125\ GeV$; $\mathcal{D}^{4^{th}jet}(E_T)$ is the E_T^{L5} distribution of the 4^{th} jet in the event matched with the 4^{th} L2 cluster, again derived from Monte Carlo; and finally $F^{turn-on}_{gjs10d}(E_T)$ is the turn-on curve as a function of the 4^{th} jet E_T^{L5} derived from "Tower-10" data. Here and in the following errors are statistical only and do not include the uncertainties returned by the fitting procedure on the curve parameters.

The value of the trigger acceptance systematic, $\Delta\epsilon^{trg}$, explained above accounts for all the requirements of the TOP_MULTI_JET trigger. It should be applied as systematic uncertainty to the trigger efficiency calculated on a TRIGSIM++ basis for inclusive $t\bar{t}$ events. Anyway, it should be noted that $\Delta\epsilon$ is calculated on a subsample of events in which there is a match between the 4^{th} jet and the 4^{th} L2 cluster. For this reason, the application of the $\Delta\epsilon$ to the whole $t\bar{t}$ sample rests on the assumption that it has there the same value as in the subsample obtained matching jets to clusters.

In order to check the reliability of this assumption, we performed a scan of Monte Carlo trigger turn on curves derived with different matching requirements for jets and clusters. The jet-to-cluster matching used for the results reported above was performed using a $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \leq 0.4$ cut. We derived $t\bar{t}$ turn-on curves with ΔR cut set at 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7 and compared them with the "Tower-10" curves determined with fixed $\Delta R \leq 0.4$ requirement. The variation of the ΔR cut on the Monte Carlo essentially varies artificially the matching efficiency, and allows in this way to check for any dependence on the matching requirement of the trigger turn-on determination

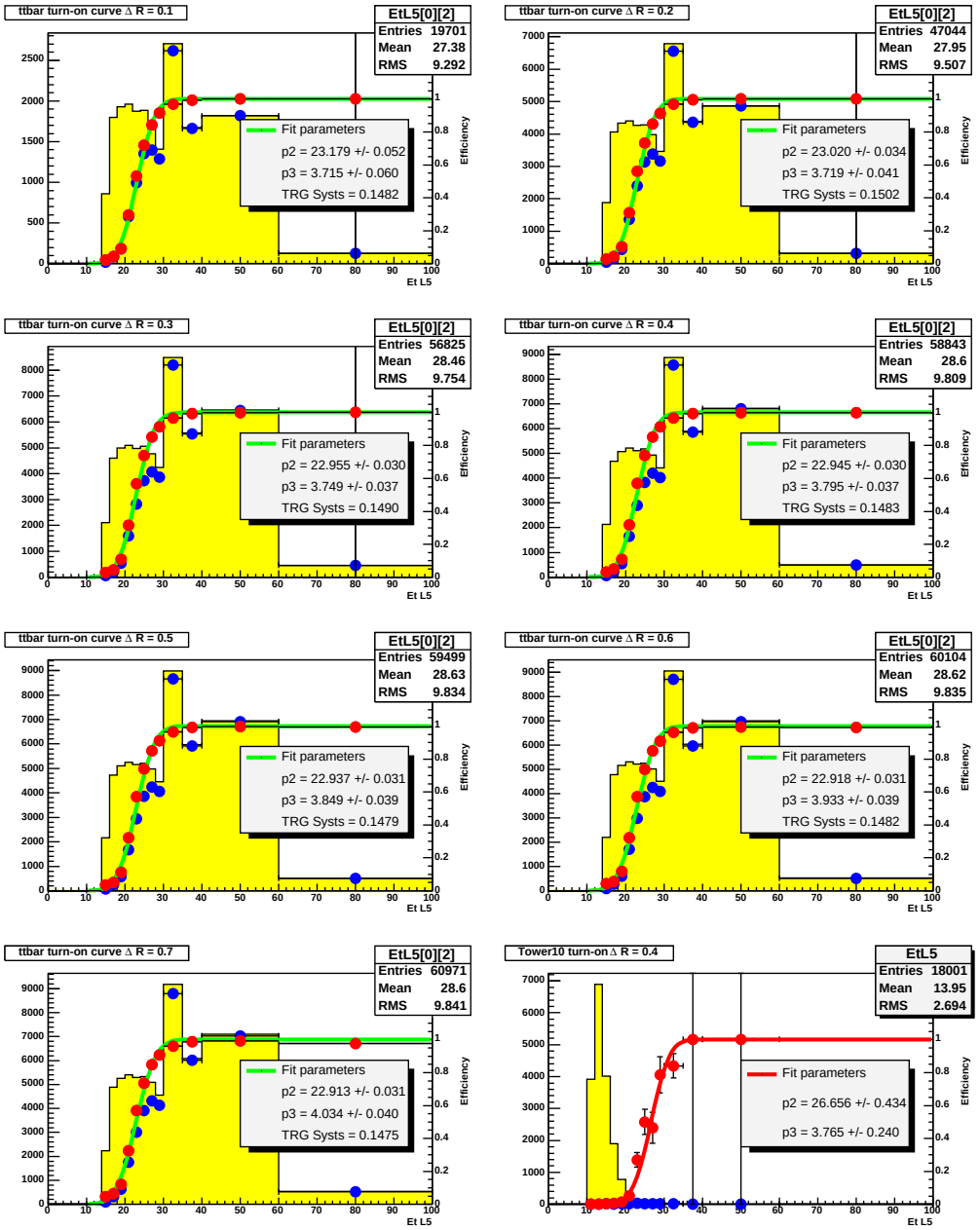


Figure C.2: Monte Carlo turn-on curves and corresponding trigger systematics calculated with different cluster-to-jet matching requirements: $\Delta R = [0.1, 0.7]$. In the top left panel of the figure, the filled (dotted) histogram represent the 4^{th} leading jet E_T^{L5} distribution when a match within $\Delta R \leq 0.1$ with the 4^{th} L2 calorimetric cluster (with $E_T \geq 15\ GeV$ and $\sum E_T^{L2} \geq 125\ GeV$) is required on Monte Carlo inclusive $t\bar{t}$ events after our kinematical selection; the corresponding turn-on curve is superimposed. Same distributions and turn-on curves are shown on the other panels for $\Delta R \leq [0.2, 0.7]$. For convenience, the bottom right panel, reports the "Tower-10" data trigger turn-on curve for $\Delta R \leq 0.4$.

jet-to-cluster ΔR cut on $t\bar{t}$ MC	0.1	0.2	0.3	0.4	0.5	0.6	0.7
calculated trigger systematic [%]	14.82	15.02	14.90	14.83	14.79	14.82	14.75

Table C.1: Variation of the trigger systematics as a function of the ΔR cut used for the jet-to-cluster matching for Monte Carlo $t\bar{t}$ events.

procedure. Results of this check are drawn in Fig. C.2.

Tab. C.1 summarizes the results of this check, and allows to conclude that the matching requirements does not affect significantly the determination of the trigger turn-on curves.

The cluster-to-jet matching was needed in order to derive the curve for a subclass of clean events, where jet and cluster hierarchies were respected passing from trigger to offline level; furthermore to avoid situations (especially as far as the Monte Carlo is concerned) in which a large amount of energy is reconstructed as a single L2 cluster by the trigger hardware but then split among different offline jets yielding effects that can not be easily accounted for in the turn-on curve comparison between data and Monte Carlo.

The conclusion of this study is that by looking at the differences between the trigger turn-on curve built-in in the $t\bar{t}$ Monte Carlo sample after the kinematical selection and the one obtained from "Tower-10" data, a trigger acceptance systematics of $(14.8 \pm 0.2)\%$ should be assumed.

