UNIVERSITÀ DI PISA

Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Specialistica in Scienze Fisiche

TESI DI LAUREA SPECIALISTICA

A New Top Mass Measurement in The Dilepton Channel

Candidato:

Marco Trovato

Relatori:

Prof. Giorgio Bellettini
Dr. Gueorgui V. Velev

Anno Accademico 2006-2007

DEDICATION

To my brother and his will-power

Contents

Introduction

"*We establish the existence of the top quark using a 67 pb^{-1} sample of $\overline{p}p$ collisions at $\sqrt{s}$ = 1.8 TeV collected with the Collider Detector at Fermilab*". This is a quotation from [11] in which the discovery of top quark was announced in 1995.

The top quark discovery completed the present picture of the fundamental constituents of the nature. Since then, the Collider Detector at Fermilab (CDF) and $D\emptyset$ Collaborations have been spending great efforts to measure its properties better. About 30 times larger than the second heaviest quark, the mass of the top has been measured with increased statistic and more and more sophisticated techniques in order to reduce as much as possible its uncertainty. This is because the top is expected to play a fundamental role in the Standard Model. The value of its mass sets boundaries on the mass of the unobserved Higgs boson, and perhaps more appealing, studies of its properties might lead to the discovery of new physics.

In the work described in this thesis the top mass is measured via top pair production and decay in the dilepton mode $t\overline{t} \rightarrow (W^+b) + (W^-\overline{b}) \rightarrow (l^+\nu b) + (l^-\overline{\nu}\overline{b})$. Dilepton events are selected by the "LTRK" signature, which is optimized for a larger statistic while preserving a fair signal to background ratio (Chapter 4).
The top mass is estimated using the so-called Template Method (Chapters 5 to 10). Event-by-event the azimuthal angles of the neutrinos are arbitrarily set, all associations of the final state observables to the decay particles of the $t\overline{t}$ system are tried and the top mass is obtained from the association which agrees best with a $t\overline{t}$ dilepton event. The procedure is iterated by varying the assumed azimuthal neutrino angles and a mass is obtained for each iteration. A suitably weighted average of the obtained masses gives the mass associated to an event (Chapter 6). Once the masses of the 236 candidate events are obtained their distribution is fitted in terms of simulated distributions ("templates") for signal and background events. The fit is performed using a likelihood distribution function. The background template is obtained adding the contributions by the competing channels

weighted according to their cross section, while mass-dependent signal templates are obtained by simulated $t\overline{t}$ events (Chapters 5, 7). The result of the two component fit indicates the best mass template, the top mass and its statistical error.

Systematic uncertainties are added to the statistical error to account for uncertainties in the parameters used in the simulations and for defects of the analysis process (Chapter 9).

The method ("Neutrino ϕ Weighting Method") was already successfully applied on a 190 and 340 pb^{-1} integrated luminosities and the result was published [45].

The measurement described in this thesis has made use of a refined Neutrino ϕ Weighting Method both in event selection (Chapter 4), and in the way to reconstruct events for a better accuracy on top mass. Furthermore the method was applied to a sample of a much larger statistic (2.1 fb^{-1} integrated luminosity).

The Laurea Candidate worked at Fermilab for over nine months in close collaboration with Dr. Igor Suslov, who had already played a key role in the previous analysis [45], and under the supervision of Dr. Gueorgui Velev and the advise by Prof. Giorgio Bellettini. It was a full time and massive effort, in which the personal contributions of the Candidate have been primarily in the selection of lepton+track events, estimation of backgrounds, particularly those relative to mismeasured events ("Fakes") which are the main background (Chapter 4), and evaluation of a number of systematic uncertainties.

Chapter 1

Standard Model and Top Quark Physics

In this chapter a brief overview of particle physic theory will be provided. The first section will deal with the Standard Model and the second one with top quark physics.

1.1 The Standard Model

Present understanding of fundamental constituents of matter and the fundamental interactions are provided by a theory called **Standard Model (SM)**.

The SM has been developed in the 1960's and 70's and has been experimentally tested in these last 50 years.

The SM incorporates two gauge theories: the theory of Quantum Chromodinamics (QCD) [1, 2], which describes the strong interactions, and the Glashow-Salam-Weinberg (GSW) theory of electroweak interactions [3, 4, 5], which unifies the weak and electromagnetic interactions. The problem to incorporate gravity in the theory is still open.

The SM describes two families of fundamental particles [1]: *fermions* and *bosons*, which have different spins. Bosons and fermions have integer and half-integer spin respectively.

The fermion families contain six quarks and six leptons arranged in three generations of weak isospin doublets. Four bosons are the interaction carriers [2]. Figure 1.1 gives a picture of the particles in SM.

[1] A particle is considered fundamental if no internal structure is revealed

[2] Each of these particles has its own antiparticle

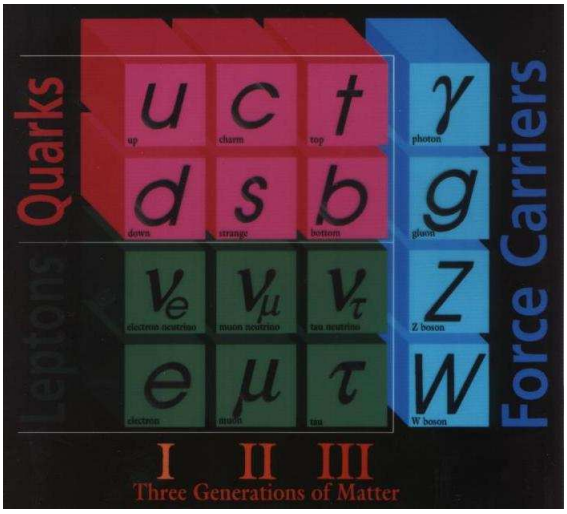

Figure 1.1: The three fermion generations and the four bosons of the Standard Model.

1.1.1 Fundamental Interactions

SM is a gauge theory. In a gauge theory the fundamental particles are described by quantized fields ψ and their interactions are expressed by the Lagrangian density $\mathcal{L}$. Noether's theorem states that an invariance of $\mathcal{L}$ under a certain transformation of fields will hide a symmetry for the system described by the Lagrangian density.

The Electroweak Theory

Electromagnetic and weak interactions theory have been unified in the electroweak theory.

The electromagnetic interaction (QED) is described by the following Lagrangian density:

$$L_{QED} = \overline{\psi}(i\partial\!\!\!/ - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + QeA^{\mu}\overline{\psi}\gamma_{\mu}\psi \tag{1.1}$$

where

- $\overline{\psi}(i\partial\!\!\!/ - m)\psi$ is the Lagrangian density of a free fermion of mass m described by a Dirac four-component spinor ψ.
- $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ describes the propagation of a free photon field and expresses the Maxwell equations in a covariant form.

- $QeA^{\mu}\overline{\psi}\gamma_{\mu}\psi$ describes the interaction between a fermion of charge Q in units of the electron charge [3] and the photon field. By analogy with classic mechanics a QED current can be defined:

$$J_{\mu}^{QED} \equiv Qe\overline{\psi}\gamma_{\mu}\psi \tag{1.2}$$

While the second and third term have a direct analogy to classical physics, the first term is pure quantum field theory.

QED has U(1) gauge Abelian local symmetry since a local change in the phase definition for the fields does not affect the electromagnetic Lagrangian density (see 1.1). The symmetry group for QED is defined as $U(1)_Q$, since conservation of the charge Q comes from local U(1) invariance.

Weak interaction phenomena show common properties between doublets of fermions ($\begin{pmatrix} e \\ \nu_e \end{pmatrix}$, $\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}$, ... see figure 1.1). Because of this reason a weak isospin is defined and weak interaction Lagrangian density is built in order to be invariant under 2-D rotation in the weak isospin space (SU(2)). Fermions are therefore grouped into doublets and singlets as follows:

$$\chi_L = \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix}_L \tag{1.3}$$

$$\chi_R = \psi_R \tag{1.4}$$

where "L" and "R" mean left-handed and right handed respectively [4]. "u" and "d" are the labels up and down fermions (according to figure 1.1).

Analogously to QED (see 1.2), three weak currents can be defined:

$$J_{\mu}^{weak,i} = \overline{\chi}\gamma_{\mu}\frac{\sigma^i}{2}\chi \quad i = 1, \ldots, 3 \tag{1.5}$$

where σ^i are the Pauli matrices, J_3 is the neutral current, J_1 and J_2 are linear combination of the charged ones.

Neutral weak current is experienced by both left and right handed fermions. Since neutrino are massless [5], charged currents are experienced only by left-handed fermions [6].

[3] e is the charge of the electron. $e = 1.60217646 \times 10^{-19}$ Coulomb

[4] a particle is left/right-handed if the direction of its spin is opposite/same as the direction of its motion.

[5] In the SM neutrino are considered massless. Their very small mass does not change our picture.

[6] This statement is true for leptons since no right-handed neutrinos have been seen yet. It is considered valid for quarks as a natural extension in the electroweak unification

Electroweak theory incorporates flavor changing charged currents and the absence of such neutral currents [7]: ${\psi^{quark}}_d \equiv (d, s, b) \rightarrow {\psi'^{quark}}_d \equiv (d', s', b')$ where (d', s', b') is expressed according to the following:

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

where V_{CKM} is the Cabibbo-Kobayashi-Maskawa quark mixing matrix.

According to the proposal presented by Glashow in 1961 [3, 4, 5], these two theories, QED and weak, have been unified in the EW theory, represented by the group SU(2)xU(1).

Since right-handed fermions experience QED and are singlets in a weak-isospin space it is convenient to define the hypercharge Y and hypercharge current J_μ^Y:

$$Y = 2(Q - t_3) \tag{1.6}$$

$$J_\mu^Y = \overline{\psi}\gamma_\mu Y \psi \tag{1.7}$$

where t_3 is the third component of the weak isospin. 1.7 is the Gell-Mann-Nishijima formula.

According to formulas 1.1,1.5,1.7 the EW Lagrangian density is written in terms of the weak fields B_μ and $W_\mu \equiv (W_\mu^1, W_\mu^2, W_\mu^3)$:

$$\begin{aligned} L_{EW} &= \overline{\psi}(i\partial\!\!\!/ - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}W^{+\mu\nu}W_{-\mu\nu} - \frac{1}{4}Z^{\mu\nu}Z_{\mu\nu} + g{\mathbf{J}^{\mathbf{weak}}}_\mu \dot{\mathbf{W}}^\mu + \\ &\quad \frac{g'}{2}J_\mu^Y B^\mu \end{aligned} \tag{1.8}$$

where

- W_μ^1 and W_μ^2 are linear combinations of $W_\mu^\pm$ (fields carried by $W^\pm$ bosons), while B_μ and W_μ^3 can be expressed in term of Z_μ^0 and A_μ (fields carried by Z^0 and γ bosons) [8].
- $\frac{1}{2}W^{\pm\mu\nu}W^{\pm\mu\nu}$, $\frac{1}{4}Z^{\mu\nu}Z_{\mu\nu}$ describe the propagation of W and Z^0 gauge boson fields. $W^{\pm\mu\nu}$ and $Z_{\mu\nu}$ are defined by analogy with QED case.

[7] The first evidence came from strangeness changing charged current. An example of this phenomenon is $\Lambda \rightarrow pe\overline{\nu_e}$

[8] $W^\pm$ and Z^0 are respectively the charged and neutral carriers of weak interaction (see 1.1)

- g, g' are defined according to:

$$e = gsin(\theta_W) = g'cos(\theta_W) \tag{1.9}$$

  θ_W being the Weinberg angle.

This Lagrangian density is invariant under transformations in the $SU(2)_L xU(1)_Y$ group [9].

However, this symmetry would be exact if fermions, W and Z bosons were massless. In order for the standard model to be compatible with the large observed masses of W and Z bosons [10], symmetry breaking must occur. This EW symmetry breaking (EWSB) would presumably also be responsible for the mass hierarchy observed for the fermions.

Spontaneous EWSB can be accomplished by the introduction of a scalar field, interaction term in the Lagrangian known as Higgs Field [7]. The existence of a massive boson, the Higgs boson, would be associated with the Higgs field. At present the existence of the Higgs boson has yet to be confirmed experimentally. Indirect experimental bounds for the SM Higgs boson mass are obtained from fits to precision measurements of the measured top and W masses.

The current (winter 2007) best fit is $m_H = 91 \pm^{45}_{32} GeV$ or $m_H < 186\ GeV$ at the 95% confidence level [8].

Quantum-Chromo-Dynamics

The last piece of the SM is QCD describing the strong force. This force is responsible for quarks "sticking" together to form composite particles (hadrons).

QCD is described by a non-abelian local SU(3) group. As for QED and weak interactions, gauge invariance of this group implies a new degree of freedom to be conserved: the *color*. However, unlike QED, the gauge symmetry group is non-abelian, causing gluons (carriers of strong interaction) to possess color charge and interact with each other.

A consequence of the interactions among gluons is that the coupling "constant" α_s depends on Q^2 (the interaction momentum transfer scale) [11]. To a first approximation in Q^2/λ^2 one has:

$$\alpha_s(Q^2) \ = \ \frac{12\pi}{(33-2f)ln(Q^2/\lambda^2)} \tag{1.10}$$

[9] $SU(2)_L$ acts only on left-handed fermions while $U(1)_Y$ acts only on right-handed one, being Y conserved in processes among right-handed fermions.

[10] $M_W = 80.41 \pm 0.05\ GeV/c^2$ and $M_Z = 91.188 \pm 0.002\ GeV/c^2$

[11] α_{QED}, the QED coupling constant, does not depend on Q^2 since QED is an abelian symmetry group.

where f is the number of different flavoured quarks with mass less than Q^2. Fits to experimental data set λ around 200 MeV [6].

QCD features two very important properties:

1. **Asymptotic freedom**. $\alpha_s(Q^2)$ becomes small at large Q^2 (see 1.10). This means that quarks and gluons interact weakly if they are within a short range. This property allows perturbation theory to be used in theoretical calculations to produce experimentally verifiable predictions for hard scattering processes.

2. **Confinement**. Colored particles are confined into colorless singlets (hadrons) by an increasing QCD potential with increasing relative distance. If quarks are forced to large relative distances the energy density in the binding color string increases and energy is materialized into colored quark pairs. Thus a hard scattered parton evolves into a shower of partons and finally into hadrons (*hadronization*).

   Even if analytically unproven, confinement is widely believed to be true because it explains the consistent failure of free quark searches.

1.1.2 Fundamental particles

Figure 1.1 shows the fundamental particles divided into three categories: gauge bosons, quarks and leptons. Tables 1.1, 1.2 summarize their most important quantum numbers. Each particle has a corresponding antiparticle with opposite electric charge [12].

Bosons	Mass (approx)	Q	force carried	force experienced
photon (γ)	$< 6 \cdot 10^{-26}$ GeV	0	EM	Grav.
$W^{\pm}$	80.4 GeV	± 1	Weak	Grav., EM, Weak
Z^0	91.2 GeV	0	Weak	Grav., Weak
gluon (g)	0 GeV	0	Strong	Grav, Strong
graviton	not yet observed.	0	Grav.	Grav.

Table 1.1: Bosons: Mass, electric charge and the force they carry and experience.

[12] Some particles can be their own antiparticles. One examples is the photon

Fermions	Mass (approx)	Q	T^3	Y/2
leptons				
electron (e)	511 KeV	-1	-1/2	-1/2
electron neutrino (ν_e)	< 2 eV	0	+1/2	-1/2
muon (μ)	106 MeV	-1	-1/2	-1/2
muon neutrino (ν_μ)	< 2 eV	0	+1/2	-1/2
tau (τ)	1.78 GeV	-1	-1/2	-1/2
tau neutrino (ν_τ)	< 2 eV	0	+1/2	-1/2
quarks				
u	1.5÷3 MeV	+2/3	+1/2	+1/6
d	3÷7 MeV	-1/3	-1/2	+1/6
c	$\sim$ 1.25 GeV	+2/3	+1/2	+1/6
s	$\sim$ 95 MeV	-1/3	-1/2	+1/6
t	170÷175 GeV	+2/3	+1/2	+1/6
b	4.2÷4.7 GeV	-1/3	-1/2	+1/6

Table 1.2: Leptons and quarks: mass, electric charge, third component of weak isospin, hypercharge.

1.2 The Top Quark

As consequence of the discovery of bottom (b) quark in 1977 the existence of its weak isospin doublet partner, the top (t) quark, has been inferred because of several reasons.

1. First of all the renormalizability of the standard model requires that the sum of electric charges of all left-handed fermions is zero.

2. The measured $\Gamma(Z^0 -> b\overline{b})$ is $384 \pm 4\ MeV$, very close to the value of 381 MeV expected if $T_3^b = -1/2$

The discovery of top quark was accomplished in 1995 at the CDF and DØ experiment [10, 11]. After more than 10 years from its discovery the world average for top quark mass is $M_{top} = 170.9 \pm 1.1(stat) \pm 1.5(syst) GeV/c^2$ [58] (more than 30 times the mass of second heaviest quark). Figure 1.2 summarizes a number of top mass measurements at CDF.