Constructing the Monte Carlo trigger turn-on curve after the kinematical selection, in some sense could over-estimate the systematic effect. Indeed, the trigger acceptance systematic evaluated using $t\bar{t}$ events before kinematical selection is found to be $(10.4 \pm 0.2)\%$. This behavior is due to the tight cut on the $\not\!\!E_T^{sig} = \not\!\!E_T/\sqrt{\sum E_T}$ which is applied for the event selection. The effect of this cut is to preferentially select $t\bar{t}$ Monte Carlo events close to the kinematical region affected by trigger turn-on effect (*i.e.* high $\not\!\!E_T^{sig}$ in part selects low $\sum E_T$ events). Unfortunately, due to the lack of statistics in the Tower-10 sample, the kinematical selection can not be applied when deriving the data turn-on curve. For this reason we chose, in order to be very conservative, to quote 14.8% as trigger systematics value, even if the value is probably over-estimated as explained above.

C.1.1 Trigger systematics and jet energy scale

The trigger turn-on curves for $t\bar{t}$ events are obtained using L5 corrected jet E_T distribution. The jet energy response is subject to systematic uncertainties, as described in Sec. 3.3.5 and Sec. 7.1.7, which could affect the trigger systematic determination, since trigger turn-

Jet energy response	standard	$+1\sigma$	-1σ
ϵ_{trg}[%]	78.25 ± 0.06	77.95 ± 0.06	78.64 ± 0.06

Table C.2: Effect of the jet energy response systematics on trigger efficiency. The Efficiencies are evaluated over a sub-sample of $t\bar{t}$ inclusive events in which there is a match between the 4^{th} L2 cluster and the 4^{th} L5-corrected jet.

on curves are derived using corrected jet E_T distributions. For this reason we evaluated the possible dependence of the trigger turn-on determination in Monte Carlo events on the jet energy scale systematics.

For each set of jet corrections (standard, and $\pm 1\sigma$) the TOP_MULTI_JET trigger efficiency is calculated as the ratio of the jet E_T^{L5} distribution integrals obtained applying or not applying the L2 TOP_MULTI_JET trigger requirements, on inclusive $t\bar{t}$ Monte Carlo events before any kinematical selection. L3 requirements are not accounted for, since they are almost 100% efficient on signal events. Results are shown in Tab. C.2.

The trigger efficiency values we found in this way are higher than the value of $(63.35 \pm 0.04)\%$ we quoted for our analysis[2] because in this case they are evaluated only on the sub-class of events in which there is a match between the 4^{th} L2 cluster and the 4^{th} L5-corrected jet.

Taking half of the difference between the efficiencies obtained with $\pm 1\sigma$ jet energy correction variation, the impact of the jet correction systematics on the trigger systematic uncertainty can be set to $\Delta\epsilon_{trg}^{JES} = 0.5\%$. The impact of the jet correction uncertainty is found in this way to be negligible, when compared to the 14.8% trigger systematic previously quoted.

C.1.2 Trigger systematics versus $\sum E_T$

The systematics effect on the trigger simulation (14.8%) dominates the total systematic uncertainty of our cross section measurement (Sec. 7.1.9). For this reason several attempts at reducing it have been pursued. As shown in [86] the trigger systematics are found to be negligible in the case of the all-hadronic top pair production cross section analysis in which, aside $N_{jet} \geq 6$ and others requirements, a tight cut on $\sum E_T$ was imposed.

Fig. C.3 shows the optimization of an extra cut on $\sum E_T$ after having applied the kinematical selection and the additional requirement for at least one positive SECVTX b-tagged jet on our analysis. The background distribution is extracted from the tagging rate parametrization application in the data sample after kinematical selection. As shown by the $S/\sqrt{N}$ ratio in the bottom right panel of Fig. C.3 there is not much left to gain from a statistical point of view by adding a $\sum E_T$ cut. Anyway from the previous reasoning a $\sum E_T$ cut could reduce the systematic on the trigger simulation.

The trigger acceptance systematic uncertainty was studied as a function of an extra analysis cut on $\sum E_T$ in the range $[50, 275]$ GeV in steps of 25 GeV.

The results of this study are shown on the first column on Tab. C.3. A $\sum E_T$ requirement at 225 (250) GeV could reduce by 10% (21%) the value of the trigger systematic we quoted in our analysis.

Anyway, the effect of any extra selection variable in general adds new sources of systematic uncertainty. This is the case for the systematic effect related to jet energy response that is indeed found to increase for tight cuts on $\sum E_T$ (middle column on Tab. C.3).

The right column on Tab. C.3 reports the sum in quadrature of the systematics uncertainties due to the trigger simulation and to the jet energy response. Unfortunately, an additional $\sum E_T$ cut does not help to reduce the overall systematic uncertainty.

[2] See Tab.4.1 in Chapter 4.

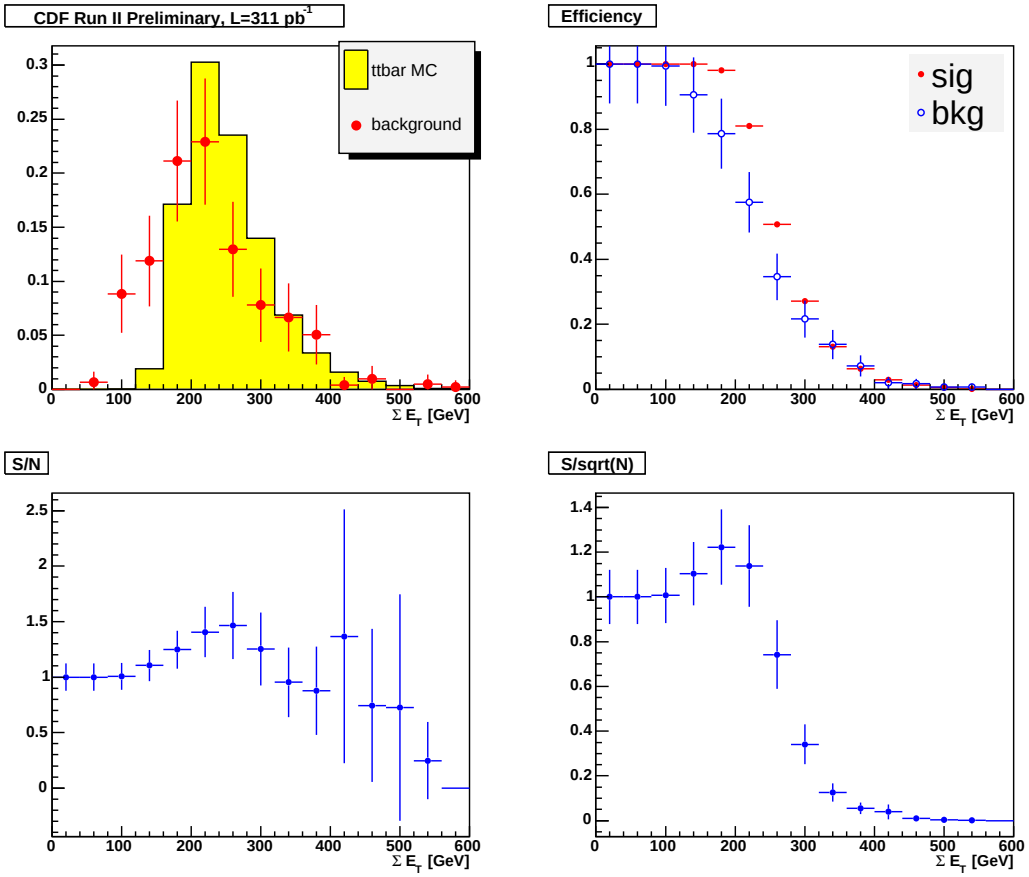


Figure C.3: Optimization of a $\sum E_T$ cut after kinematical selection and ≥ 1 positive tag requirements. Top left: distributions for signal and background. Top right: efficiencies both for inclusive $t\bar{t}$ Monte Carlo events and background as a function of the applied cut. Bottom left: signal to background ratio. Bottom right: signal over square root of background ratio (cut gain).

Fig. C.4 shows graphically the results of Tab. C.3.

C.1.3 Trigger systematics versus N_{jet}

Another attempt to reduce the effect of the trigger systematic on the cross section measurement consists in splitting the measurement by jet multiplicity. Indeed, the trigger systematic is reduced when computed over sub-samples of $t\bar{t}$ events selected by requiring $N_{jet} = 5, 6$, or 7. However, as explained before for a cut on $\sum E_T$, if from one hand the cross section measurement made by jet multiplicity can decrease the overall trigger systematics impact, on the other hand it increases the contributions from the jet energy response uncertainties to the measurement: a fixed cut on the event jet multiplicity is indeed affected by promotion or regression effects as far as the number of jets is concerned due to jet energy response variations within $\pm 1\sigma$. For these reasons, even this attempt is found to not provide any net improvement.

$\sum E_T$ cut	Trig Syst [%]	JES Syst [%]	Total [%]
50	14.84	1.48	14.92
75	14.84	1.48	14.92
100	14.86	1.48	14.93
125	15.06	1.54	15.14
150	15.89	2.23	16.05
175	16.37	5.10	17.15
200	15.40	9.19	17.94
225	13.49	13.68	19.21
250	11.73	16.81	20.50
275	10.39	18.93	21.59

Table C.3: Trigger and jet energy response systematics as a function of $\sum E_T$ cut.

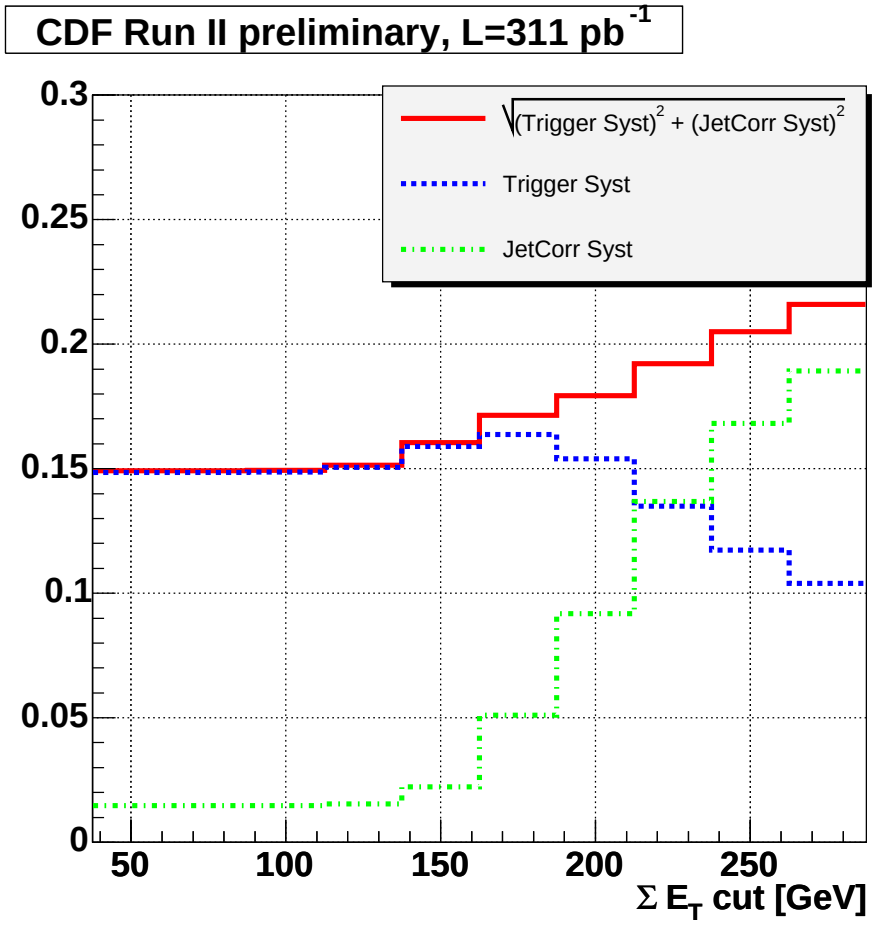


Figure C.4: Trigger and jet energy response systematics as a function of $\sum E_T$ cut.

C.2 Tower-10 data vs $b\bar{b}$ and multi-jet Monte Carlo

In the introduction of this Appendix we mentioned that the observed discrepancies in terms of trigger turn-on curves can partly originate from kinematical properties of the

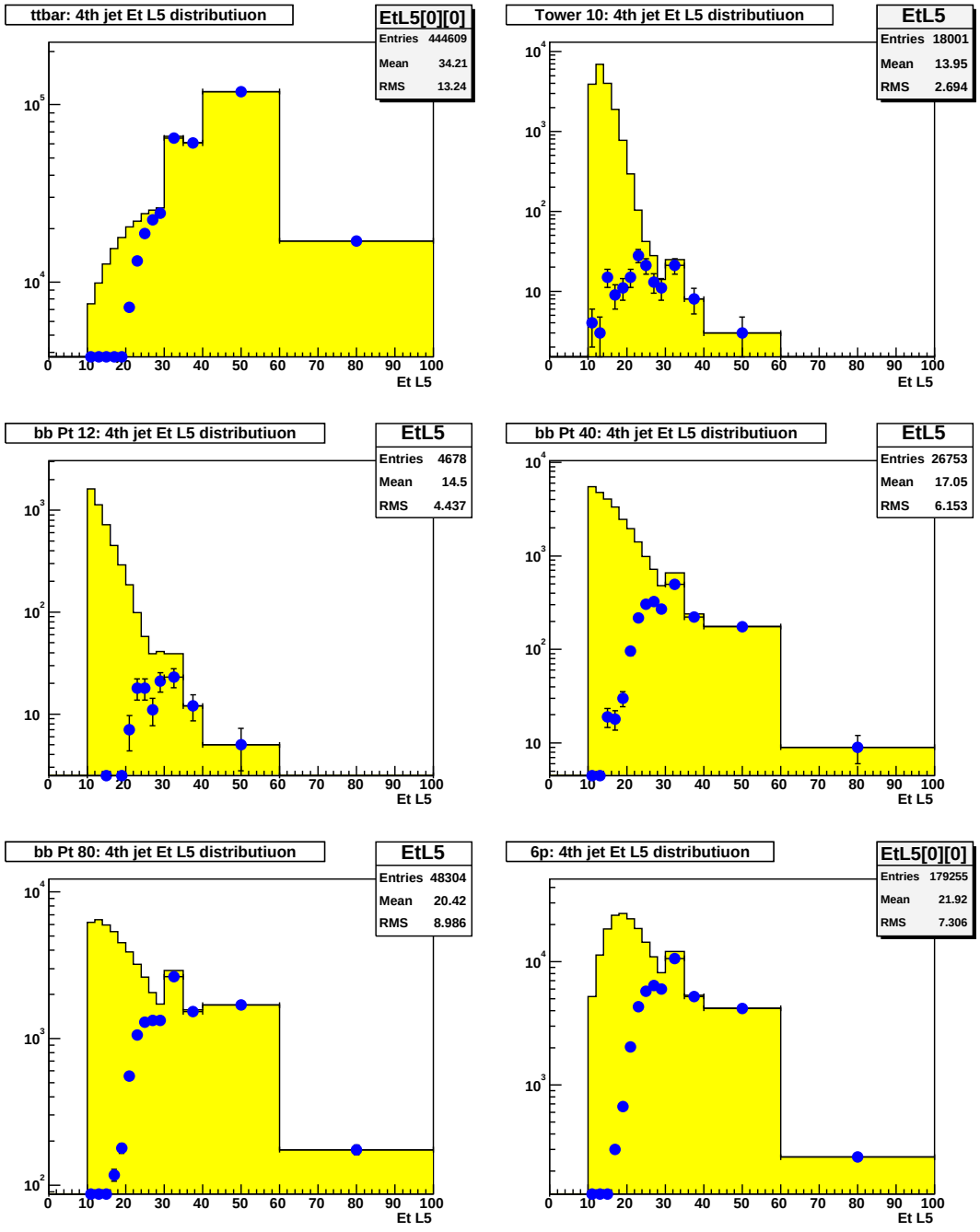


Figure C.5: Fourth-jet E_T^{L5} distributions for different Monte Carlo samples and Tower-10 data.

considered samples. In this section some studies performed in order to distinguish the pure trigger simulation from the kinematics-induced systematics will be described.