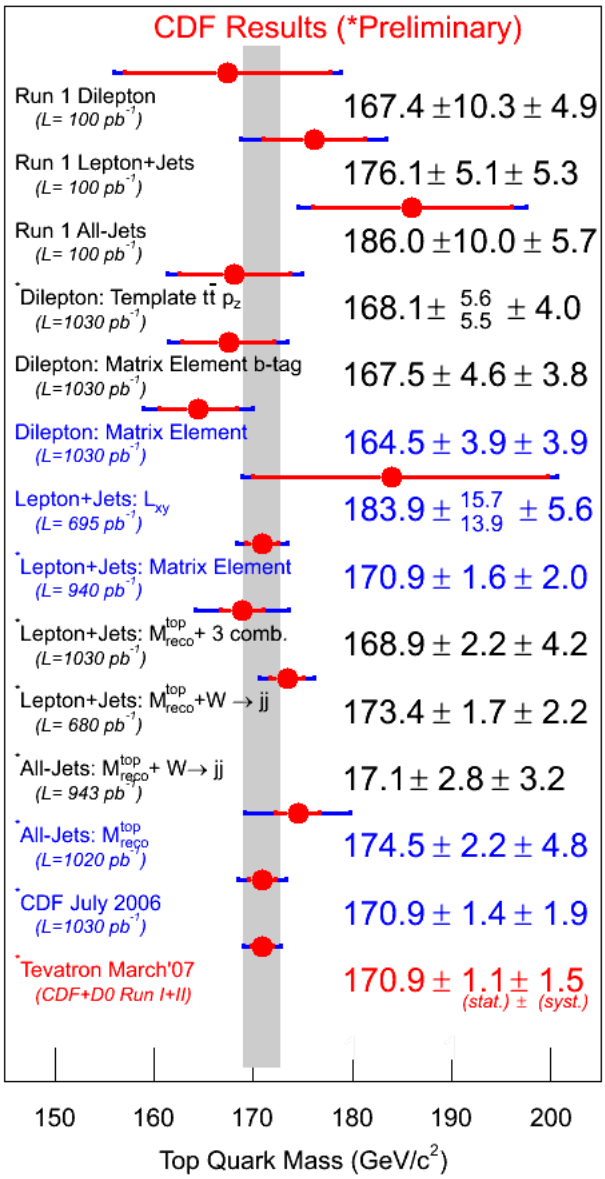

Figure 1.2: Top quark mass measurements in Run I and Run II at CDF (summer 2007).

1.2.1 Top Quark production and $t\bar{t}$ cross section

At CDF top quarks are produced predominantly in pairs via $q\bar{q}$ collisions [13]. Even if quarks and gluons are bound within hadrons, at large momentum transfer they can be considered free (asymptotic freedom). An example of such a process is shown in figure 1.3. Each parton carries a fraction x of the hadron total momentum.

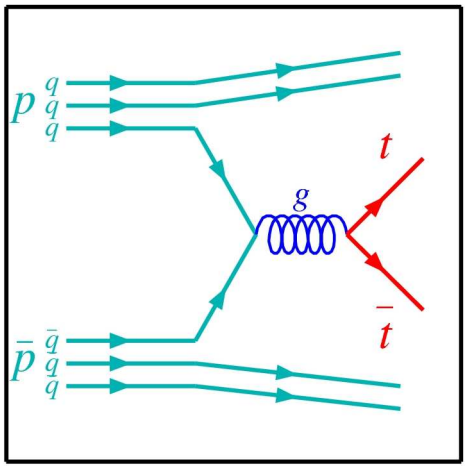

Figure 1.3: $p\bar{p} \rightarrow t\bar{t}$.

Parton distribution functions (PDF) $f(x, \mu)$ for gluons and quarks are shown in figure 1.4. In this figure the scale of the process μ has been set of the same order of M_{top}. Only two partons participate to the $t\bar{t}$ production, while the other ones are "spectators". At Tevatron Collider energies q-q scattering dominates the production process. If q and $\overline{q}$ carry the same momentum fraction x the threshold fraction for $t\bar{t}$ production is:

$$x_{thre} = \frac{2M_{top}}{\sqrt{s}} \tag{1.11}$$

At the Tevatron the center-of-mass energy $\sqrt{s}$ of the $p\bar{p}$ system is 1.96 TeV: therefore the minimum fraction required is ~ 0.18 [14].

[13] Single top production via the electroweak force has not yet been observed.
[14] this value is 0.025 at LHC ($\sqrt{s} = 14TeV$)

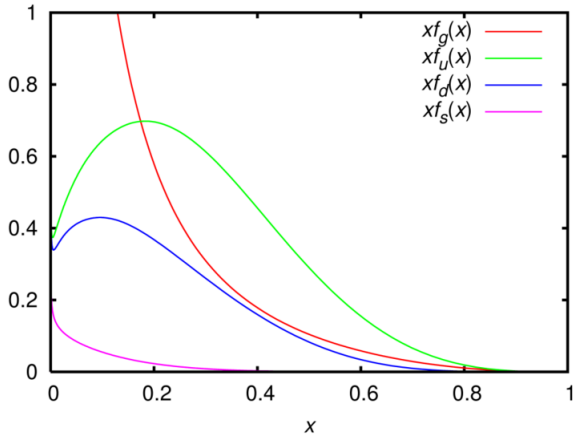

Figure 1.4: Parton distributions functions [42]

According to leading orders SM predictions $t\bar{t}$ pair are expected to be produced via $q\bar{q}$ annihilation ($q\bar{q} \rightarrow t\bar{t}$) approximately in the 85% of the cases and via gluon fusion ($g\bar{g} \rightarrow t\bar{t}$) in the remaining ones ([45]). Feynman diagrams for these processes are shown in figure 1.5. Next-to-leading order processes are expected to have a weight of 30% [15] in the estimated $\sigma^{NLO} = 6.7 \pm^{0.7}_{0.9}\ pb$ for $M_{top} = 175 GeV/c^2$ [12]. Some NLO order diagrams for $q\bar{q}$ annihilation are shown in figure 1.6. Diagrams of NLO order processes for gluon fusions are similar.

The cross-section is computed as:

$$\sigma(p\bar{p} \rightarrow t\bar{t}; s) = \sum_{i,j} \int_0^1 \int_0^1 f_i(x_1) f_j(x_2) \hat{\sigma}_{i,j}(q_i q_j \rightarrow t\bar{t}; x_1 x_2 s) dx_1 dx_2 \qquad (1.12)$$

where the i,j indices run over the parton families and gluon, $\hat{\sigma}_{i,j}$ is the cross sections at parton level, f_i is the PDF (see figure 1.4) for the i parton carrying a x fraction of the proton (antiproton) momentum.

With a cross section about 4 times smaller the top quark is also expected to be produced in the single-top channel via the weak interaction. The main processes that contribute to single top production are shown in figure 1.7

[15] see [13] for further details.

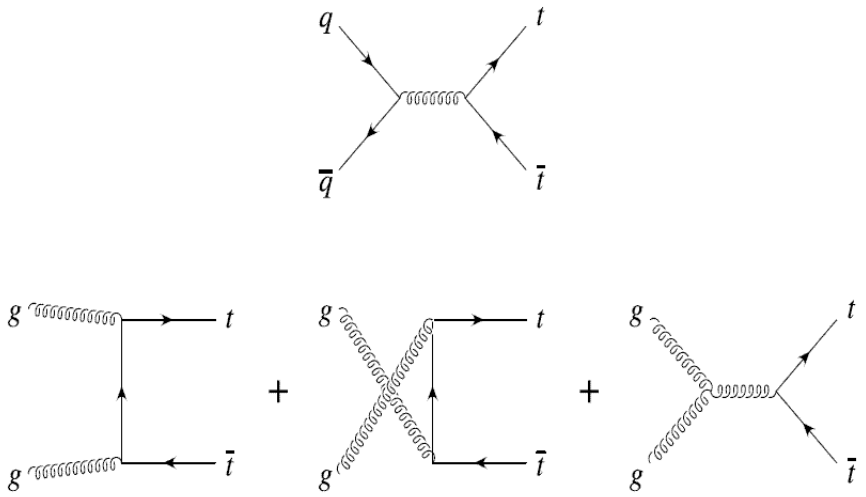

Figure 1.5: Leading order Feynman diagrams contributing to top quark pair production at hadron colliders.

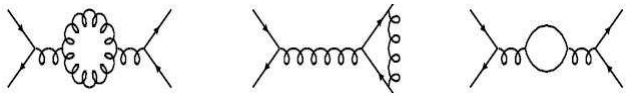

Figure 1.6: Some NLO Feynman diagrams for $t\bar{t}$ production via $q\bar{q}$ annihilation

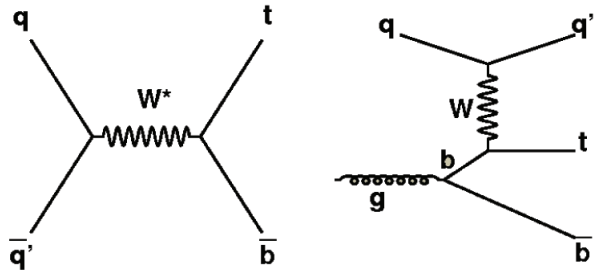

Figure 1.7: Single top production: leading processes.

1.3 Top Quark Decay

The top quark decay is mediated by the weak force. Since off-diagonal CKM matrix elements V_{ts} and V_{td} are much smaller than V_{tb} a BR of 100% can be assumed with an accuracy of $O(10^{-3})$. Therefore the intermediate state for a decaying $t\bar{t}$ pair is defined as follows:

$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \tag{1.13}$$

Moreover, the large mass of top quark results in a lifetime ($\tau_t \approx 0.5 \times 10^{-24} s$) shorter than the timescale required for quarks to hadronize. Because of this top quark can be considered free from "the confinement effect"..

Final $t\bar{t}$ states are classified according to the decay products of the two W's. Because of the large energy scale available lepton masses can be neglected and $W^-(W^+) \rightarrow l^-(l^+)\overline{\nu}_l(\nu_l)$ is expected to have the same BR independently of the lepton (l) flavor.

Three $t\bar{t}$ final state configurations (channels) are possible: all-hadronic, all-leptonic ("dilepton") and semi-leptonic ("lepton+jets" or "single-lepton"). The relative branching ratio (BR) of each final state are reported in table 1.3.

Decay mode	relative BR	Channel
$t\bar{t} \rightarrow (qq'b)(qq'b)$	36/81	Full-hadronic
$t\bar{t} \rightarrow (qq'b)(e\nu_e b)$	12/81	Lepton + jets
$t\bar{t} \rightarrow (qq'b)(\mu\nu_\mu b)$	12/81	Lepton + jets
$t\bar{t} \rightarrow (qq'b)(\tau\nu_\tau b)$	12/81	Lepton + jets (τ)
$t\bar{t} \rightarrow (e\nu_e b)(\mu\nu_\mu b)$	2/81	Dilepton
$t\bar{t} \rightarrow (e\nu_e b)(\tau\nu_\tau b)$	2/81	Dilepton (τ)
$t\bar{t} \rightarrow (\mu\nu_\mu b)(\tau\nu_\tau b)$	2/81	Dilepton (τ)
$t\bar{t} \rightarrow ((e\nu_e b)(\mu\nu_e b)$	1/81	Dilepton
$t\bar{t} \rightarrow (\mu\nu_\mu b)(\mu\nu_\mu b)$	1/81	Dilepton
$t\bar{t} \rightarrow (\tau\nu_\tau b)(\tau\nu_\tau bb)$	1/81	Dilepton (τ)

Table 1.3: Relative BR of the decay modes of the $t\bar{t}$ system

Antiparticles are not specified in table 1.3. q, q' can be u $(\overline{u}), d$ $(\overline{d})$, c $(\overline{c})$, s $(\overline{s})$.

A different experimental signature in the detector is associated with a different channel.

- The *all-hadronic channel*: both W bosons decay to quarks, resulting in a final state having a signature of six jets.

- The *lepton+jets channel*: one W decays to a lepton and neutrino and the other one to quarks, resulting in a signature of a high momentum lepton, four jets and missing transverse energy [16].

- the *dilepton channel*: Both W bosons decay to leptons resulting in a signature of two high momentum leptons, two jets and large missing transverse energy.

Relative branching ratios for any final state configurations are shown in the pie chart 1.8. $t\bar{t}$ events involving at least a τ lepton are grouped together and shown in this plot for completeness. They are much more difficult to study than other channels.

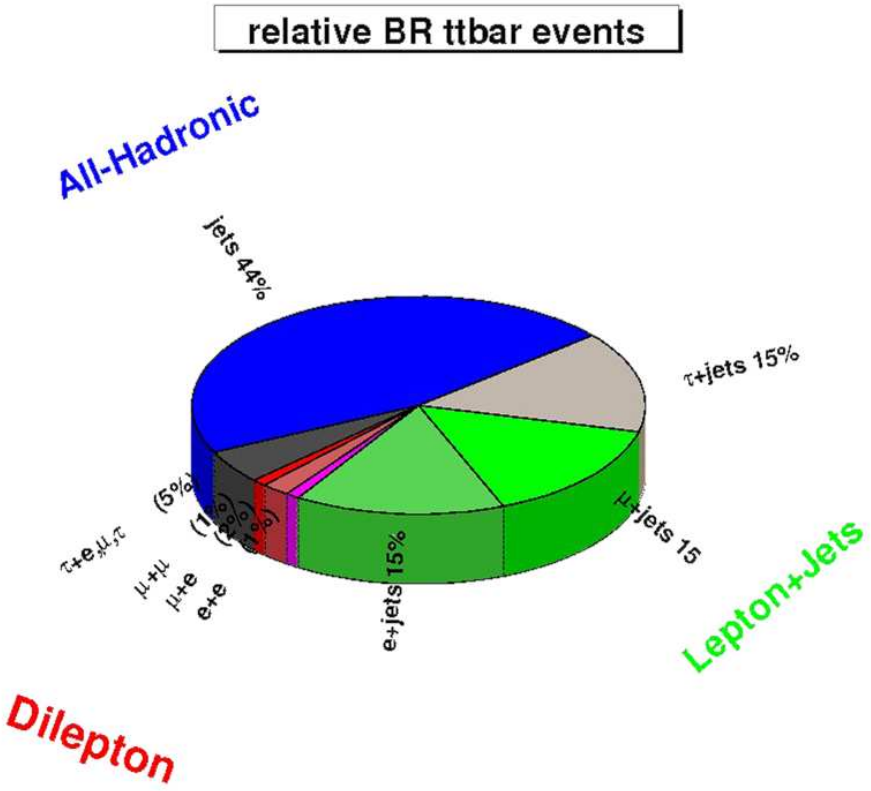

Figure 1.8: Relative BRs of $t\bar{t}$ decay.

[16] Assuming a zero transverse energy of the initial state, the missing transverse energy ($\not\!\!E_T$) is defined as the imbalance in final state transverse energy. For further details see 4.1

1.3.1 Importance of top mass measurement

The top mass measurement is one of the main goals of CDF. This measurement is extremely important because some parameters of the standard model depend on it.

The W boson mass contains loops corrections with a contribution by the top quark.

$$M_W = \frac{\frac{\pi\alpha}{\sqrt{2}G_F}}{sin^2\theta_W(1+\Delta R)} \tag{1.14}$$

where G_F is the Fermi coupling constant. The contribution to ΔR involving the top mass is: $(\Delta R)_{top} \approx \frac{3G_F M_t^2}{8\sqrt{2}\pi^2 tan^2\theta_W}$. $(\Delta R)_{top}$ is computed at NLO, whose Feynman diagram is shown in figure 1.9.

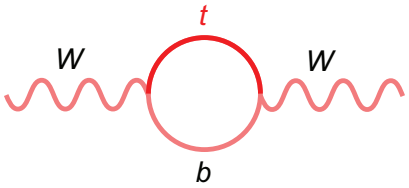

Figure 1.9: Virtual top quark loop contribution to the W mass.

Similar corrections at NLO are applied to the Z mass. The related NLO Feynman diagram is shown in figure 1.10.

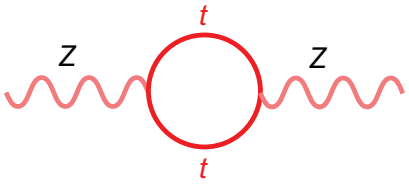

Figure 1.10: Virtual top quark loop contributing on Z mass

Therefore a top mass measurement allows to impose conditions on possible contributions by other virtual loops, including sub-dominant Higgs boson loops, to the physical W and Z boson masses.

Because of this, top mass measurements are asked to be as precise as possible in order to provide bounds to the Higgs mass via limits to virtual Higgs loops imposed primarily by the measured W mass. The present situation is illustrated by figure 1.11.

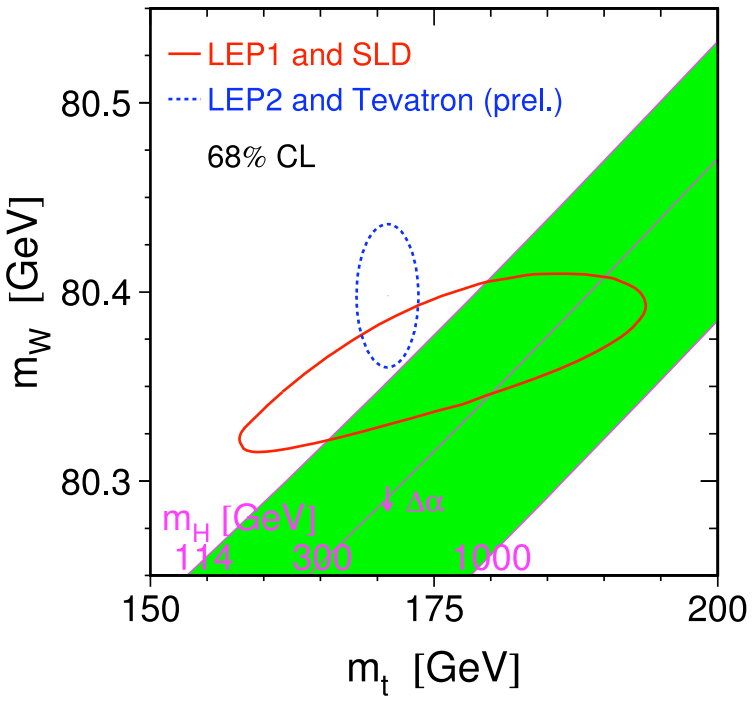

Figure 1.11: Top and W mass data constraining the Higgs mass. Expected correlations between M_t and M_W as a function of Higgs mass are shown as shaded band. Ellipses are drawn to span 1σ variation in top and W mass.