Fig. C.5 shows the fourth leading jet E_T^{L5} distributions for different Monte Carlo samples before (filled histogram) and after (dotted histogram) the application of TOP_MULTI-_JET L2 requirements, as well as for Tower-10 data. The top left distribution refers to

Sample	Tower-10	$t\bar{t}$	$b\bar{b}$, $P_T > 12\ GeV$	$b\bar{b}$, $P_T > 40\ GeV$	$b\bar{b}$, $P_T > 80\ GeV$	$6-P$
$< E_T^{4^{th}} >$	13.95	34.21	14.50	17.05	20.42	21.92

Table C.4: Mean of 4^{th} jet E_T distributions for Tower-10 data, $t\bar{t}$, $b\bar{b}$ and multijet Monte Carlo samples. E_T values are in GeV.

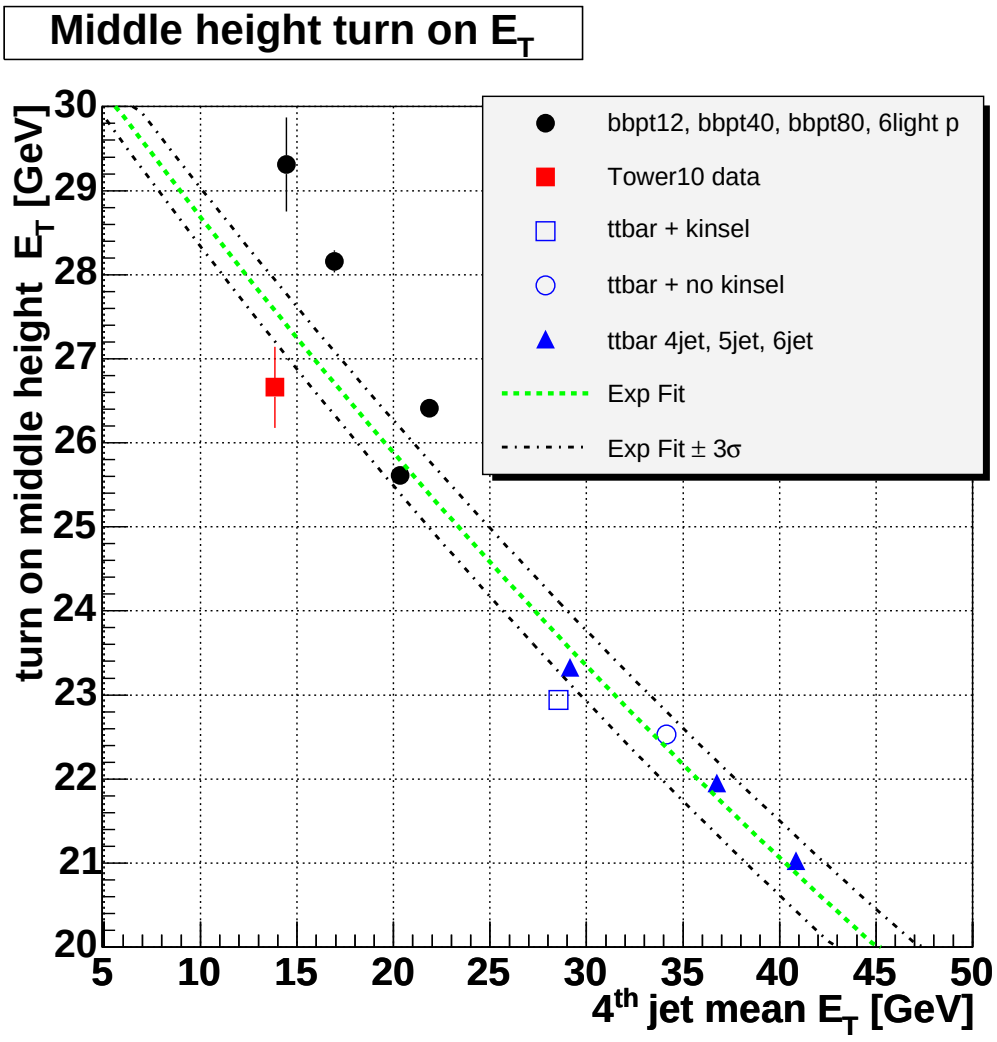


Figure C.6: Turn-on middle height E_T as a function of the $< E_T^{4^{th}} >$ of the sample.

inclusive $t\bar{t}$ Monte Carlo events before any kinematical selection. On the top right panel, the same distributions are provided for Tower-10 data. These E_T distributions, with the exception of the application of the kinematical requirements for the $t\bar{t}$ sample, were used to derive the trigger systematics affecting our analysis, as explained in the previous section. The mean of the distributions are significantly different for the two samples, being 34.21 GeV and 13.95 GeV for $t\bar{t}$ and Tower-10 data, respectively.

Additional Monte Carlo samples were considered in this study in order to investigate any trigger turn-on curve dependence on the hardness of the 4^{th} jet E_T spectra (as usual,

a jet-to-cluster matching is required within $\Delta R = 0.4$). In particular we used PYTHIA $b\bar{b}$ samples generated with a $P_T^{min} = 12,\ 40,\ 80\ GeV$ and an ALPGEN+HERWIG six-light-partons sample. The fourth leading jet E_T^{L5} distributions for these processes are shown in the middle left, middle right, bottom left and bottom right panels of Fig. C.5, respectively.

Tab. C.4 report the fourth cluster-matched mean E_T^{L5} for each of the considered samples. In the case of $b\bar{b}$ samples, the fourth jet originates from initial or final state radiation effects and thus the distributions do not usually include jets from b-quarks.

The E_T distribution for these additional samples are closer to the spectrum observed in Tower-10 data, than that obtained from the $t\bar{t}$ sample. For this reason, following the same recipe highlighted in the previous section we derived trigger turn-on curves for each of these additional Monte Carlo samples.

The E_T corresponding to the middle height of the turn-on curves as a function of the mean of the fourth jet E_T distribution is shown in Fig. C.6. Turn-on middle heights are drawn for inclusive $t\bar{t}$ events after the kinematical selection as well as for $t\bar{t}$ events with at least 4, 5, and 6 jets too.

There is a clear dependence of the E_T corresponding to the turn-on curve middle-height on the hardness of the sample, as shown by the exponential fit to the points. In particular, Fig. C.6 shows that the sample closer to the Tower-10 data, in terms of E_T spectra is the $b\bar{b},\ P_T > 12\ GeV$.

As a consequence of these studies it is reasonable to ask how much the trigger systematic, so far quoted at the 14.8% level, could change using the $b\bar{b}, P_T > 12\ GeV$ sample, instead of the $t\bar{t}$, to quantify the trigger systematic effect.

Fig. C.7 shows the effect of the Tower-10 data turn-on curve application on various Monte Carlo samples. The filled distributions refers to the fourth E_T^{L5} spectra before trigger requirements, except the L1 of the TOP_MULTI_JET; the dotted distributions refer to the E_T distribution obtained from Monte Carlo samples after the application of the requirement of at least four level 2 cluster exceeding 15 GeV and $\sum E_T^{L2} \geq 125\ GeV$. On the other hand, the open histograms represent the Monte Carlo E_T distribution obtained by applying the Tower-10 data trigger turn-on curve.

The trigger systematic values evaluated by using these $b\bar{b}$ and multijet Monte Carlo samples, compared to the one obtained from $t\bar{t}$ events before kinematical selection are reported in Tab. C.5.

Using the $b\bar{b}, P_T > 12\ GeV$, which has the closer E_T spectrum compared to the Tower-10 data, the trigger systematic value is found to be $(0.9 \pm 0.1)\%$.

This shows that by studying trigger turn-on curves in samples with similar kinematical properties, trigger systematic effect could be set at the percent level. Trigger simulation indeed is found to perform well even in high jet multiplicity events, as demonstrated by the value of the trigger systematic (1.7%) found using the six-light-parton Monte Carlo sample.

The studies here presented indicate that large part of the previously quoted trigger systematic (14.8%) is yielded by the different kinematical characteristics of the samples used for the trigger turn-on curves determination, rather than real pathologies concerning the modelling of the trigger behavior in the simulation.

From this point of view we could imagine that extrapolating the Tower-10 data along the dependence shown in Fig. C.6, towards E_T values characteristic of the $t\bar{t}$ production,

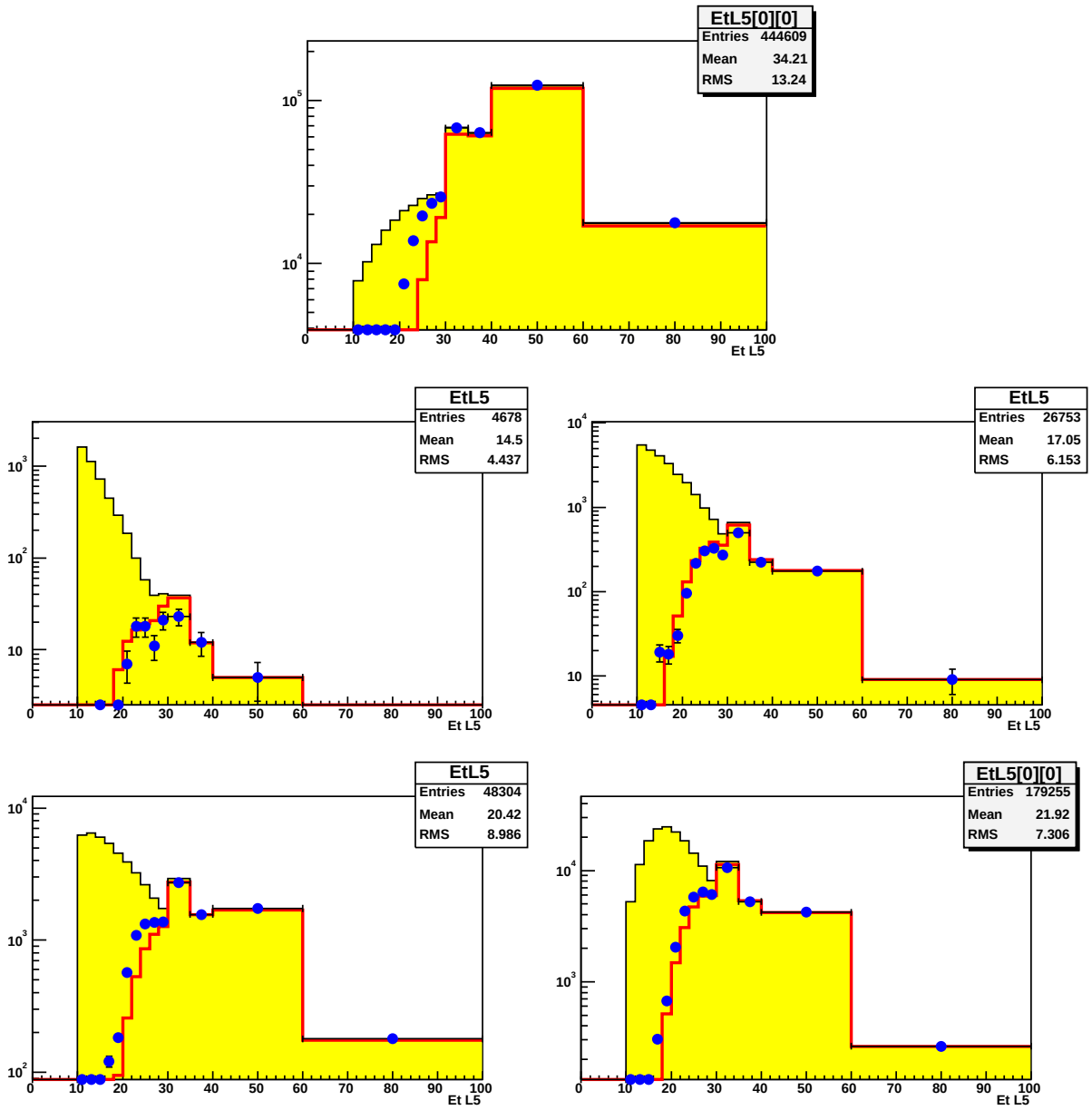


Figure C.7: Fourth jet E_T^{L5} distribution for different Monte Carlo samples (Filled histograms): the top panel correspond to $t\bar{t}$ events, middle left (right) panel refers to $b\bar{b}, P_T > 12\ GeV$. ($b\bar{b}, P_T > 40\ GeV$). Finally, bottom panels refers to $b\bar{b}, P_T > 80\ GeV$ (left) and $6P$ (right) samples. The impact of the TOP_MULTI_JET L2 requirement is shown by the dotted histogram while the impact of the Tower-10 data turn-on curve is shown by the open histogram.

we would find trigger systematic effects limited at the percent level. This extrapolation is partly justified by the low trigger systematic found using multijet Monte Carlo sample, which indicates that the trigger simulation can easily manage high jet multiplicity events. On the other hand, since the trigger simulation is done emulating bit-by-bit the trigger hardware, no surprises are expected to appear as far as the extrapolation of Tower-10