Low M_H values seem to be favored by current measured top and W masses.

An accurate measurement of top mass is also important in some extensions of SM. The minimal supersymmetric extension of SM sets a top mass dependent upper bound on the lightest neutral Higgs (h) mass. By using the word average top mass, the light SUSY Higgs boson h mass is expected to be below 140 GeV/c^2. If no h below this limit is found in the future, minimal SUSY will be ruled out

Chapter 2

Accelerator Complex

The Tevatron in Batavia, Illinois, US is the first large-scale superconducting synchrotron in the world. Originally named the *Energy Doubler* since as a proton-synchrotron it was reaching twice the energy of the original Fermilab facility (the "Main Ring"), it began operation in 1983 in fixed target mode and in 1985 as a proton-antiproton collider.

Since 1985 various periods of collider or fixed target operations or shut down for upgrading the machine alternated with each other. The on-going (2007) collider operation period is named Run II. The present $p\bar{p}$ energy ($\sqrt{s}$) in center of mass system (C.M.S.) is $1.96 TeV$.

After the Tevatron another high energy hadron collider (the *Large Hadron Collider*) is expected to collide proton beams in 2008. According to the expectations it will reach $\sqrt{s} = 14\ TeV$.

Along the Tevatron ring there are two apartment building-sized collider detectors, CDF (Collider Detector at Fermilab) and D$\emptyset$, that have undergone extensive upgrades during the 6 years long (1996 to 2001) preparations for Run II. Both experiments host over 600 physicists from allover the world.

A schematic layout of the Fermilab accelerator complex is shown in figure 2.1:

In section 2.1 this complex (accelerator feeding the Tevatron, and Tevatron Collider) will be briefly described.

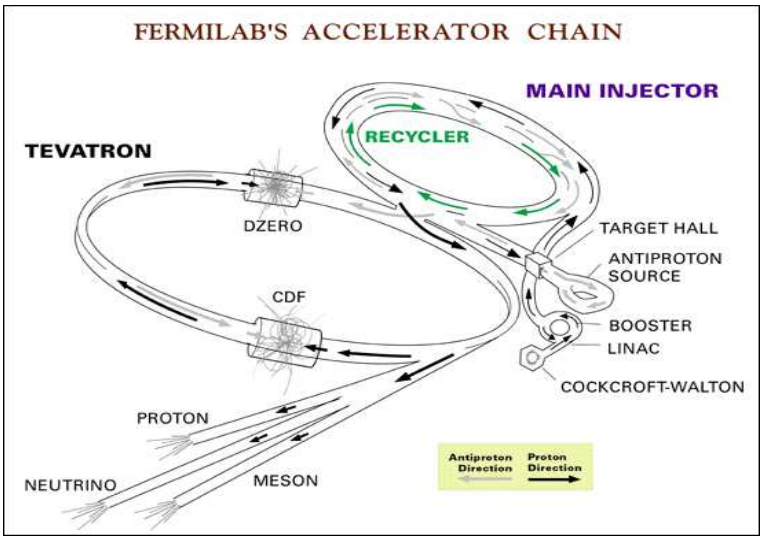

Figure 2.1: The accelerator complex of the Fermi National Accelerator Laboratory.

2.1 Acceleration Chain

2.1.1 Cockroft-Walton electrostatic Preaccelerator

Acceleration begins with a *Cockroft-Walton electrostatic Preaccelerator*. Here H^- ions are accelerated from a ion source to 750 KeV.

2.1.2 Linac

Ion bunches at 750 KeV are fed into the *Linear Accelerator (Linac)*. The Linac (figure 2.2) is approximately 140 m long and comprises two sections. In the first one five accelerating cavities with a drift tube design fed by a single RF generator accelerate ions to approximately 166 MeV. The second one, comprising 7 RF cavities fed by a more modern set of Klystron amplifiers, ramps ions by additional 400 MeV [1]. At the Linac exit the ion beam strikes a thin carbon target and turns into a proton beam by electron stripping.

[1] The 750 final energy is the result of a Linac upgrade, that took place in 1993 and increased the boost in the second Linac sector from 200 MeV to 400 MeV. This effort allowed to double the number of protons per bunch and to increase by about 50 % the production rate of antiprotons.

Figure 2.2: upstream view of the 400 MeV section of the Linac

2.1.3 Booster

Stripped protons enter the *Booster*, a 8-GeV synchrotron whose diameter is about 150 m. To maintain a constant circular orbit the dipole magnetic field in the Booster increases from 0.74 Tesla to 7 Tesla during acceleration.

Both Linac and the Booster provide pulses up to 5×10^{12} protons at a rate of about 5 Hz for antiproton production every $1.5s$, or 6×10^{10} protons per bunch in series of 5 to 7 bunches, repeated 9 times per second (in average). Figure 2.3 shows a view of the Booster.

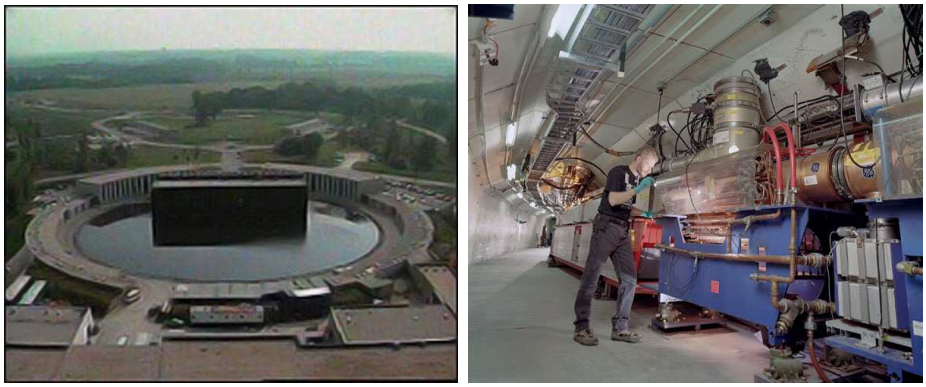

Figure 2.3: View of the booster building (left) and of the Booster tunnel (right)

2.1.4 Main Injector

From the Booster the proton beam is fed into the Main Injector whose role is either to accelerate protons as needed for injection in the Tevatron or to deliver beam to the antiproton production target.

The original Tevatron injector was the Main Ring, built to provide primarily 400 GeV protons to fixed target experiments. The main ring limited aperture was a limit to the whole accelerator performances. The *Main Injector* was designed to solve this problem and located in a separate tunnel for an easier operation of the complex.

The **Main Injector** is a synchrotron with circumference of about $3km$. It is able to accelerate protons of 8 GeV energy up to 150 GeV. It operates at 120 GeV for antiproton production, while 150 GeV protons are delivered to the Tevatron.

The Main Injector is also used to give beam to a number of fixed target experiments, noticeably on secondary neutrino beams.

An inside view of the Main Injector tunnel is shown in picture 2.4.

Figure 2.4: The Main Injector (blue magnets on the ground), completed in 1999, and the Recycler (green ring)

2.1.5 Debuncher and Recycler: Antiproton Production and Storage

A pulse of 5×10^{12} protons at $120GeV$ is extracted every $\sim$ 2.2 seconds from the Main Injector and directed to the antiproton station, a rotating 7 cm-thick

target made of nickel alloys containing chromium, iron and other metals. The resulting particles spray contains some antiprotons with a broad momentum and wide-spread spatial distribution.

A cylindrical lithium lens (760 T/m) focuses the particles produced around the forward direction. Negative particles in a 35 mrad cone about the forward direction are selected by a 1.5 T pulsed dipole magnet, focused by strong magnetic lenses and injected in the *Debuncher Storage Ring*. Typically, 21 antiprotons per 10^6 protons on target are collected.

In the **Debuncher ring**, a rounded triangular-shaped synchrotron with mean radius of 90 meters stochastic cooling and bunch phase rotation [2] to reduce momentum spread while increasing time spread.

After each beam pulse the Debuncher is emptied. The antiproton bunches (8 ± 0.018 GeV) are transferred with a 60%-70% efficiency to the **Antiproton Accumulator**, a 75 m mean radius storage ring of larger acceptance housed in the same tunnel as the Debuncher (see a sketch in figure 2.5). In the Accumulator multiple beam pulses are stacked and $\overline{p}$ are further cooled to increase the antiproton phase space density.

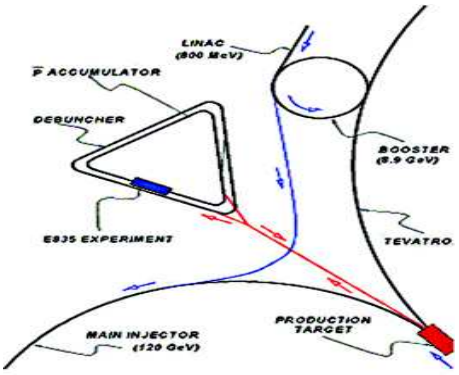

Figure 2.5: Fermilab antiproton Accumulator: a zoom-in of figure 2.1

For the time being problems in antiproton collection, cooling and stacking are among the main causes limiting the final Tevatron luminosity [3].

A further improvement of the antiproton source is the **Recycler**, a post-accumulator storage ring of constant 8 GeV energy, located in the Main-Injector

[2] Stochastic technique is a way of narrowing the particle distribution in transverse and longitudinal momentum around the average value. There is no associated beam-loss. This goal is achieved by applying iteratively a mechanism which recognizes deviation from the spatial orbit of a $8 - GeV$ antiproton in upstream sensors and makes appropriate correction downstream.

[3] as written above a very small fraction of the protons incident on target produces antiprotons and only a part of these can be stored.

(see figure 2.1) enclosure and composed of permanent magnets. Because of its larger acceptance it can store an antiproton current up to over 2.5 Amps, much larger than the Accumulator.

When a new store is ready for collider operation, the antiprotons are transferred from either Accumulator or the Recycler to the Main Injector in order to boost their energy up to 150 GeV. Antiprotons are then transferred to the *Tevatron*, where an opposite proton beam of the same 150 GeV energy was previously stored.

2.2 The Tevatron

The **Tevatron** is a $1 - km$ radius synchrotron. As written in section 2.1.4 the Tevatron receives protons and anti-protons from the Main Injector at 150 GeV. At injection 36 bunches composed typically of 300×10^9 p are transferred with a timing separation of 396 ns from each other.

Protons and antiprotons orbit are in the same vacuum pipe. Electrostatic separators reduce to a negligible amount the unwanted interactions, by keeping the beams away from each at all points in the orbit helix [4], except at the collision points.

Protons and antiprotons are accelerated to 980 GeV. A tour of the Tevatron takes about 21 μs. About one minute is needed to reach the final beam energy [5].

High-power focusing quadrupole magnets minimize the beam section at the interaction regions to maximize the collision rate.

The resulting transverse beam distributions are approximated by 2D Gaussian functions, with $\sigma_T \approx 30\ \mu m$. The typical longitudinal dimension of a bunch is 60-70 cm. The event source is roughly distributed longitudinally as a Gaussian with $\sigma_z = 28\ cm$ [6] .

Tevatron bunches are organized in three trains (see figure 2.6). Within a train the inter-bunch time is 396 ns while inter-train time is 2.6 μs. The intra-train empty sectors allow enough time for fast kicker magnets to abort the beam into a dump before the arrival of the next train in case of emergency.

[4] Intrabeam distance is typically 5 times the sum of the beam widths (in a Gaussian approximation)

[5] The Tevatron comprises about 1000 superconducting magnets including 772 dipoles. Each dipole is approximately 6 m in length and 4 tons in weight. The superconducting coils are made up of niobium-titanium wires embedded in copper. A 4400 A current in the dipoles provides a 4.2 T magnetic field. All superconducting magnets are kept at 4 K temperature

[6] The about 28 cm length of the interaction region is determined by the overlap of the two approximately longitudinally Gaussian bunches.

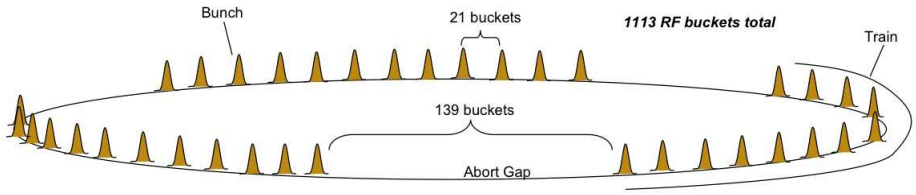

Figure 2.6: Bunch structure of the Tevatron beams in Run 2

The record peak luminosity [7] reached by the Tevatron is $\sim 2.92\text{x}10^{32}\ cm^{-2}s^{-1}$ (figure 2.7) corresponding to about 5 interactions per bunch-crossing.

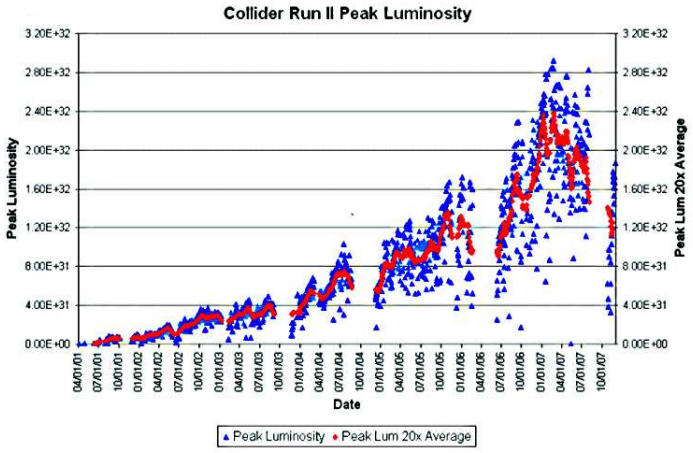

Figure 2.7: Run 2 Peak luminosity in the stores and its average, as a function of calendar date (up to January 2007).

[7] In January 2007

The integrated luminosity delivered up to August 2007 is $\sim 3\ fb^{-1}$. (figure 2.8). By Autumn 2010 the integrated luminosity is expected to reach 6-8 fb^{-1}.

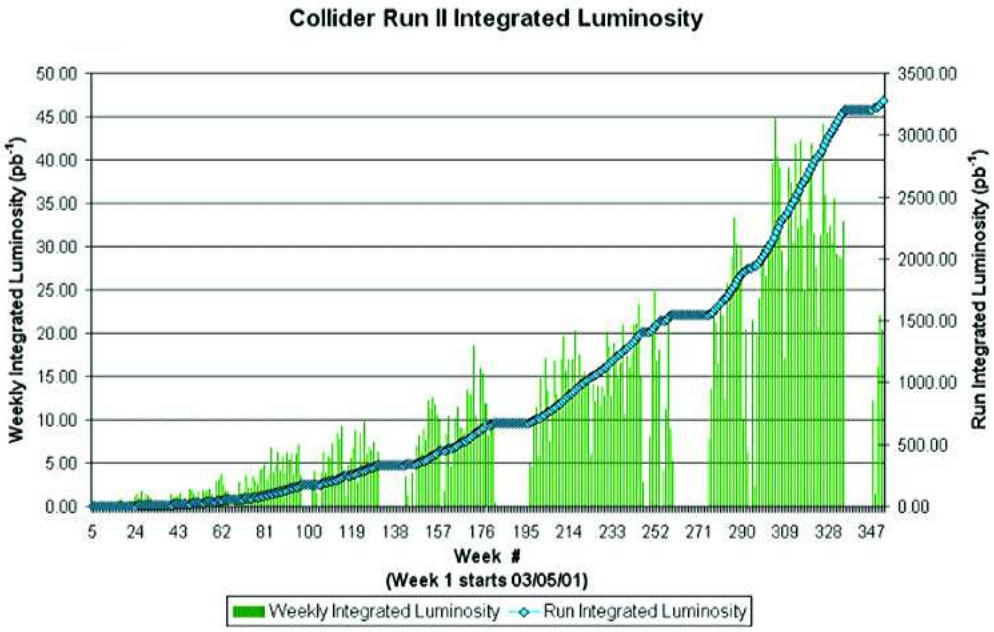

Figure 2.8: The integrated luminosity in Run 2 weeks. Empty periods are due to Tevatron shutdowns

Chapter 3

The Run II CDF Detector at Fermilab

The Collider Detector at Fermilab (CDF) was designed to study $p\bar{p}$ collisions at the Tevatron. Commissioned in 1987 it was upgraded in 2001 in order to be adapted to the higher collision rate coming from the increased instantaneous luminosity delivered by the accelerator.

A cylindrical (r, ϕ ,z) coordinate-system is used to describe CDF geometry. The origin is the geometric center of the detector. The z-axis points along the direction of protons. Polar and azimuthal angles relative to the z-axis are used to describe particle trajectories.

It is often convenient to use a polar variable invariant under boost along $\hat{z}$. This variable is the *rapidity* defined as

$$y = \frac{1}{2} ln(\frac{E + pcos(\theta)}{E - pcos(\theta)}) \tag{3.1}$$

where E, p, θ are respectively the energy, momentum and polar angle of the considered particle. At high energies and away from very forward angles $y \approx \eta \equiv -ln[tan(\theta/2)]$ [1].

3.1 Overview

The Run II Detector (see figure 3.1, 3.2) is composed of several components, each optimized for a specific task.