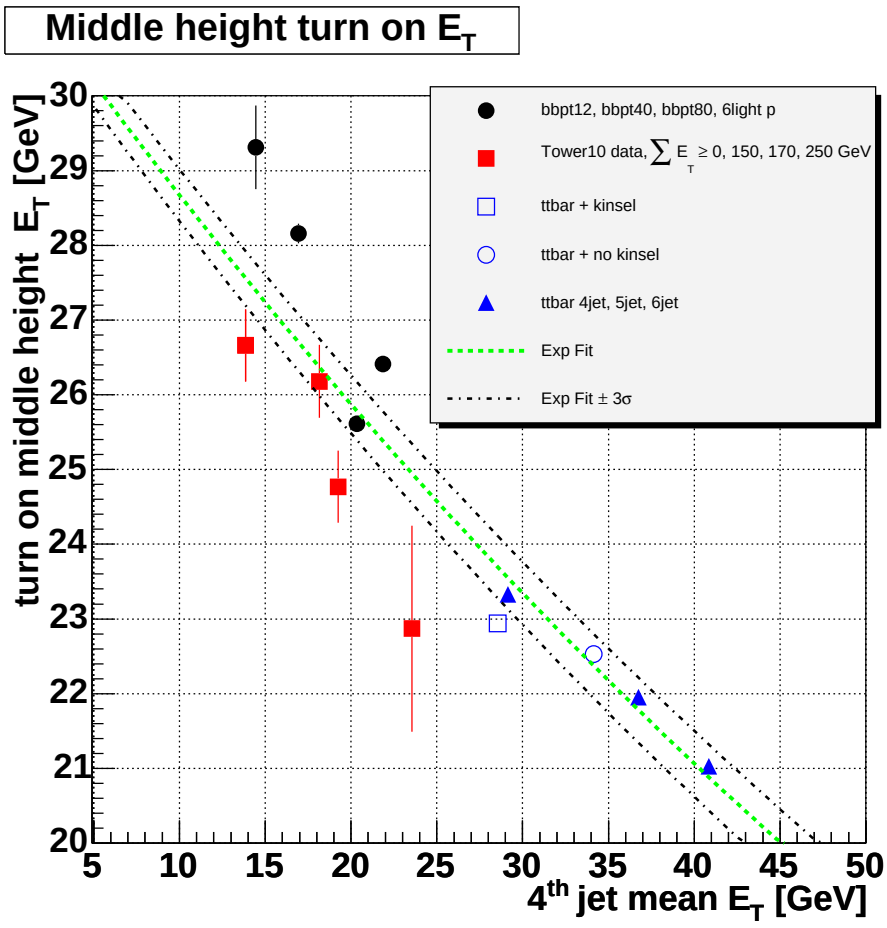


Figure C.8: Turn-on middle height E_T as a function of the $< E_T^{4^{th}} >$ of the sample. For Tower-10 data points $\sum E_T \geq 0,\ 150,\ 170\ 250\ GeV$ is required respectively.

Sample	$t\bar{t}$	$b\bar{b}$, $P_T > 12\ GeV$	$b\bar{b}$, $P_T > 40\ GeV$	$b\bar{b}$, $P_T > 80\ GeV$	$6 - P$
$< E_T^{4^{th}} >$	34.21	14.50	17.05	20.42	21.92
$\Delta\epsilon^{trg}$	10.4 ± 0.2	0.9 ± 0.1	1.3 ± 0.1	3.4 ± 0.1	1.7 ± 0.1

Table C.5: Mean of 4^{th} jet E_T distributions and calculated trigger systematic uncertainties for $t\bar{t}$, $b\bar{b}$ and multijet Monte Carlo samples. E_T and systematic values are in GeV and percent, respectively.

data turn-on curve to events with more energetic jets is concerned.

A reliability check on this assumption has been performed by re-evaluating Tower-10 data turn-on curves applying some $\sum E_T$ cuts. In particular, Fig. C.8 reports the trigger turn-on middle height E_T as a function of the 4^{th} jet mean E_T for Tower 10 data satisfying $\sum E_T \geq 0,\ 150,\ 170,\ 250\ GeV$, respectively. The data are shown to follow the same dependence already discusses as far as Monte Carlo samples is concerned.

Variable	Symbol	Input Value	Output Value
Integrated Luminosity (pb^{-1})	$\mathcal{L}$	311 ± 18	311 ± 19
Observed Tags	N_{obs}	127	–
Expected Tags	N'_{exp}	57.4 ± 8.1	57.3 ± 7.7
Kin. efficiency (%)	ϵ_{kin}	4.878 ± 0.431	4.888 ± 0.503
Ave. b-tags per $t\bar{t}$ event	ϵ_{tag}^{ave}	0.7889 ± 0.0458	0.7891 ± 0.0486

Table C.6: Input and output values of the likelihood maximization.

C.2.1 Cross section results with reduced trigger systematics

The evaluation of the trigger acceptance systematic as the sum in quadrature of the systematic values found using $b\bar{b}, P_T > 12\ GeV$ and $6P$ samples yields 2.0%. With this value, we can re-evaluate the cross section measurement uncertainties.

Following the same approach described in Sec. 7.2 we can modify the input values for the kinematical efficiency systematic on $t\bar{t}$ accordingly to the newer estimate of the trigger systematics.

Doing so we end up with the following results. The input and output parameter of the likelihood maximization are quoted in Tab.C.6. The measured cross section value is found to be:

$$\sigma_{t\bar{t}} = 5.8 \pm 1.2(stat)\ ^{+0.9}_{-0.7}(syst)\ pb \quad \text{(C.2)}$$

$$= 5.8\ ^{+1.5}_{-1.4}\ pb, \quad \text{(C.3)}$$

to be compared with the previous quoted cross section value of:

$$\sigma_{t\bar{t}} = 5.9 \pm 1.2(stat)\ ^{+1.4}_{-1.0}(syst)\ pb \quad \text{(C.4)}$$

$$= 5.9\ ^{+1.8}_{-1.6}\ pb. \quad \text{(C.5)}$$

The effect of the adoption of the reduced trigger systematic values decreases the overall uncertainty of the measurement from 28% to 25%.

Cross section dependence on M_{top}

As in Sec. 7.2, additional samples of inclusive $t\bar{t}$ Monte Carlo events generated with different M_{top} values in the range $[130, 230]$ GeV/c^2 are used to compute the cross section measurement dependence on M_{top}; as shown in Figure C.9, the cross section changes by ± 0.05 pb for each ∓ 1 GeV/c^2 change in the top mass from the initial value of 178 GeV/c^2. For instance, we measure $\sigma_{t\bar{t}} = 6.0 \pm 1.2(stat.)^{+0.9}_{-0.7}(syst.)$ pb for $M_{top} = 175$ GeV/c^2. The change is due to the varying signal selection efficiency with top quark mass.

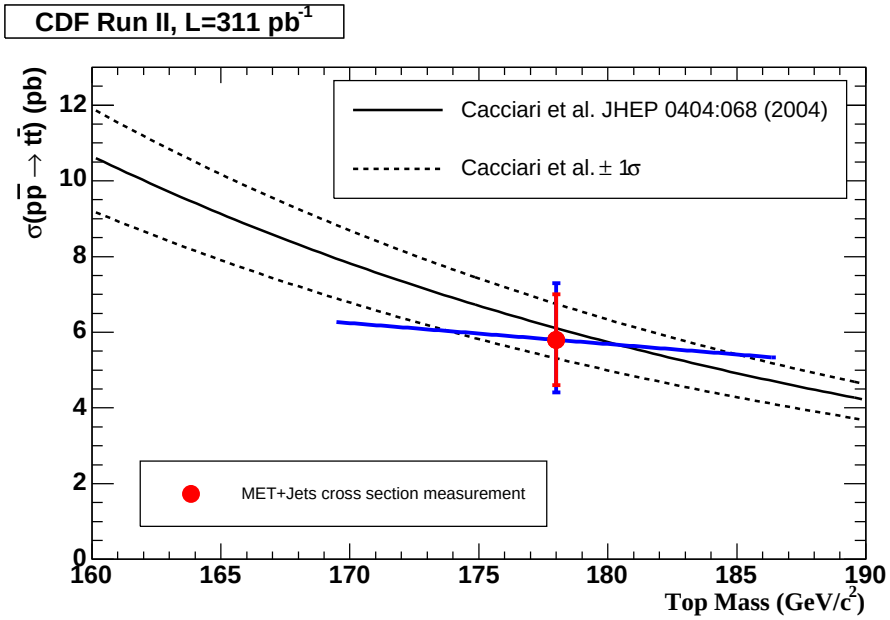


Figure C.9: Cross section measurement dependence on M_{top}.

Acknowledgments

The first persons I would like to thank are Giovanni, my supervisor, Tommaso, who followed me since the beginning of my work within CDF, and Dario, for the opportunity they gave me to do my Ph.D. in such an exciting particle physics experiment.