[1] In CDF literature one distinguishes η_{det}, which is relative to the geometrical center of detector, and η, which is measured with respect to the interaction point z_0 where particles originated. Usually the former symbol is used for describing the detector geometry and the latter for outgoing particles. For simplicity the same symbol will be used in both cases.

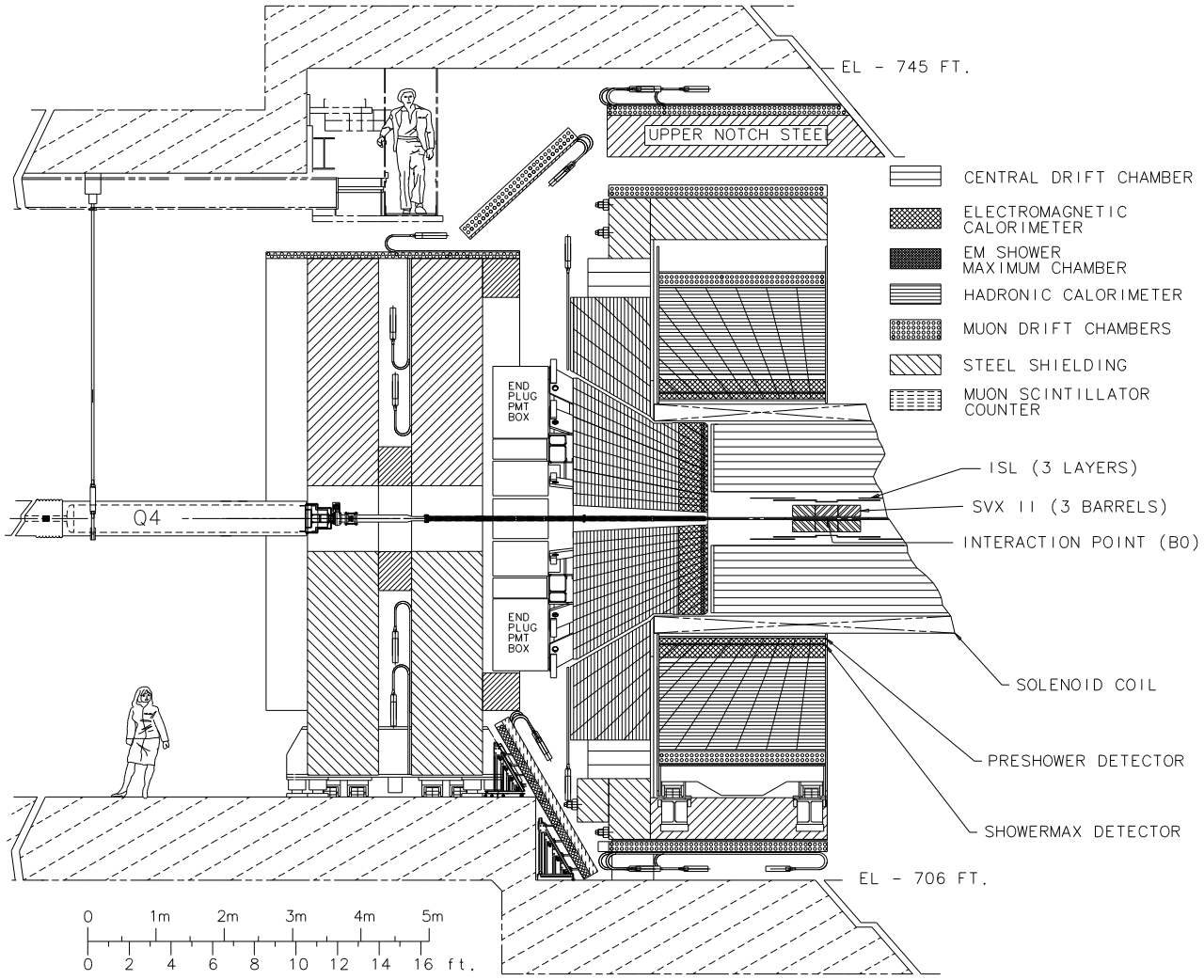

Figure 3.1: Elevation view of the CDF II detector

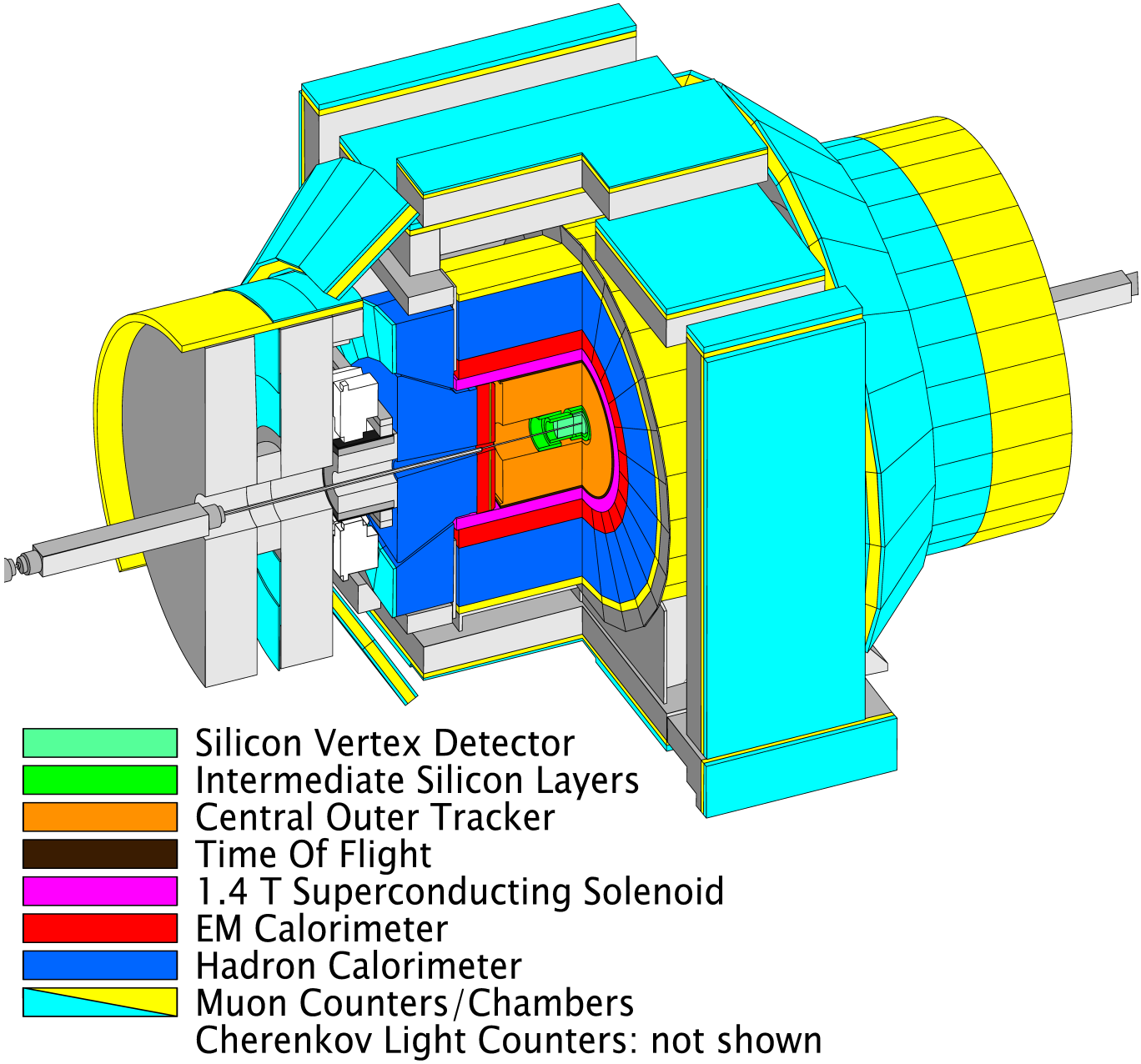

Figure 3.2: Artist view of the CDF Run II detector.

Starting from the interaction point and following the path of an outgoing particle within acceptance there are:

- a tracking system enclosed by a superconducting solenoid (1.5 m in radius and 4.8 m in length), which generates 1.4 T magnetic field parallel to the beam axis. The magnetic field is nearly uniform in the tracking region.
- finely segmented calorimeters
- planar drift chambers backed by scintillation counters.

In the next sections these sub-systems will be discussed.

Some of the components (the time-of-flight detector, the Cerenkov Luminosity Counters, the small angle spectrometers on the beam pipe, etc.) of the full CDF II detector have been neglected since are not directly related with this thesis.

A detailed description of the upgraded detector can be found in [43].

3.2 Tracking System

Charged particles within the tracking system acceptance encounter an inner silicon tracking system and outer gas drift-chamber as shown in figure 3.3.

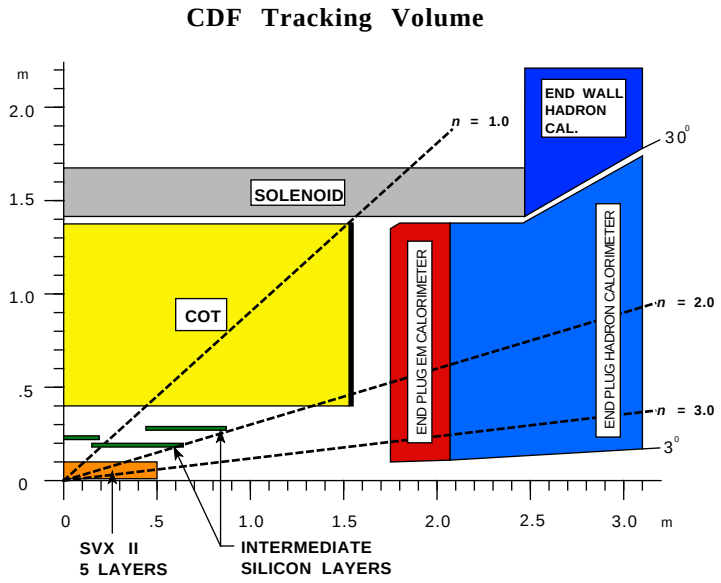

Figure 3.3: Longitudinal view of the Tracking System of the CDF II detector.

Within the solenoid field they follow helical trajectories which are measured by the system in order to measure their momentum.

3.2.1 Inner Tracker

The Inner Tracker is composed of eight layers (seven at $\theta = 90°$) of silicon sensors arranged in approximately cylindrical sub-systems coaxial with the beam-pipe: Layer 00 (L00) [14, 15], the Silicon Vertex Detector (SVX II), and the Intermediate Silicon Layers (ISL) [16, 17]. Figure 3.4 zooms on the Inner Tracker from two different point of views.

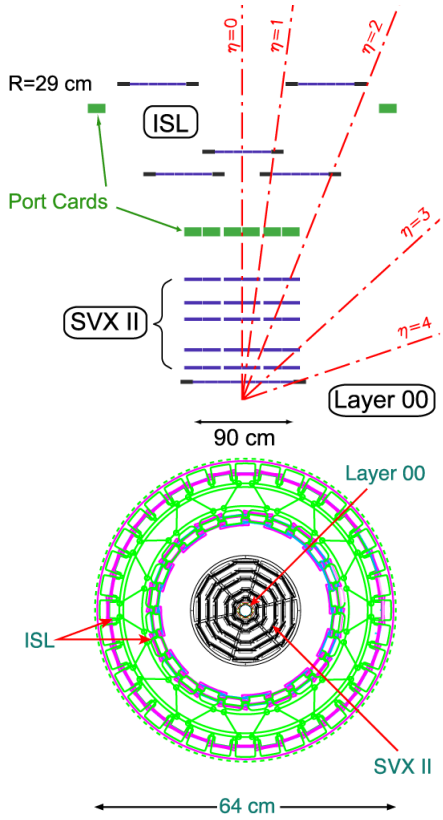

Figure 3.4: Upper: cutaway transverse to the beam of the three inner tracker sub-systems. Lower: sketch of the silicon detector in a x/y projection.

All silicon microstrip sensors have a space resolution of $\sim 12\ \mu m$ in the direction transverse to the beam. They also provide z-measurements with reduced accuracy.

▶ L00

L00 is 90 cm long, radiation hardened cylindrical arrangement of silicon strip detectors, that is mounted directly on the beam pipe: it is at radial distance of 1.35 to 1.62 cm from the beam axis. L00 purpose is to improve the track impact parameter (the measured distance of minimum approach to the beam axis) resolution ($\sim$ 25/30 μm) and compensate for the degraded performance, due to radiation damage, of the other silicon sub-detectors.

▶ SVX II

SVX II, shown in an artist view 3.5, extends radially from 2.5 cm to 10.7 cm.

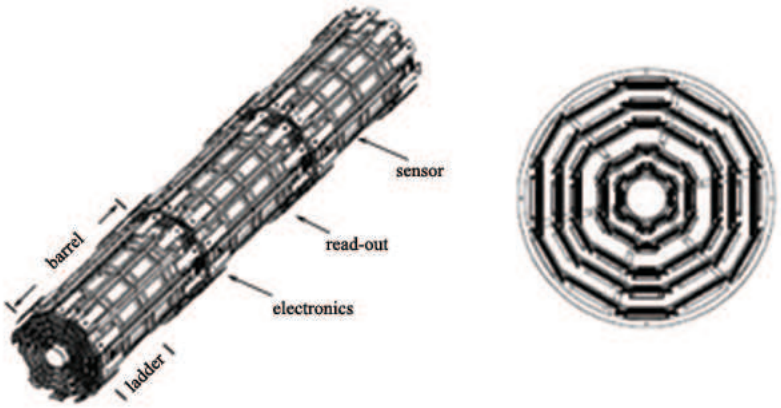

Figure 3.5: Isometric (left) and end-view (right) of the CDF Silicon Vertex Detector

It is segmented into three 29 cm barrels along the z axis: this allows for a $\eta < 2.0$ coverage. Each barrel carries 5 layers of double-sided microstrip wafers. Four silicon wafers are aligned on common support structures called ladders. Twelve concentric ladders make a layer.

The double-sided imprint of the wafers allows for 3D position measurements: one side of the wafer has strips along the beam axis, the other one has either 90°or 1.2°stereo strips.

This sub-detector has a 12 μm resolution on the single hit in the direction transverse to the beam on the single hit, and provides also some dE/dx information.

► **ISL**

The main ISL purpose is to compensate for incomplete coverage of other sub-detectors in the region $|\eta| > 1$ by providing precision tracking at $1 < |\eta| < 2$. It consists of 5 layers of double sided silicon wafers (same wafers as for SVXII). Four layers are at $1 < |\eta| < 2$ (at radii of 20 and 28 cm, as shown in figure 3.4), one layer is at $|\eta| < 1$.

The combined resolution of the CDF inner trackers for high momentum tracks is $\sim 40\ \mu m$ in impact parameter and $\sim 70\ \mu m$ along z direction [14].

3.2.2 Central Outer Tracker

The main tracker at CDF is the Central Outer Tracker (COT). The COT has a cylindrical shape and is radially right outside the ISL. Its internal and external radii are 43 and 137 cm respectively. COT provides full tracking in the central region ($|\eta| < 1$), even if its maximum geometrical acceptance reaches up to $|\eta| < 2$ (see figure 3.3), where tracking performances are reduced.

The COT is composed of 4 axial and 4 stereo [2] superlayers of azimuthal cells. Each cell [3] has alternated sense and field shaping wires (figure 3.6). Within the cell width, the trajectory of a particle is sampled 12 times (by sense wires spaced 0.583 apart). Figure 3.6 shows a portion of the COT endplate.

Figure 3.7 shows the equipotential electrical field lines in a cell.

Inside the solenoid magnetic field, the drifting electrons experience a Lorentz force which rotates their path. The cells are tilted by 35°with respect to the radial direction in order to make the electrons drifting perpendicularly to the radius for optimal momentum resolution [4].

The single hit position has been measured with an uncertainty of $\sim 140\ \mu m$ which translates into an overall transverse resolution $\sigma(P_T)/P_T = 0.15\%\ P_t[GeV/c]$ [18], P_T being the transverse momentum of the tracked particle.

3.2.3 Calorimeters

The CDF calorimeter measures the particle energy by absorbing their total energy and providing a signal proportional to it. Calorimeter information is also used to

[2] stereo superlayers are tilted at ±2°with respect to the z direction. Axial layers provide tracking information in $r - \phi$, stereo layers are also sensitive to the z direction.

[3] The number of cells in a superlayer increases linearly with the radius.

[4] For best momentum resolution, the optimal correlation between drift time and hit distance from wire is for electrons drifting in the direction transverse to the radius

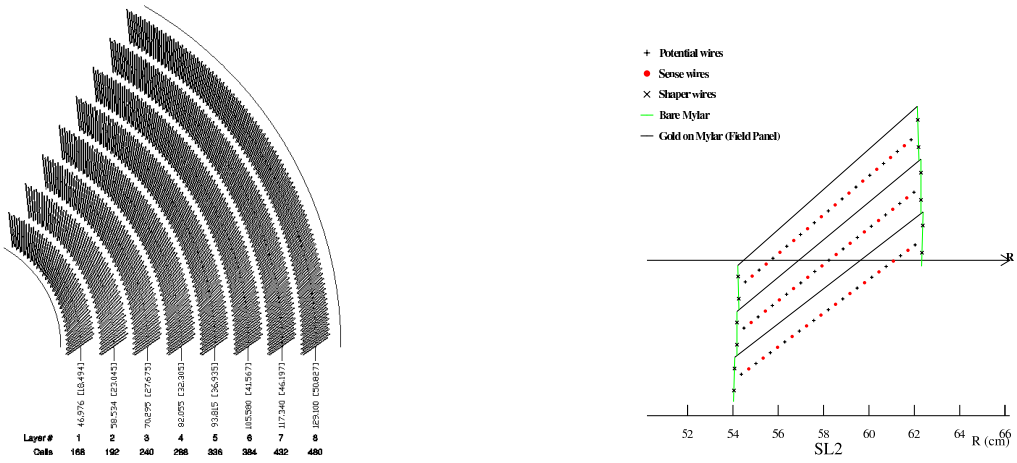

Figure 3.6: Slots housing the wire holding fixtures of one quadrant of a COT endplate (left). Drift and field wires in three cells (right). The horizontal arrow shows the radial direction.

estimate the transverse energy of weakly interacting particles, such as neutrinos, by computing the imbalance in the total transverse energy. CDF uses scintillator sampling calorimeters divided into an inner electromagnetic and an outer hadronic compartment. Both calorimeters are segmented into projective towers. Each tower consists of alternating layers of passive absorber material (Pb in the front and Fe in the rear compartment) and plastic scintillator for shower sampling. The light from the scintillator plates is read out through wavelength-shifting (WLS) bars or plates and light guides by photo-multiplier tubes (PMT) (see figure 3.8).

High energy electrons and photons generate an electromagnetic shower which is mostly absorbed in the front calorimeter compartment [5]. For charged particles heavier than the electron, radiative energy losses are negligible to a first approximation.

Hadrons interact with the detector matter mostly through inelastic collisions with nuclei of the absorbing medium. Particles produced in the nuclear interactions can loose their energy by ionization and secondary nuclear interactions. Mixed electromagnetic and hadron showers that originate in this process are absorbed in the entire (front + rear compartments) calorimeter.

CDF calorimeters provide full azimuthal coverage and up to 3.6 in $|\eta|$ [6]. They include the Central Electromagnetic Calorimeter (CEM) and the Hadronic Calorime-

[5] In the case of photons and electrons the showers are composed mainly of electrons, positrons and photons.

[6] Calorimeters (Miniplugs) for measurement of diffractive processes are not considered in this thesis. More information is available at [25]

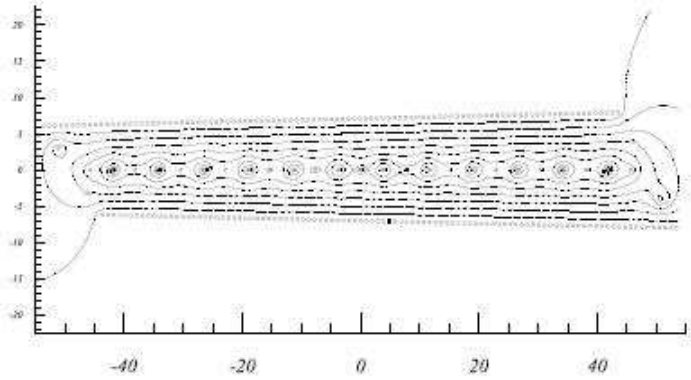

Figure 3.7: Equipotential lines in a COT cell

ter (CHA) in the $|\eta| < 0.9$ region, the Endwall Hadronic Calorimeter (WHA) at $0.9 < |\eta| < 1.3$ and the electromagnetic and hadronic plug calorimeters (PEM, PHA) at $1.1 < |\eta| < 3.6$ (figures 3.1, 3.3).

Central Calorimeter

The central calorimeters, CEM, CHA and WHA are composed of two parts joining at the geometrical center of CDF [7]. Central calorimeters are azimuthally divided into 24 wedges covering $\sim 15°$in ϕ each. Each wedge is divided into projective towers of size $\delta\eta = 0.11$.

The **CEM calorimeter** is made of 31 alternate layers of 0.5 cm thick plastic scintillator plates and 0.32 cm thick lead absorbers: the total amount of material is $18 \cdot X_0$ (X_0 is the electron radiation length). The CEM energy resolution is:

$$\sigma_{E_T}/E_T = 13.5\%\sqrt{E_T[GeV]} \oplus 2\% \tag{3.2}$$

E_T being the energy of an electron or a photon hitting the calorimeter perpendicularly to its front.

CEM also includes two additional specialized detector: the Central Electron Strip Chambers (CES) and the Central Preshower (CPR). CES is a combined strip/wire gas proportional chamber embedded in CEM at about $6 \cdot X_0$ [8]. The CES purpose is to measure the position and the shape of electro-magnetic showers

[7] In this zone, $\eta = 0$, there is an un-instrumented area about 20 cm thick in the z-direction ("crack")

[8] the maximum energy density in the longitudinal development of the electromagnetic shower is expected at about $6 \cdot X_0$

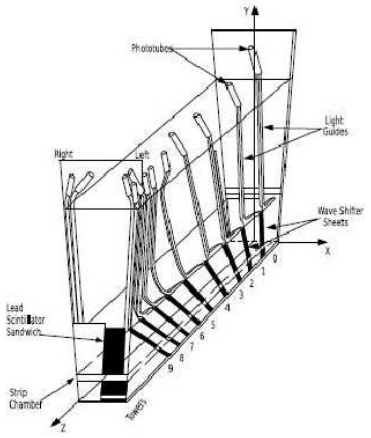

Figure 3.8: Light-shifter plates connected to light guides and to photomultipliers of the front electromagnetic compartment of a central calorimeter wedge.

in both transverse plane and longitudinal direction. CES resolution is about 1 cm in z and 1 mm in r-ϕ. CPR is a set of scintillator tiles located in front of the calorimeter wedges which help distinguishing electrons from charged hadrons by gauging their probability of showering in the detector material prior to entering the calorimeter [20].

The **CHA calorimeter**, located behind CEM, is composed of 32 alternate layers of 1 cm thick plastic scintillator and 2.5 cm thick steel. The **WHA calorimeter** employs the same technology as CHA except for the smaller number of layers (15) and the larger thickness of the radiator plates (5 cm). The total calorimeter thickness is $\sim 4.7\ \lambda_0$ (λ_0 is the charged pion absorption length) for both CHA and WHA.

Resolutions of CHA and WHA for perpendicular particle entrance are:

$$CHA: \quad \sigma_{E_T}/E_T = 50\%\sqrt{E_T[GeV]} \oplus 3\% \tag{3.3}$$

$$WHA: \quad \sigma_{E_T}/E_T = 75\%\sqrt{E_T[GeV]} \oplus 4\% \tag{3.4}$$

Plug Calorimeters

The **PEM calorimeters** (see figure 3.9) have the same structure as CEM: same tower segmentation in η, but finer in ϕ (7.5°wide ϕ bins) for $|\eta| < 2.11$, 22 layers of 4.5 mm thick lead alternate with 22 layers of 4 mm thick scintillator [21].

The PEM transverse energy resolution is:

$$\sigma_{E_T}/E_T = 16\%\sqrt{E_T[GeV]} \oplus 1\% \tag{3.5}$$

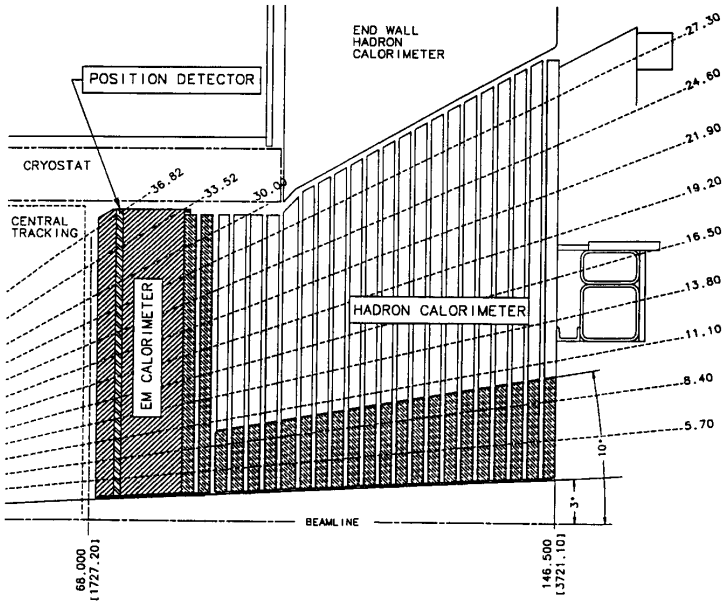

Figure 3.9: Longitudinal view of Plug Calorimeters

As for CEM, PEM is equipped with a shower maximum detector (PES). Details are reported in [22].

PHA, located behind PEM, has the same tower segmentation. The technology is the same as for CHA, with 23 layers of 2 cm thick steel absorber alternating with 6 mm thick scintillator. The total amount of material corresponds to ~ 4.7 λ_0. PHA resolution is:

$$\sigma_{E_T}/E_T = 80\%\sqrt{E_T[GeV]} \oplus 5\% \tag{3.6}$$

3.2.4 Muon Detectors

Although muon interact mostly electromagnetically as the electrons, because of their much larger mass they can cross a much larger amounts of material before losing a significant fraction of their energy [9]. For this reason systems dedicated to detect muons are located in the outermost shell of the detector. Muon momenta are measured in the tracker.

[9] At Tevatron energies muons interact in calorimeters as minimum ionizing particles (MIP).

Four independent systems are used to detect muons in the $|\eta| < 1.5$ region: the Central Muon Detectors (CMU), the Central Muon Upgrade Detectors (CMP), the Central Muon Extension (CMX), the Intermediate Muon Detectors (IMU). The $\eta - \phi$ coverage of the Run II muon detectors is shown in figure 3.10.

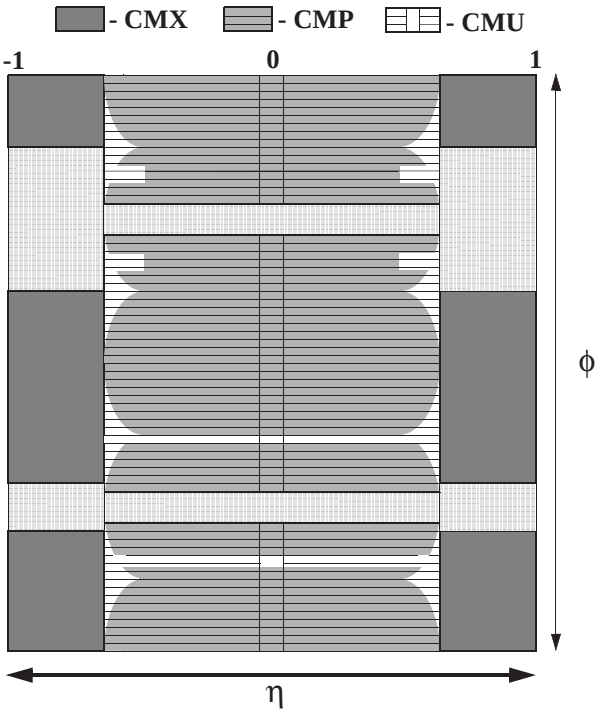

Figure 3.10: Coverage of muon detectors in the (η,ϕ) space.

Muon detectors share common features [23, 24]. They consists of stack of rectangular drift chamber modules [10] composed of single-wire cells. Stacks are four layers deep with laterally displaced cells from layer to layer to compensate for cell edge inefficiencies. The difference in drift-electrons arrival-times between neighbour cells provides a typical resolution of 250 μm for the hit position in the transverse plane. Charge division at the wire ends measures the z coordinate with a 1.2 mm resolution.

Chambers are coupled with scintillator counters in order to suppress backgrounds due to secondary interactions in the beam pipe material and to cosmic

[10] chambers are filled with a mixture of argon and ethane (50% each)

rays.

A muon candidate is reconstructed when a short track segment (stub) in the muon chambers corresponds to the extrapolation of a COT track..

The **CMU detector** is behind CHA at a radius of 347 cm from the beam axis and covers the $|\eta| < 0.6$ region. CMU consists of 144 modules with 16 cells each. The **CMP detector** is arranged to enclose the $|\eta| < 0.6$ region in an approximately central box (see 3.10). Scintillator layers (CSP) on the outermost side of the CMP chambers allow identifying bunch crossing. CMU/CMP system is called CMUP. It detects muons with a minimum energy of ~ 1.4 GeV.

The **CMX detector** extends the muon identification in $0.6 < |\eta| < 1$ region. As for CMP cells are sandwiched to scintillators (CSX).

The forward region of muon system is the **IMU detector** ($1.0 < |\eta| < 1.5$). The associated scintillator counters sub-system is named BSU.

3.2.5 Trigger System

At Tevatron Run II the interaction rate is typically 2.5 MHz (the bunch crossing frequency), which is much higher than any possible event recording rate (order of 100 Hz).

However, $p\bar{p}$ interactions are mostly inelastic, elastic and diffractive with no significant momentum transfer. More interesting events have cross-sections from 10^3 to 10^{12} times smaller than inclusive $p\bar{p}$ cross section (for example $\sigma_{p\bar{p}\to t\bar{t}} \approx 7pb$).

The identification of the interesting events is accomplished by dedicated fast online electronics, called the **Trigger System**, which evaluates the information from the detector and makes an accept/reject decision in real time. The trigger system (see figure 3.11) is a three-tier system, where each level of electronics performs a slower but more accurate event reconstruction and applies a tighter filter with increasing trigger level according to a set of pre-defined conditions.

■ Level 1

The level 1 (L1) is a synchronous pipeline system where up to 42 subsequent events can be stored for $\sim 5.5\ \mu s$ while the hardware is making a decision. If no acceptance decision is made up to that time the event is rejected [11]. L1 decisions are made in average in about 4 μs: no dead time is expected from this level. L1 rejects $\sim 97\%$ of the events and typical output rate is ~ 50 kHz.

The L1 decision is generated by:

[11] In order to be as fast as needed by the no-dead time condition L1 employs only hardware

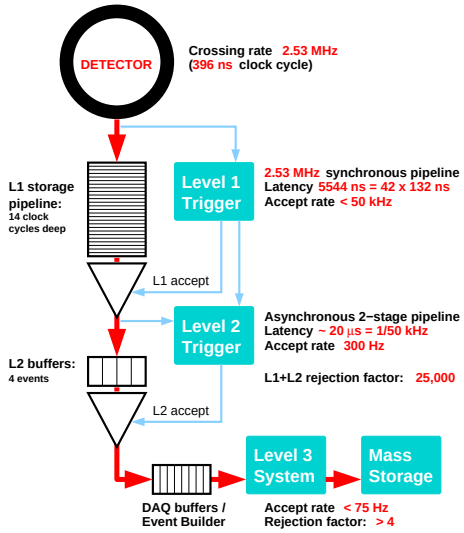

Figure 3.11: CDF trigger block diagram

- XFT (extremely fast tracker), which reconstructs approximate tracks ($P_T \geq 1.5 GeV$) in the transverse plane by exploiting information from COT superlayers [12]. These tracks can be propagated to the calorimeter and to the muon chambers to contribute to higher trigger levels.
- the calorimeter trigger, which indicates large energy releases in the electromagnetic or hadronic cells (these can be seeds for electrons or jets identification at higher levels)
- the muon trigger, which matches XFT tracks to stubs (see 3.2.4) in the muon chambers.

■ Level 2

Events accepted at level 2 (L2) are sent to 4 asynchronous buffers. Buffers are used to store events until a decision is made. Because of the limited size of buffer deadtimes may occur. L2 purposes are:

- to add the energy deposited in the towers in small regions around L1 seeds, as an approximate measure of an electron or jet energy

[12] It searches the 4 axial SL for track segments, then the Linker Board tries to link together at least three of them.

- to combine calorimeter and CES (see 3.2.3) information to signal candidates electrons
- to reconstruct a full COT track and associate it to an outer muon stub (see 3.2.4) in order to improve muon signature
- to indicate tracks with a large impact parameter by means of the Silicon Vertex Tracker (SVT) which allows to trigger on secondary vertexes from decay of long-lived beauty hadrons.

L2 acceptance rate is $\sim$ 300 Hz (rejection factor $\sim$ 150)

The block diagram of L1 and L2 with the involved subdetectors is schematized in figure 3.12.

■ Level 3

Level 3 (L3) is a software trigger. L3 addresses event objects delivered by L2 to the Event Builder (EVB), which reconstructs the entire event with the same accuracy as in the offline analysis.

The final decision to accept an event is made on the basis of a list of observables indicating candidate events of physical interest (top production events, W/Z events, Drell-Yan events, etc.).

Accepted events exit L3 at a rate of up to 100 Hz and are permanently stored on tapes.

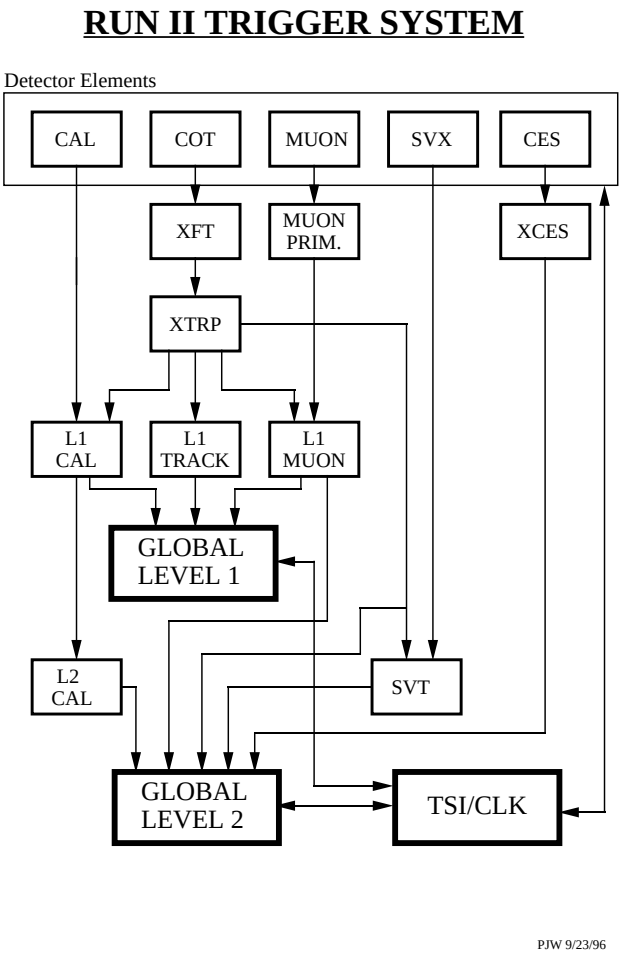

Figure 3.12: Block diagram of level 1 and level 2 triggers. The involved sub-detectors are indicated.