I would then thank all the members of the CDF Padova group; in particular Tommaso and Luca who helped me in moving my first steps in the tangle of the CDF software and who supported and motivated me during the first stage of this work.

I would also thank, Nicola and Patrizia for their friendly presence during my stays at Fermilab. A particular thanks to Julien, for his friendship and support, and to Donatella for her useful comments and suggestions.

A special mention to Saverio with whom I shared many moments either at the Department of Physics of Padova or at Fermilab in the attempt of completing our Laurea and Ph.D. theses; to Delli, who accompanied me during the ever increasing "smoking-breaks", for his friendship; and finally, to Matteo and Roberto for their useful help.

I am grateful to my parents, to Giovanna and Anna for their ever lasting support. A special thanks to my little brother Giacomo, for the beautiful moments he always grants to me. This thesis is dedicated to them.

A final consideration and thank to all my friends who never have forgotten me during my long stays far away.

List of Figures

List of Tables

Bibliography

[1] F. Mandll and G. Shaw, *Quantun Field Theory*, Revised Edition, John Wiley & Sons, (1999).

[2] M. Herrero, "The standard model", arXiv:hep-ph/9812242, (1998).

[3] F. Abe *et al.* [CDF Collaboration], "Evidence for top quark production in anti-p p collisions at s**(1/2) = 1.8-TeV", Phys. Rev. Lett. **73**, (1994) 225 [arXiv:hep-ex/9405005].

[4] S. Abachi *et al.* [DØ Collaboration], "Observation of the top quark", Phys. Rev. Lett. **74** (1995) 2632 [arXiv:hep-ex/9503003].

[5] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, "Evidence For The 2 Pi Decay Of The K(2)0 Meson", Phys. Rev. Lett. **13** (1964) 138.

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

[7] J. J. Aubert *et al.* [E598 Collaboration], "Experimental Observation Of A Heavy Particle J", Phys. Rev. Lett. **33** (1974) 1404.

[8] J. E. Augustin *et al.* [SLAC-SP-017 Collaboration], "Discovery Of A Narrow Resonance In E+ E- Annihilation", Phys. Rev. Lett. **33** (1974) 1406.

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

[10] 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.

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

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

[13] LEP collaborations. A combination of Preliminary LEP Electroweak Measurement and contraints on the Standard Model. $CERN - PPE$, 95-172, 1995.

[14] G. Arnison *et al.* [UA1 Collaboration], "Experimental Observation Of Isolated Large Transverse Energy Electrons With Associated Missing Energy At S**(1/2) = 540-GeV", Phys. Lett. B **122** (1983) 103.

[15] G. Arnison *et al.* [UA1 Collaboration], "Experimental Observation Of Lepton Pairs Of Invariant Mass Around 95-GeV/C**2 At The Cern Sps Collider", Phys. Lett. B **126** (1983) 398.

[16] G. L. Kane and M. E. Peskin, "A Constraint From B Decay On Models With No T Quark", Nucl. Phys. B **195** (1982) 29.

[17] A. Bean *et al.* [CLEO Collaboration], "Improved Upper Limit On Flavor Changing Neutral Current Decays Of The B Quark", Phys. Rev. D **35** (1987) 3533.

[18] S. Eidelman *et al.* [Particle Data Group], "Review of particle physics", Phys. Lett. B **592** (2004) 1.

[19] M. Carena *et al.* [Higgs Working Group Collaboration], "Report of the Tevatron Higgs working group", arXiv:hep-ph/0010338.

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

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

[22] CDF and DØ collaborations, "Combination of CDF and DØ results on the top-quark mass", arXiv:hep-ex/0507091, (2005).

[23] P. Azzi *et al.* [CDF Collaborattion], "Combination of CDF and D0 results on the top-quark mass", arXiv:hep-ex/0404010.

[24] t. L. E. Group, t. S. electroweak and h. f. groups [the SLD Collaboration], "Precision electroweak measurements on the Z resonance", arXiv:hep-ex/0509008, (2005).

[25] LEP EW WG Plots for the Summer 2005 Conferences, http://www.cern.ch/LEPEWWG.

[26] I. I. Y. Bigi, Y. L. Dokshitzer, V. A. Khoze, J. H. Kuhn and P. M. Zerwas, "Production And Decay Properties Of Ultraheavy Quarks", Phys. Lett. B **181** (1986) 157.

[27] L. H. Orr, "Decay Versus Hadronization For Top Quarks Produced In Hadron Colliders", Phys. Rev. D **44** (1991) 88.

[28] K. Koba *et al.*, Proceeding of the IEEE Particle Accelerator Conference 2003, p1736.

[29] Fermilab Beams Division, http://www-ad.fnal.gov/runII/index.html.

[30] Fermilab Beams Division, "Fermilab Linac Upgrade. Conceptual Design", FERMILAB-LU-Conceptual Design, http://www-lib.fnal.gov/archive/linac/fermilab-lu-conceptualdesign.shtml.

[31] Fermilab Beams Division, "The Linac rookie book", http://www-bdnew.fnal.gov/operations/rookie_books/Linac_PDF/Linac.html

[32] Fermilab Beams Division, "The Booster rookie book", http://www-bdnew.fnal.gov/operations/rookie_books/Booster_PDF/Booster_TOC.html.

[33] Fermilab Beams Division, http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html.

[34] Fermilab Beams Division, "The antiproton source rookie book", http://www-bdnew.fnal.gov/pbar/documents/PBAR_Rookie_Book.pdf.

[35] Fermilab Beams Division, "The Recycler ring", http://www-recycler.fnal.gov/recycler_main.html.

[36] G. Jackson, "Operational experience with the permanent magnet Recycler ring at Fermilab," *Prepared for 7th European Particle Accelerator Conference (EPAC 2000), Vienna, Austria, 26-30 Jun 2000.*

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

[38] F. Abe *et al.*(The CDF II Collaboration), "The CDF II Technical Design Report", FERMILAB-PUB-96-390-E, Oct 1996.

[39] P. C. Bhat and W. J. Spalding, "Fermilab collider Run II: Accelerator status and upgrades", AIP Conf. Proc. **753** (2005) 30 [arXiv:hep-ex/0410046].

[40] H. Wenzel, "Tracking in the SVX", CDF Internal Note 1790, (1992).

[41] 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). Update to Proposal P-909: "Physics performance of CDF II with an Inner Silicon Layer and a Time of Flight detector", http://www-cdf.fnal.gov/upgrades/btb-update_jan99.ps,(1999).

[42] J. Marriner, "Secondary vertex fit with mass and pointing constrains", CDF Internal Note 1996, (1993).

[43] W.-M. Yao, and K. Bloom, "Outside-In silicon Tracking at CDF", CDF Internal Note 5991, (2002).

[44] P.Azzi, G. Busetto, A. Ribon, "Histogram tracking in the COT", CDF Internal Note 5562, (2001).

[45] , C. Hays *et al.*, "Inside-Out tracking", CDF Internal Note 6707, (2003).

[46] , H. Studie *et al.*, "Vxprim in Run II", CDF Internal Note 6047, (2002).

[47] J. E. Huth *et al.*, Proceedings of the 1990 DPF Summer Study on High Energy Physics, Snowmass, Colorado, edited by E.T. Berger, World Scientific, Singapore, 1992.

[48] Jet Energy and Resolution WG, "Generic Jet Scale Correction for Run II, CDF Internal Note 7358, (2005).

[49] A. Bhatti *et al.*, "Determination of the jet energy scale at the Collider Detector at Fermilab", arXiv:hep-ex/0510047, (2005).

[50] D. Tsybychev *et al.*, "A study of missing E_T in Run II minimum bias data", CDF Internal Note 6112, (2002).

[51] D. Glenzinsky *et al.*, "A detailed study of the SECVTX algorithm", CDF internal note 2925, (1994).

[52] C. Cmiao *et al.*, "Update of SECVTX tagging on Run 1B data", CDF internal note 3178, (1995).

[53] T. Affolder *et al.*, Phys. Rev. D **64**, 032002 (2001).

[54] J. Guimares *et al.*, "SecVtx Optimization Studies for 5.3.3 Analyses", CDF internal Note 7578, (2005).

[55] M. Paulini, "B Lifetimes, Mixing and CP Violation at CDF", Int. J. Mod. Phys. **A14**, 2791 (1999).

[56] S. Belforte *et al.*, "SVT – Silicon Vertex Tracker – Technical Design Report", CDF internal Note 3108, (1996).

[57] J. Troconiz *et al.*, "Update in the search of a Neutral Heavy Scalar in the Multijet Data sample", CDF internal Note 4129, (1998).

[58] A. Cerri, R. Miquel, "Studying the SVT Efficiency and Resolution with J/Psi Data", CDF Note 5838, (2002); and B. Ashmanskas *et al.*, "Performance of the CDF Online Silicon Vertex Tracker", CDF Note 5854, (2002), Published Proc. 2001 IEEE Nuclear Science Symposium (NSS) and Medical Imaging Conference (MIC), San Diego, CA, 4-10 November 2001. FERMILAB-CONF-02/035-E.

[59] K. Anikeev *et al.*, "Event Builder and Level-3 for aces", CDF Internal Note 5793, (2001).

[60] S.Klimeno *et al.*, "The CDF Run II luminosity monitor", CDF Internal Note 4330, (1998).
S. Klimeno *et al.*, "CLC production and mechanical assembly", CDF Internal Note 5119, (1999).
D. Acosta *et al.*, "A first look at the CLC luminosity measurements", CDF Internal Note 6052, (2002).

[61] 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).

[62] T.Affolder *et al.*, "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8\ TeV$", Phys.Rev D**64**, 032002 (2001)

[63] T.Affolder *et al.*, "Measurement of the top quark mass with the collider detector at Fermilab", Phys.Rev D**63**, 032003 (2001)

[64] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, Computer Physics Commun. 135 (2001) 238.

[65] G. Marchesini, B.R. Webber, G.Abbiendi, I.G.Knowles, M.H.Seymour and L.Stanco Computer Physics Communications 67 (1992) 465; G.Corcella, I.G.Knowles, G.Marchesini, S.Moretti, K.Odagiri, P.Richardson, M.H.Seymour and B.R.Webber, JHEP 0101 (2001) 010.

[66] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky and W. K. Tung, "New eneration of parton distributions with uncertainties from global QCD analysis", JHEP **0207** (2002) 012.

[67] J.Lewis, P.Avery, "CLEOMC: The CDF Interface to the CLEO Monte Carlo (QQ)", CDF Internal Note 2724, (1994).

[68] R. Brun, R. Hagelberg, M. Hansroul and J. C. Lassalle, "Geant: Simulation Program For Particle Physics Experiments. User Guide And Reference Manual", CERN-DD-78-2-REV.

[69] S. Carron, M. Kruse, M. Gold, "Parametric Model fot the Charge Deposition of the CDF RunII Silicon Detectors", CDF Internal Note 7598, (2005).

[70] W. Riegler *et al.*, "COT detector physics simulations", CDF Internal Note 5050, (1999).

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

[72] M. Mangano *et al.*, "ALPGEN, a generator for hard multiparton processes in hadronic collisions", hep-ph/0206293, (2002).

[73] M. Acciarri *et al.* [L3 Collaboration], "Measurement of the branching ratios b $\rightarrow$ e nu X, mu nu X, tau nu X and nu X", Z. Phys. C **71** (1996) 379.

[74] T. Dorigo, L. Scodellaro, "b-Specific Jet Corrections for Higgs Boson Mass Reconstruction", CDF Internal Note 6647, (2003); and T. Dorigo, L. Scodellaro, "Narrowing the Higgs mass peak: the hyperball algorithm", CDF Internal Note 6450, (2004).

[75] P. Lujan *et al.*, "Efficiency of SecVtx B-Tag used for 2005 Winter Conferences", CDF Internal Note 7343, (2005).

[76] A. Holloway *et al.*, "SecVtx tagging scale factor for 5.3.3 analyses", CDF Internal Note 7445, (2005).

[77] T. Wright, "SecVtx B-tag efficiency using Muon transverse momentum", CDF Internal Note 7448, (2005).

[78] H. Bachacou *et al.*, "Combinig the SecVtx B-tagging scale Factors for 5.3.3 analyses", CDF Internal Note 7480, (2005).

[79] V. D. Barger and R. J. N. Phillips, "Collider Physics", Addison-Wesley (1987).

[80] http://www-cdf.fnal.gov/internal/dqm/goodrun/v7/goodv7.html.

[81] C. Hill *et al.*, "Electron Identification in Offline Release 5.3", CDF Internal Note 7309, (2005).

[82] V. Martin, "High-Pt Muon ID Cuts and Efficiencies for use with 5.3.1 Data and 5.3.3 MC", CDF Internal Note 7367, (2005).

[83] K. Copic, M. Tecchio, "Event vertex studies for dilepton events", CDF Internal Note 6933, (2004).

[84] E. Thomson *et al.*, "Top and ElectroWeak Event Classification Module for CDF Run II", CDF Internal Note 5947, (2003).

[85] G. Cortiana *et al.*, "$t\bar{t} \to \tau + jets$ nt5 analysis update", CDF Internal Note 7553, (2005).

[86] G. Cortiana, "Top Multijet trigger efficiency study for the all-hadronic and tau+jets t tbar decays", CDF Internal Note 7473, (2005).

[87] G. Cortiana *et al.*, "Strategy for an event selection reoptimization aimed at the $t\bar{t} \to \tau + jets$ search", CDF Internal Note 7382, (2004).

[88] G. Cortiana *et al.*, "Background studies for the $t\bar{t} \to \tau + jets$ search", CDF Internal Note 7292, (2004).

[89] G. Cortiana *et al.*, "A Preliminary event selection for the $t\bar{t} \to \tau + jets$ search", CDF Internal Note 7291, (2004).

[90] G. Cortiana *et al.*, "Validation and retuning of τ ID cuts for the $t\bar{t} \to \tau + jets$ search", CDF Internal Note 6847, (2004).

[91] O. Gonzalez *et al.*, "Uncertainties due to PDFs for the gluino-sbottom search", CDF Internal Note 7051, (2004).

[92] G. Cortiana *et al.*, "$t\bar{t}$ cross section measurement in the Run II multijet dataset", CDF Internal Note 6914, (2004).

[93] S. Klimenko, J. Konigsberg, and T.M. Liss, FERMILAB-FN-0741 (2003).

[94] CDF collaboration, "Combination of CDF top pair production cross section measurements", CDF Note 7794, (2005).

[95] L.Lyons, D.Gibaut and P.Clifford, NIM A270, 110-117 (1988); L.Lyons, A.Martin and D.Saxon, Phys. Rev. D 41, 3 (1990); and A.Valassi, NIM A500, 391 (2003).

[96] A. Anastassov *et al.*, "Tau Reconstruction Efficiency and QCD Fake Rate for Run2", CDF Internal Note 6308, (2003).

[97] P. Murat, "Measurement of the $\sigma(p\bar{p} \rightarrow W) \cdot BR(W \rightarrow \tau\nu)$ and $BR(W \rightarrow \tau\nu)/BR(W \rightarrow e\nu)$ with Run II data", CDF Internal Note 6010, (2003).