Chapter 4

Object Identification and Event Selection

Information of the CDF sub-detectors are combined in order to reconstruct the "physical objects" (leptons, jets, etc.) of interest for the analysis (section 4.1). In the top candidate event sample these objects are used by a machinery, called "event selection", in order to skim the events of interest (section 4.4). The event selection machinery is driven by the peculiarities of the $t\bar{t}$ signature in the dilepton channel (section 4.2) and by the need of reducing the number of background events [1](section 4.3).

4.1 Object Identification

Outgoing particles from $p\bar{p}$ interactions are identified using the information from tracker, electromagnetic and hadronic calorimeters and muon detectors. Detector sub-systems are described in chapter 3, whose notation is going to be used in the rest of this Chapter.

4.1.1 Tracks

In the tracking system charged particles travel on a longitudinal helix in the solenoidal magnetic field. The helix curvature (C) allows for particle P_T measurement, since:

$$P_T = Bq\rho = \frac{Bq}{2C} \tag{4.1}$$

[1]The standard way to optimize selections is to maximize $S/\sqrt{S+B}$, S and B being the expected signal and background.

where B is the magnetic field, ρ the curvature radius in the transverse plane, q the charge of the particle.

- **COT+SVX Tracking Algorithms**

The following tracking algorithms exploit the information from the COT and SVX sub-detectors:

✓ *Stand-alone*: only SVX information are used.

✓ *Inside-Out*: SVX response is used to find seed tracks and COT hits are linked. This algorithm allows tracking beyond $|\eta| = 1$ region.

✓ *Outside-In*: COT tracks are used as input and the available SVX hits are linked. This algorithm allows tracking in $|\eta| < 1$ region.

Since this analysis uses the Outside-In (OI) algorithm to reconstruct the central tracks ($|\eta| < 1$) only OI is going to be described. More details about the other algorithms are available in [26, 29].

The Outside-In algorithm implements the following steps:

1. every COT super-layer is looked for groups (segments) of three aligned hits.
2. Threefold groups are fitted by a straight line. All close hits in the same super-layer [2] are attached and the fit is repeated until a segment is completed.
3. A track is reconstructed from the segments by the standard CDF linking algorithms [3].
4. If a track is reconstructed, stereo hits are attached to the track by the linking algorithms.
5. A global fit is performed taking into account corrections for field non-uniformity and the modeling of electron drift.
6. If available, SVX hits are attached to the track and the fit is redone.
7. If more tracks have hits in common the track with the largest number of hits is accepted [4].

[2] the distance between hit and fitted straight line must be less than 1 mm

[3] At CDF two algorithms have been implemented: the "segment linking" and "histogram linking" [27].

[4] If two or more tracks score equal, the one with the best fit is selected.

- **Phoenix algorithm**

The "Plug" region ($1 < |\eta| < 3.6$) is not covered by the COT: a calorimetry-seeded tracking algorithm, named "Phoenix algorithm", is needed to identify tracks.

Phoenix algorithm reconstructs the track of an electron (Phoenix electron or PHX) by using:

i. the reconstructed 3D point where the $p\bar{p}$ interaction took place (primary vertex).

ii. the 3D position of the EM shower estimated by PES.

iii. the transverse energy deposited in the calorimeter.

The primary vertex and shower position are used as seed for the track helix and the helix curvature is varied to match the transverse energy. If such a match is found available hits from SVX are added to allow for a better fit.

4.1.2 Electrons

Electron reconstruction starts when a cluster in the electromagnetic calorimeter ("EM cluster") is found. Iterating over all calorimeter tower seeds [5], the EM candidate clusters are defined by adding the transverse energies of the seed towers and their neighbours [6]. For CEM neighbours towers can be 2 on each side in η direction, while for PEM the maximum cluster size is 2x2 including the seed. A cluster is found if $E_T^{em} > 2\ GeV$ and $E_T^{had}/E_T^{em} < 0.125$, where E_T^{em} and E_T^{had} are the EM and hadronic energies in the cluster. The clustering algorithms look for EM objects in the CEM and PEM separately.

CES/PES [7] information are used to estimate the cluster center and tracks (see 4.1.1) are linked to the cluster. An electron is finally identified if the cluster is matched with at least one track [8].

4.1.3 Muons

Muons are identified by stubs in muon chambers (see 3.2.4) and COT tracks. Muon stubs must be matched to tracks. Tracks must satisfy certain quality criteria in order to be considered for the matching [30].

[5] $E_T \geqslant 2\ GeV$ in a tower seed

[6] a tower is defined by using electromagnetic and hadronic calorimeters info.

[7] depending on the region where the cluster is reconstructed

[8] If this criterion is not satisfied the cluster may be a photon.

At Tevatron energies ionization losses dominate in the passage of muons through matter. The small ionization energy lost by muons in the calorimeter provides a significant signature.

Muons can be faked by hadrons whose showers leak from the back of the hadronic calorimeters, or by cosmic rays [9].

4.1.4 Jets

Outgoing quarks and gluons undergo the hadronization process whereby hadron showers are created around the primary parton direction. Figure 4.1 represents the steps from parton to jet level.

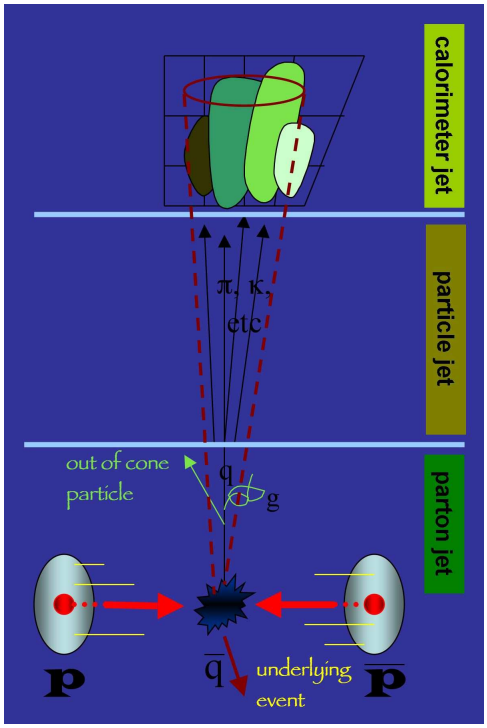

Figure 4.1: From the initial partons to the final jets.

Jets are the only objects that can be used to obtain information about final state partons of the hard interaction. Clusters of energy deposited in localized

[9] Cosmic muons can be rejected by imposing tight requirements on their distance from the beam line.

areas may indicate the presence of jets. However, there is no universal way of defining a jet because of the lacking information about the hadronization process.

■ Jet Reconstruction

A number of jet reconstruction algorithms have been developed at CDF. Some algorithms make use of tracking information in searching for charged jets or in measuring their transverse momenta. The jet algorithm used in this analysis, called JETCLU [31], relies only on calorimeters.

JETCLU belongs to the family of cone algorithms. The cone radius is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{4.2}$$

where $\Delta\eta$ and $\Delta\phi$ are the distances in pseudorapidity and azimuthal angle between a tower center and a cluster axis.

JETCLU proceeds through the following steps:

1. searching for towers with $E_T \equiv E^{em} sin\theta^{em} + E^{had} sin\theta^{had} > 1\ GeV$, where θ^{em} (θ^{had}) is the polar angle of the vector pointing from the primary vertex [10] to the geometrical center of the electromagnetic (hadron) tower.

2. The towers above 1 GeV are marked as precluster seeds and ordered with increased E_T.

3. All seeds in a 49-towers square centered on the highest E_T seed, define a precluster. Seeds cannot belong to more than one cluster. The centroid of the precluster is calculated as the E_T-weighted center of the seeds.

4. A $\Delta R = 0.4$ [11] cone is drawn around the centroid and all $E_T > 0.1\ GeV$ towers in the cone are summed. The centroid is then recalculated as follows:

$$(\eta^{bPC}, \phi^{bPC}) = (\frac{\sum_j \eta_j E_{Tj}}{\sum_j E_{Tj}}, \frac{\sum_j \phi_j E_{Tj}}{\sum_j E_{Tj}}) \tag{4.3}$$

5. The step above is iterated until $(\eta^{bPC}, \phi^{bPC})_i \approx (\eta^{bPC}, \phi^{bPC})_{i+1}$, i being the ith iteration. Jet candidates are now defined.

6. Since jet candidates may share some amount of transverse energy (see figure 4.2),

   - jet candidates are merged if the shared E_T is more than 75% of the less energetic E_T jet. Jet centroid is then recalculated.

[10] reconstructed position of the $p\bar{p}$ collision

[11] Cone radii used at CDF in other analysis are also 0.7 and 1.0

- if not, shared E_T is assigned to the closest jet candidate.

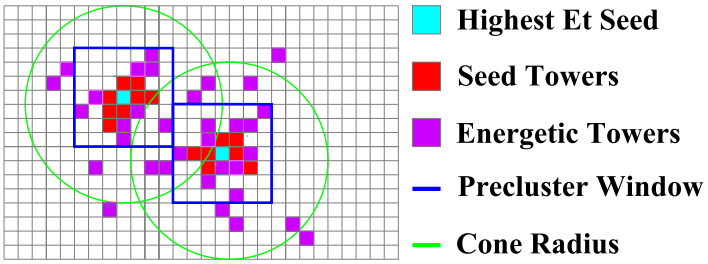

Figure 4.2: Jet reconstruction by the JETCLU cone algorithm. See text for significance of symbols.

A jet is univocally identified by this procedure. The four-momentum of the jet is computed from its towers [12]:

$$E = \sum_{i=1}^{N_{towers}} (E_i^{em} + E_i^{had}) \tag{4.4}$$

$$P_x = \sum_{i=1}^{N_{towers}} (E_i^{em} sin\theta_i^{em} + E_i^{had} sin\theta_i^{had}) cos\phi_i^{em} \tag{4.5}$$

$$P_y = \sum_{i=1}^{N_{towers}} (E_i^{em} sin\theta_i^{em} + E_i^{had} sin\theta_i^{had}) sin\phi_i^{em} \tag{4.6}$$

$$P_z = \sum_{i=1}^{N_{towers}} (E_i^{em} cos\theta_i^{em} + E_i^{had} cos\theta_i^{had}) \tag{4.7}$$

where x, y are the Cartesian coordinates in the transverse plane. In this scheme the towers are handled as massless particles. Therefore:

$$P = \sqrt{P_x^2 + P_y^2 + P_z^2} \tag{4.8}$$

$$P_T = \sqrt{P_x^2 + P_y^2} \tag{4.9}$$

$$E_T = E \times \frac{P_T}{P} \tag{4.10}$$

$$y = \frac{1}{2} ln \frac{E + p_z}{E - p_z} \tag{4.11}$$

[12] Two schemes have been defined at CDF: "Snowmass scheme" [32], and "E-scheme". The latter is used in this analysis.

$$\phi \;=\; arctan(\frac{P_y}{P_x}) \tag{4.12}$$

■ Jet Energy Corrections

The 4-momenta assigned to a jet (4.4-4.7) suffer for both detector inaccuracies and reconstruction algorithm imperfections. A set of corrections to the jet reconstructed energy ("raw energy") have been developed to overcome these limits.

These corrections are applied in a sequence of levels (of "L-levels") in order to correct for each bias independently.

- *η-Dependent correction (L1)*

  The L1 correction accounts for non-uniformities in calorimeter response along η. This correction is obtained by studying the P_T-balancing in djet events [36, 37]. Djet events are selected in order to have one jet ("trigger jet") in the $0.2 < |\eta| < 0.6$ region [13]. The other jet, called "probe jet", is free to span over the $|\eta| < 3$ region (see figure 4.3).

  The P_T balancing fraction is formed:

$$f_b \equiv \frac{\Delta P_T}{P_T^{ave}} = \frac{P_T^{probe} - P_T^{trigger}}{(P_T^{probe} + P_T^{trigger})/2} \tag{4.13}$$

  where P_T^{probe} and $P_T^{trigger}$ are the transverse momenta of the probe and the trigger jets respectively. f_b, is used to define the correction factor [14]:

$$\beta \equiv \frac{2+ < f_b >}{2- < f_b >} \tag{4.14}$$

  which is in average equal to $P_T^{probe}/P_T^{trigger}$ [15]. In figure 4.4 β is shown as a function of η.

  The L1 correction is thus defined as $f_{L1}(R = 0.4, E_T^{raw}, \eta) = 1/\beta(\eta)$.

  The error on f_{L1} varies from 0.5% to 7.5%.

[13] This region is far away from detector cracks and is expected to have a stable response.

[14] $< f_b >$ is the average in the η bin chosen for this study

[15] The definition 4.14 reduces the sensitivity of the measurement to the non-Gaussian tails which affect the ratio $P_T^{probe}/P_T^{trigger}$. See [36] for further details.

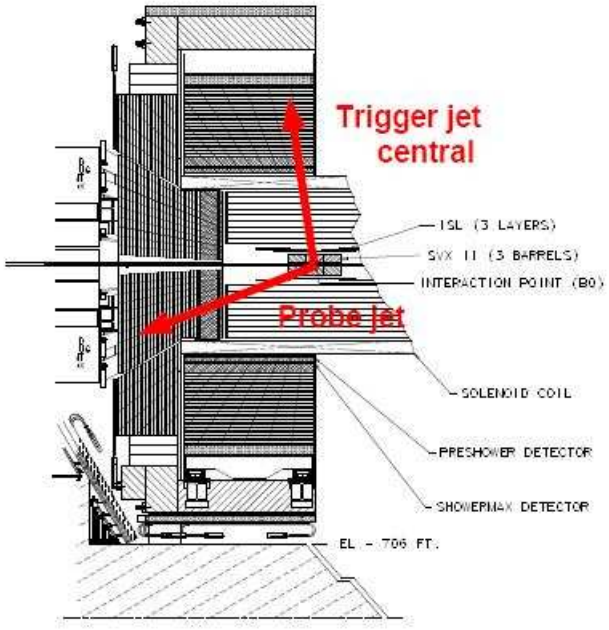

Figure 4.3: Schematic drawing of djet balancing

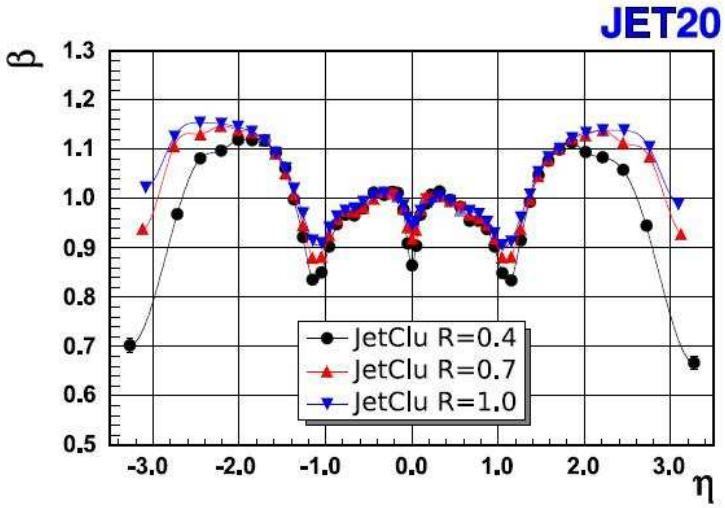

Figure 4.4: η-dependent energy scale correction factor for three different cone radii. JET20 sample comprises the events satisfying the trigger level 1 request to have at least one tower above 20 GeV.

- *Multiple Interaction Corrections (L4 [16])*

  Energy from additional interactions [17] may fall into the jet cone and affect its overall energy. A correction for this effect is extracted using a sample of minimum bias events [38]. A cone is drawn in random positions within the $0.1 < |\eta| < 0.7$ region and the corresponding E_T is measured. The E_T distribution is parametrized as a function of the number of vertexes of quality [18] (figure 4.5) in the event. This function is used as a luminosity-dependent correction factor.

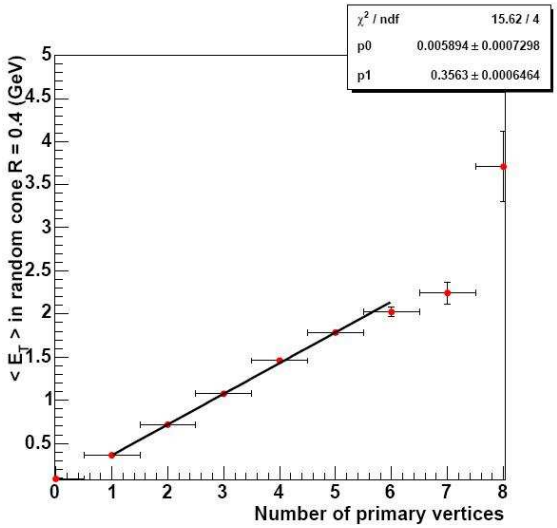

Figure 4.5: E_T in R=0.4 cone as function of the number of reconstructed primary vertexes in minimum bias events.

  The correction factor uncertainty, depending on luminosity, event topology and vertex reconstruction efficiency, is about 15%.

- *Absolute Energy Scale Corrections (L5)*

  While L1 and L4 accounted for corrections at calorimeter level, L5 steps back to particle level. The procedure used to estimate the L5 correction factor is as follow [39]:

[16] L2 and L3 have survived in the CDF jargon but are not used anymore. L2 was used in Run I to remove any time dependence of the calorimeter PMT's, L3 was used in a brief period during Run II to account for differences between MC and data observed in photon+jets events. These differences have been later cured.

[17] At the current maximum luminosity the average number of $p\overline{p}$ interactions per bunch crossing is about 5 (see chapter 2).

[18] Vertexes composed at least by two COT tracks.

- A large event sample is generated with the full CDF detector simulation.
- Jets are reconstructed at both calorimeter (C) and hadron (H) level [19]. P_T^C and P_T^H transverse momenta are assigned to C and H jets respectively.
- Each C jet is associated to a H jet if $\Delta R < 0.1$.
- The expected probability $\mathscr{P}(P_T^C|P_T^H)$ to measure P_T^C, given P_T^H, is extracted.
- $Max_{P_T^H}(\mathscr{P}(P_T^C|P_T^H))$ is taken as correction factor (see figure 4.6).

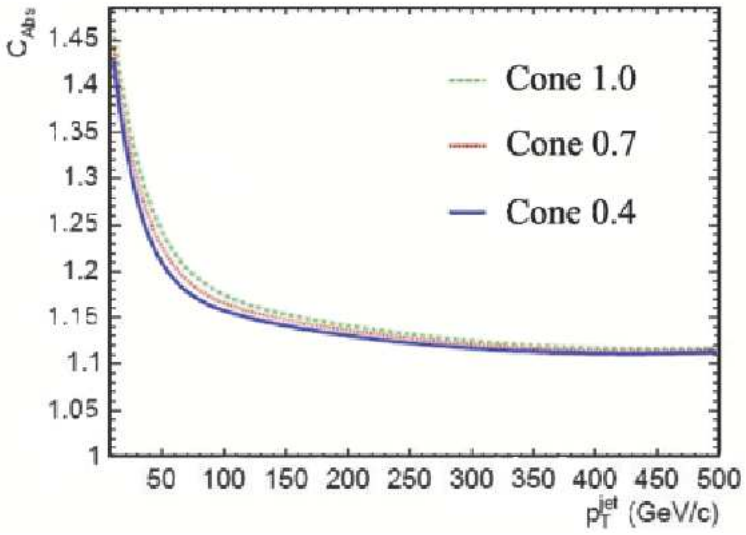

Figure 4.6: Absolute corrections for different cone sizes as a function of jet transverse momentum.

L5 correction factor uncertainty is about 1.5-3%. Main contributions to this uncertainty arise from calorimeter simulation, modeling of fragmentation and stability of the calorimeter.

- *Underlying Event (L6) and Out-Of-Cone (L7) Corrections*

  Reconstructed jet energies in hard $p\bar{p}$ interactions may contain contributions by particles created by soft spectator interactions or by gluons from initial state radiation. These contributions are called underlying event. On the other hand a fraction of the parton energy may be lost outside the jet

[19] This procedure does not distinguish between jets from gluons or quarks.

cone because of final state gluon radiation, fragmentation effects or low P_T particles bending in the magnet field.

L6 and L7 purpose is to correct for these effects. L6 and L7 correction factors are obtained from simulated djet processes by applying the same procedure as for the L5 corrections (more details are described in [39]).

The correction factors described above are applied to the raw energy E_T^{raw} [20] as follows (in this analysis a cone radius $R = 0.4$ is used):

$$E_T(E_T^{raw}, \eta) = (E_T^{raw} \cdot f_{L1}(E_T^{raw}, \eta) - A_{L4}) f_{L5}(E_T^{raw}) \; - \; A_{L6} \; + A_{L7}(E_T^{raw}) \quad (4.15)$$

where $f_{L1},\ A_{L4},\ f_{L5},\ A_{L6},\ A_{L7}$ are the correction factors for each indicated level.

Formula 4.15 define the **Jet Energy Scale (JES)** [21].

4.1.5 Missing Transverse Energy

Neutrinos cannot be detected at CDF. In $t\bar{t}$ dilepton events some indirect information on the two neutrinos ($\not\!\!E_T$ or MET) can be derived by the momentum conservation in the transverse plane assuming the transverse momentum of the interacting parton system to be zero. This is shown by equation 4.16 [22]:

$$\overrightarrow{\not\!\!E_{T_{raw}}} = -\sum_i E_i sin\theta_i \hat{n}_i \quad (4.16)$$

where the sum is over all towers with a deposited energy more than 0.1 GeV. $\hat{n}$ is the transverse versor pointing from the CDF geometrical center to the tower, whose polar angle is θ.

$\not\!\!E_{T_{raw}}$ must be corrected for various reasons (muons P_T, interaction vertex position, etc.), which will be discussed in section 4.4.2.

4.2 $t\bar{t}$ Dilepton Channel Signature

In 1.3 the three final state channels of the $t\bar{t}$ system have been briefly described.

[20] Correction factors can be equally applied to the raw transverse momentum P_T^{raw} and E_T^{raw}, since both P_T^{raw} and E_T^{raw} rely on calorimeter information.

[21] Some correction levels can sometimes be skipped. Later in the Chapter L4, L6 and L7 will be neglected in event selection.

[22] The assumption $m_\nu = 0$ has been made

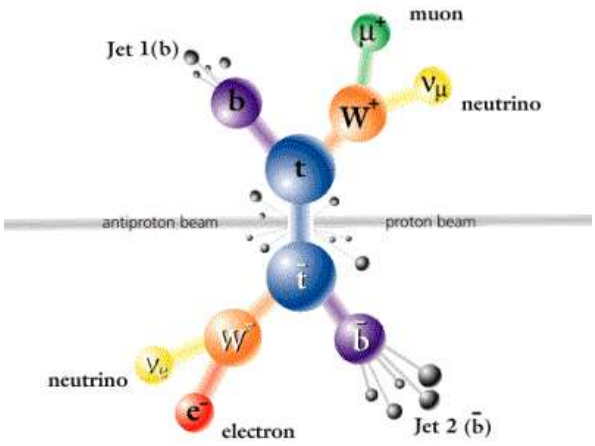

Figure 4.7: a $t\bar{t}$ dilepton event at Tevatron

The dilepton final state signature (see figure 4.7) consists of:

- Two large transverse momentum leptons.
- Large $\not\!\!E_T$, which is assumed to be due to the two-neutrinos in the W decays.
- Two jets assumed to be originated from b-quarks.
- The possibility of additional jets from initial and final state radiation.

4.3 Backgrounds

Background processes (BG) can be divided into two groups: physical and non-physical background. The former is due to non-$t\bar{t}$ events simulating top dilepton events because of similar kinematic features, while the latter is caused by detector imperfections.

The main expected BG are discussed here.

4.3.1 Physical Backgrounds

▶ *Diboson*

this BG includes WW, WZ, ZZ [23] events in which respectively

[23] $p\bar{p} \rightarrow WW, WZ, ZZ + X$, X being any additional system.

- W's decay leptonically
- Z decays leptonically and W hadronically
- One Z decays leptonically and the other hadronically

Those events can cause physical background if they happen to have two or more jets from initial or state radiation. In the case of WZ and ZZ, missing E_T may occur from mis-measured jets or leptons.

▶ *Drell-Yan*

Charged lepton pairs generated by Z boson or virtual photon [24] can simulate the dilepton $t\bar{t}$ signature. A minimum number of two jets and $\not\!\!E_T$ must occur because of the same reasons as in WZ and ZZ events.

4.3.2 Non-Physical Backgrounds

▶ *Fakes*

Fakes are due to partons being confused with leptons. At detector level this effect can be described as jet faking the signature of a lepton. W+jets and $t\bar{t}$ single lepton (1.3) processes are expected to be the main contributions to Fakes background.

4.4 Event Selection

At CDF two selections have been performed so far in the dilepton channel: the dilepton (DIL) and the lepton+track (LTRK) selections [25]. Most selection criteria are quite similar (with a large overlap, see figure 4.8) in both algorithms except for the requirements defining the track [26] lepton.

Since **LTRK** selection has been used for this analysis, LTRK is going to be described in this section. DIL selection will be described in appendix B.

4.4.1 Data Samples

In order to be included in the data sample, runs in the datasets (datasets are listed in table E.1) are required to belong to the "good run lists", which take into account the optimal status of the beam and of sub-detectors. Good run lists used in this analysis are reported in table E.2 [27].

[24] $p\bar{p} \rightarrow Z/\gamma^* + X$
[25] CDF jargon
[26] See later in the Chapter for more details
[27] Each good run list in table E.2 is referred to the triggered lepton in the event.

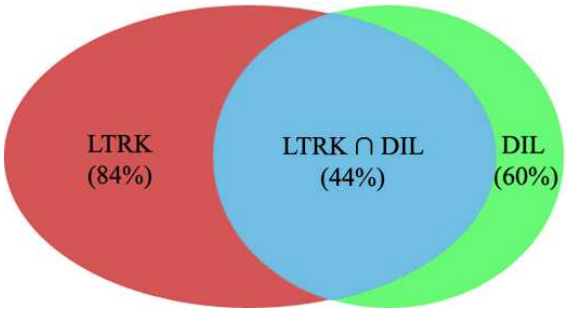

Figure 4.8: LTRK and DIL samples represented as ensembles. LTRK, DIL and the overlap between LTRK and DIL are normalized to the union of DIL and LTRK (events in either LTRK or DIL). LTRK scores 84%, DIL scores 60%. Shared events by the two samples (light blue) are 44%. Percentages refer to an integrated luminosity of $\sim 350\ pb^{-1}$ [45].

4.4.2 Requirements on the Physical Objects

Objects in candidate $t\bar{t}$ dilepton events must satisfy the following requirements [28].

▲ Tight Lepton (tl)

Four type of leptons are considered in the LTRK analysis: CEM, CMUP, CMX and PHX [29]. A tight lepton must have $E_T \geq 20\ GeV$ ($P_T \geq 20\ GeV$ for muons), associated track $|z_0| < 60\ cm$ and satisfy the following requirements defined according to the lepton type.

- *CEM*:

  --→ Associated track must have $P_T > 10\ GeV$ and be well-measured (≥ 3 axial and ≥ 2 stereo superlayers with ≥ 5 hits in the COT).

  --→ $E/p < 2$ if $E_T < 50\ GeV$ (to reject hadrons accompanied by photons)

  --→ $E_{had}/E_{em} < 0.055 + 0.00045 \times E$

  --→ Shower profile as determined in electron test beam measurements

- *Muon*:

  --→ A well-measured track with acceptable track reconstruction probability

  --→ $E_{had} < 6 + max(0, 0.028 \times (p - 100))GeV$

[28] Only main cuts are reported below. The full list is reported in [40, 41]

[29] Leptons are labelled with the name of the calorimeter they hit. The exception is the PHX electron, which is identified using the plug calorimeter info (see 4.1.1)

--→ $E_{em} < 2 + max(0, 0.0115 \times (p - 100))GeV$

--→ Track $|d_0| < 0.02\ cm$ if track is reconstructed with silicon hits; $|d_0| < 0.2\ cm$ otherwise.

--→ if the muon is a CMUP the extrapolated track must be matched to both CMU and CMP stubs. If the muon is a CMX the exit radius of the COT track must be $\geq 140\ cm$.

- *PHX*:

--→ the track must include at least 3 silicon hits

--→ $E_{had}/E_{em} < 0.05$

--→ the shower in PES must be at $1.2 < |\eta| < 2$

--→ if $E_T \leq 100\ GeV$ the shower profile must match the one obtained in test beam

--→ the quality χ^2 of the PEM cluster match to the track must be ≤ 10

--→ the distance between PEM cluster and PES shower must be $\leq 3\ cm$

Tight electrons (muons) must be isolated, i.e. the ratio between the additional E_T (P_T) in the R=0.4 cone centered on the electron (muon) and the E_T (P_T) of the electromagnetic associated cluster (COT associated tracks) must be < 0.1.

Tight leptons must be triggered. They are selected with the following decreasing priority: CEM, CMUP, CMX, PHX. If two tight leptons of the same type are found the highest-E_T is chosen as first.

▲ Track Lepton (trkl)

A track lepton (also called second or loose lepton) is a reconstructed track (see 4.1.1), which satisfies the following requirements:

▷ ≥ 24 axial and ≥ 20 stereo COT hits

▷ ≥ 3 available silicon hits if its trajectory indicates that it should

▷ $|{d_0}^{corr}| < 250\ \mu m$, where ${d_0}^{corr}$ is the beam - corrected track impact parameter [30]

▷ An acceptable track reconstruction probability

[30] The impact parameter is corrected for the displacement of the reconstructed primary vertex with respect to the CDF geometrical center. The beam inclination relative to the ideal line induces a z-dependent correction to the impact parameter. See [34] for further details

- ▷ $|z_0^{tl} - z_0^{trkl}| < 5\ cm$, where z_0^{tl} and z_0^{tl} are the z coordinates of the closest approach to the z-axis for tl and trkl respectively. They are measured with respect to the CDF geometrical center.
- ▷ the track isolation, defined as $\frac{\sum P_T}{\sum P_T + P_T^{trkl}} < 0.1$ where the sum runs over all additional $P_T > 0.5\ GeV$ tracks within a R=0.4 cone centered on the trkl [31].

▲ Jets

Jets must satisfy the following requirements

1. $E_T^{corr} > 20\ GeV$, where E_T^{corr} is the raw jet transverse energy (4.10) corrected up to L5 without L4 corrections (see 4.1.4).
2. $|\eta| < 2$

Jets are also required to be fully outside the R=0.4 cones whose axis are the tl and trkl directions.

▲ Missing E_T

Starting from 4.16, $\not\!\!E_T$ is corrected as follows: [32].

- The distance of the interaction vertex from the geometrical vertex. $\not\!\!E_T$ is recalculated with all vectors in 4.16 as having the origin in the interaction vertex.
- If the tl is a muon its momentum is used in building $\not\!\!E_T$ rather than the calorimeter energy.
- If the transverse energy in the track tower is much less than the track P_T, the track P_T rather than the calorimeter energy is used in building $\not\!\!E_T$.
- If a track within a jet cone has a transverse momentum larger than the jet transverse energy, the track P_T rather than the jet E_T is used in building $\not\!\!E_T$.
- L1, L5 jet energy corrections. $\not\!\!E_T$ is recalculated accounting for the corrected transverse energy of the jets.

After corrections, missing E_T is required to be $> 25\ GeV$.

[31] The track isolation is defined in a slightly different way with respect the tight leptons
[32] More details are given in [34].

4.4.3 Vetoes

Several topological vetoes are implemented in the LTRK selection in order to reduce the impact of BG.

1. Cosmic Removal: Muons tagged as cosmic rays are rejected. The distance between the muon track and the interaction point, timing info from COT and calorimeters are used for the tagging [33].

2. Conversion Removal: Electrons coming from the conversion of photons to e^+e^- pair are rejected.

3. Z veto/$\not\!\!E_T$: if the invariant mass of the system tl-trkl [34] is within the Z mass window ([76,106] GeV/c^2) $\not\!\!E_T > 40\ GeV$ is required. This veto aims to reject the Drell-Yan events in which $\not\!\!E_T$ arises from mis-measured jets or leptons.

4. $\Delta\phi(\not\!\!E_T - trkl)$: $\not\!\!E_T$ parallel or antiparallel to the trkl may be faked by the mis-measurements of leptons in Drell-Yan events. Therefore a minimum angular separation between the $\not\!\!E_T$ direction and trkl is required ($5^\circ < \Delta\phi(\not\!\!E_T - trkl) < 175^\circ$)

5. $\Delta\phi(\not\!\!E_T - tl)$: events with $\Delta\phi(\not\!\!E_T - tl) < 5^\circ$ are rejected. The main reason of this cut is the same as above. An additional motivation is to reduce the acceptance for $Z/\gamma^* \rightarrow \mu\mu$ as electron + trkl events, where very high-P_T muons emit a bremsstrahlung photon [47].

6. $\Delta\phi(Jet - \not\!\!E_T)$: events with $\Delta\phi(jet, \not\!\!E_T) < 25^\circ$ are rejected. All jets with $E_T > 10\ GeV$ are considered. This veto is dropped if $\not\!\!E_T > 50\ GeV$. The purpose of this cut is to reject events with mis-measured jets.

7. Opposite-Sign (OS): An opposite charge between tl and trkl is required. This aims at reducing fakes (see 4.3.2) contributions to the selected sample since fakes have the same chance for both charge signs.

The requirement on physical objects (4.4.2) and vetoes described above are the core of the LTRK selection which is applied on data.

In order to avoid double-counting fakes from Diboson and $t\bar{t}$ single lepton events one more "technical" cut with respect to data is implemented for the following categories:

- Signal, Diboson (WW): tl and trkl are required to come from W's

[33] Further details are available at [35].

[34] trkl is assumed to be of the same flavor as tl.

- Diboson (WZ): tl and trkl are required to come from the Z.

Fakes selection, which is performed on data, is described in appendix C.

Chapter 5

Top Quark Mass Measurement Methods

The approach to measure the top mass adopted in this analysis is the **Template Method (TM)**: the experimental distributions of variables sensitive to the top mass are compared with the ones built from simulated events ("templates") (section 5.1). Several sensitive variables, such as the jet transverse energies, the decay length of the tagged b-quarks in the transverse plane (section 5.2.3), the top reconstructed mass M_t^{reco} [1] can be chosen to implement a TM.

Other two approaches, the Matrix Element (ME) and the Ideogram methods, are used at CDF. ME estimated the top mass by maximizing the M_{top}-dependent probability to have the same configuration of final products in the event as it is observed. The ideogram method combines TM and ME. Details are available in [80, 81].

A number of TM methods will be discussed in this chapter.

5.1 The Dilepton Channel

The Dilepton channel has two neutrinos in the final state (see figure 4.7): this means that 6 constraints involving the neutrinos cannot be imposed by the conservation equations.

As can be understood from figure 4.7 the $t\bar{t}$ decay chain involves 10 parameters, out of which only 4 are measurable. Moreover the $t\bar{t}$ system decay has 5 vertexes

and for each of them we can write the conservation relations as in equation 5.1 [2]

[1] M_t^{reco} is the mass reconstructed by exploiting the information of the top decay products in the event. See next Chapter or [45] for details.

[2] The overall energy and longitudinal momentum conservation cannot be used, because small

$$\overrightarrow{P}_T^{\,t} + \overrightarrow{P}_T^{\,\bar{t}} = -\overrightarrow{E}_T^{\,UnderEv} \quad (5.1)$$

$$P_\mu^b + P_\mu^{W^+} = P_\mu^t \quad (5.2)$$

$$P_\mu^{\bar{b}} + P_\mu^{W^-} = P_\mu^{\bar{t}} \quad (5.3)$$

$$P_\mu^{l^+} + P_\mu^{\nu} = P_\mu^{W^+} \quad (5.4)$$

$$P_\mu^{l^-} + P_\mu^{\bar{\nu}} = P_\mu^{W^-} \quad (5.5)$$

Where $E_T^{UnderEv}$ is the underlying event transverse energy (See 4.1.4 for the definition of underlying event). P_μ indicates the 4-momentum vector.

In the system above neutrinos can be assumed to have zero mass, W mass is known ($M_W = 80.4\ GeV/c^2$), top and antitop masses are assumed to be equal to each other.

Finally, we have 23 equations out to 24 unknown (see Appendix A): therefore the system is underconstrained, a so called *-1C kinematics*. One must assume as known at least one more parameter in order to be able to reconstruct kinematically the events and find for each event a value for the mass of the top. In CDF three analysis have been developed with three different parameters assumed as known: the "Full kinematic analysis", the "Neutrino weighting algorithm" and the "Neutrino ϕ weighting method".

Once the best top mass is found in each event, the three methods share the template method for reconstructing the top mass. The experimental mass distribution is fit in terms of a likelihood function which is a mix of a background template and of a top-mass dependent signal template. The templates are probability density functions (p.d.f's) obtained from simulated events. The total likelihood function takes the form [3]:

$$\mathscr{L}(M_{top}) \equiv \mathscr{L}_{shape}(M_{top}) \times \mathscr{L}_{backgr} \quad (5.6)$$

with

$$\mathscr{L}_{shape}(M_{top}) \equiv \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \prod_{n=1}^{N} \frac{n_s \cdot f_s(m_n|M_{top}) + n_b \cdot f_b(m_n)}{n_s+n_b} \quad (5.7)$$

$$\mathscr{L}_{backgr} \equiv e^{\left(\frac{-(n_b-n_b^{exp})^2}{2\sigma^2_{n_b^{exp}}}\right)} \quad (5.8)$$

f_s, f_b are the signal and background, parametrized as continuous functions and weighted by an expected number of signal (n_s) and background (n_b) events, estimated by the fit. m_n is the top mass reconstructed in the nth event, N is the total

angle particles are lost by the CDF detector around the beampipe.

[3] The likelihood 5.6 is used by the Full kinematic and Neutrino algorithm method. One more term is added in the likelihood by the Neutrino ϕ weighting algorithm (see Chapter 6 for details)

number of observed events in data.
n_b^{exp}, $\sigma^2_{n_b^{exp}}$ are obtained from a separate measurement of the cross section [47].

We will give more details about how to obtain the likelihood function in the next chapter when we will discuss our template method. Now we will describe briefly the Full kinematic and Neutrino algorithm method. The Neutrino ϕ weighting algorithm, which we have adopted in this analysis, will be discussed in detail in the next Chapter.

5.1.1 Full kinematic analysis (KIN)

The Full kinematic method reconstructs the top mass by assuming as known the total longitudinal momentum of the $t\overline{t}$ system, $p_z^{t\overline{t}}$.
Although the distribution should be centered around zero for any top mass, its width might depend on top mass. However, simulations show that this quantity is a zero-centered Gaussian with a width $\sigma = 220\ GeV/c^2$, which is insensitive to M_t in the range $M_t = 140 - 200\ GeV/c^2$. In the single lepton channel this expectation can be checked with data, since $p_z^{t\overline{t}}$ can be measured. The MonteCarlo analysis are validated by comparing the distribution obtained in simulations with data in lepton+jets channel as it is shown in figure 5.1 [4].

In figure 5.2 we show the $P_z^{t\overline{t}}$ distribution for dileptonic events: the width is slightly reduced because of the limited η coverage of the CDF detector.
So KIN method adds to the system 5.1 the equation

$$P_z{}^t + P_z{}^{\overline{t}} = P_{t\overline{t}}^z \tag{5.9}$$

Where $P_{t\overline{t}}^z$ is randomly drawn from Gaussian with zero mean and 195 GeV/c^2 width. Thanks the equation 5.9 the system 5.1 is now solvable.
Since it is composed on non linear equations, this methods uses the iterative numerical procedure called *Newton's method*. Generally speaking this method starts from a system expressed by the vectorial equation $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ and requires an initial guess $\mathbf{x} = \mathbf{x_0}$ for the unknown vector which should be as much close as possible to the true root. The local Jacobian is calculated to allow the procedure to be repeated with a new value according to the recursive equation

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \frac{\mathbf{F}(\mathbf{x}^n)}{J_{\mathbf{F}}(\mathbf{x}^n)} \tag{5.10}$$

until a satisfactory [5] solution is found.

[4] Of course the $p_z^{t\overline{t}}$ distribution is the same in any top decay channel.
[5] $|\Delta x^n| = |x^{n+1} - x^n| << 1$

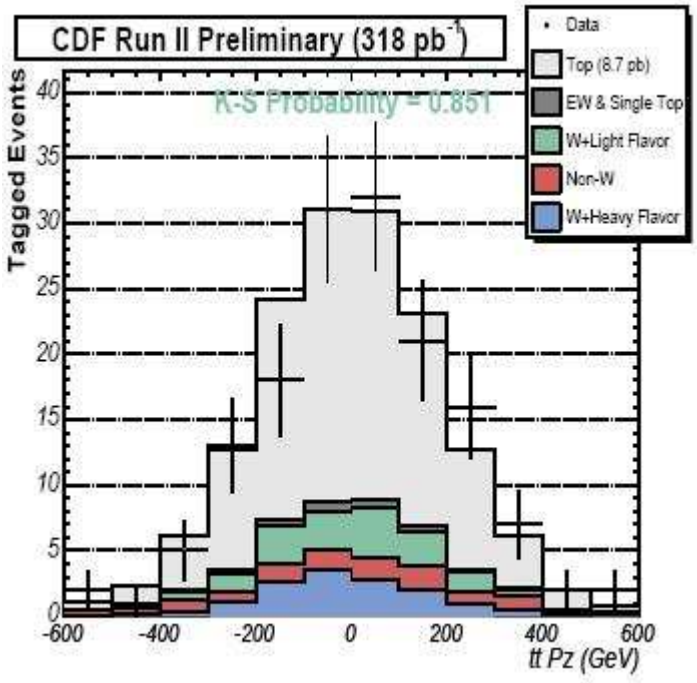

Figure 5.1: Longitudinal momentum of the $t\bar{t}$ system in lepton+jets data, compared to simulations.

This method should converge after a few iterations if the initial guess is not so far from the true root. If no solution is found the assumption on M_t and M_W are replaced by the following equations:

$$M_t \;=\; M_{\overline{t}} \pm 2.0\ GeV/c^2; \quad M_{W^\pm} \;=\; 80.4 \;\pm\; 3.0\ GeV/c^2 \tag{5.11}$$

to allow for the finite precision with which we know M_t and M_W.

The full constrained system, schematized as $F(\mathbf{x})$, can be simplified into a set of three equations shown below:

$$f_1(p_x^{\nu_1}, p_y^{\nu_1}, p_z^{\nu_1}) \;=\; (E_{l_1} + E_{\nu_1})^2 - (\overrightarrow{P}_{l_1} + \overrightarrow{P}_{\nu_1})^2 - m_W^2 = 0 \tag{5.12}$$

$$f_2(p_x^{\nu_1}, p_y^{\nu_1}, p_z^{\nu_1}) \;=\; (E_{l_2} + E_{\nu_2})^2 - (\overrightarrow{P}_{l_2} + \overrightarrow{P}_{\nu_2})^2 - m_W^2 = 0 \tag{5.13}$$

$$\begin{aligned} f_3(p_x^{\nu_1}, p_y^{\nu_1}, p_z^{\nu_1}) \;=\; & (E_{l_1} + E_{\nu_1} + E_{b_1})^2 - (\overrightarrow{P}_{l_1} + \overrightarrow{P}_{\nu_1} + \overrightarrow{P}_{b_1})^2 - \\ & (E_{l_2} + E_{\nu_2} + E_{b_2})^2 + (\overrightarrow{P}_{l_2} + \overrightarrow{P}_{\nu_2} + \overrightarrow{P}_{b_2})^2 = 0 \end{aligned} \tag{5.14}$$

These equations are easily solvable if the momentum of the second neutrino is expressed in terms of the momentum of the first one:

$$p_x^{\nu_2} \;=\; \not{E}_{T_x} \;-\; p_x^{\nu_1} \tag{5.15}$$

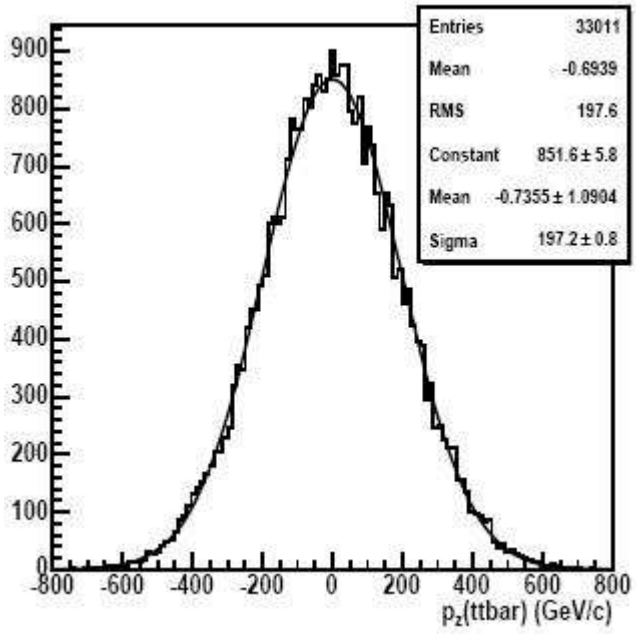

Figure 5.2: $P_z^{t\bar{t}}$ distribution for events passing DIL selection. These events are from MC sample with $M_t = 175\ GeV/c^2$. The continuous curve is a Gaussian fit to the distribution.

$$p_y^{\nu_2} = \not{E}_{T_y} - p_y^{\nu_1} \tag{5.16}$$

$$p_z^{\nu_2} = p_z^{t\bar{t}} - (p_{b_1 z} + p_{l_1 z} + p_{\nu_1 z} + p_{b_2 z} + p_{l_2 z}) \tag{5.17}$$

$$\tag{5.18}$$

Applying Newton's method we determine the first of two pairs of solutions for the two neutrinos [6].
Since there are two possible associations among b-quark jets and leptons, we have in total 8 solutions per each step of Newton method iteration.
In order to allow for the range of possible $p_z^{t\bar{t}}$ (shown in figure 5.2), as well as the finite resolutions of measured momenta and MET, in a current analysis [51] the above procedure is repeated 10^4 times per each possible solution. Each time, $p_z^{t\bar{t}}$ is randomly drawn from a Gaussian with zero mean and 195 GeV/c width. The jet Energy and MET are similarly smeared by Gaussian functions according to their estimated resolutions, while the measured jet angles and the lepton momenta which are measured with small errors are assumed to be exactly known.

[6] Notice that we are solving quadratic equations for both neutrinos, so we have 2×2 possible (p^{ν_1}, p^{ν_2}) solutions

Once we have all the kinematical variables we can calculate the $t(\overline{t}) \rightarrow W^{\pm} + b(\overline{b}) \rightarrow l^{\pm} + \nu(\overline{\nu}) + 2jets$ invariant mass, by using jet and lepton information and the obtained solutions for neutrinos. Notice that we can have up to 10^4 invariant masses per each of the 8 solutions. We choose the most probable value (MPV) of a spline fit to the mass distribution as the "raw top quark mass". The mass indicated by an event must be selected among these eight masses.
These solutions are split into two classes according to the lepton-to-jet pairing and in each class the solution that minimizes the invariant mass of the $t\overline{t}$ system is chosen as the favored one [7].
The mass distributions of the two solutions are compared (see as example figure 5.3) and the solution with the largest number of entries provides the measured top mass.

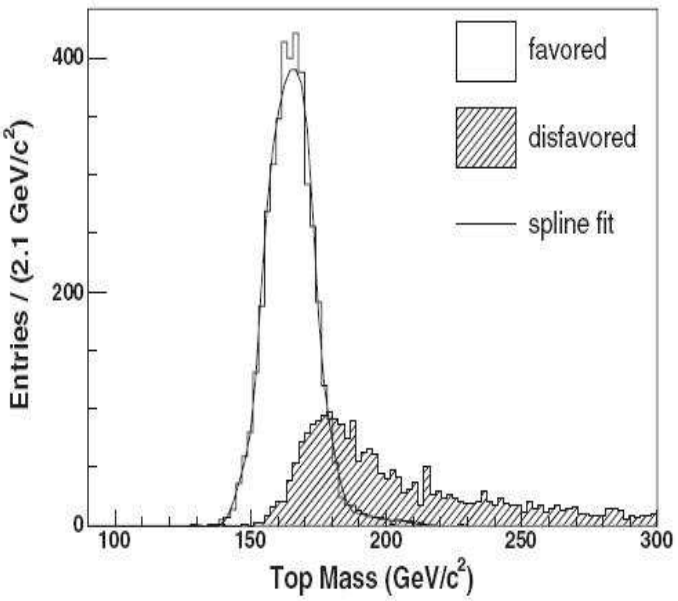

Figure 5.3: Smeared top quark mass distributions of the favoured and disfavoured lepton+jet pairings from the KIN method applied to a *HERWIG* Monte Carlo $t\overline{t}$ event with $M_t = 175\ GeV/c^2$. Also shown for the favoured pair is the spline fit used to select the top quark mass

In the above procedure events with less than 1000 entries in the invariant mass histograms are not taken into account.

[7] The efficiency of this method has been estimated to be 84% on MC generated events with $M_t = 178\ GeV$ [51]

5.1.2 Neutrino weighting algorithm (NWA)

The second method for estimating the top quark mass from dilepton events is the Neutrino weighting algorithm. For the time being the method has been applied by CDF to a 360 pb^{-1} LTRK event sample.

The strategy of this method is to solve for the neutrino and antineutrino momenta, independently of the measured missing energy, by constraining the system 5.1÷5.5 with additional assumptions on the $t\overline{t}$ decay. The neutrino/antineutrino configurations are weighted with a function $\not\!\!E_T$ (and m_t !) dependent function and used to create a probability distribution.

The NWA weight function is constructed according to the following procedure.

- We assume values for the top quark and W boson masses, neutrino and antineutrino pseudo-rapidities, and the lepton-jet.
- We apply the conservation equations to the top quark decay and obtain up to two possible equations for the neutrino
- We repeat this procedure on the antitop decay
- To each of the four obtained pairs of (ν, $\overline{\nu}$) solutions is assigned a probability (weight) according to the measured $\not\!\!E_x$ and $\not\!\!E_y$ and their uncertainties σ_x and σ_y respectively. The formula for the weight is shown below:

$$w_i \;=\; e^{-\frac{(\not\!\!E_x \;-\; p_x^{\nu} \;-\; p_x^{\overline{\nu}})^2}{2\sigma_x^2}} \;\times\; e^{-\frac{(\not\!\!E_y \;-\; p_y^{\nu} \;-\; p_y^{\overline{\nu}})^2}{2\sigma_y^2}} \tag{5.19}$$

  We use $\sigma_x \;=\; \sigma_y \;= 15\ GeV$, which is obtained from $t\overline{t}$ generated events with $m_t \;=\; 178\ GeV/c^2$ [8].

- Since (ν, $\overline{\nu}$) 4 solutions have *a priori* equal probability, we sum the four contributions:

$$w(m_t,\ \eta_\nu,\ \eta_{\overline{\nu}},\ l-jet) \;=\; \sum_{i=1}^{4} w_i \tag{5.20}$$

- Since the neutrino polar angles are a-priori equally likely, we repeat the above steps for many possible (η_ν, $\eta_{\overline{\nu}}$) pairs.
  MC studies indicates that neutrino $\eta's$ are uncorrelated (as seen in Figure 5.4) and follow a Gaussian distribution centered at zero with width of about 1.

[8] In practice, however, the performance of the algorithm is insensitive to the particular choice of the $\not\!\!E_T$ resolution

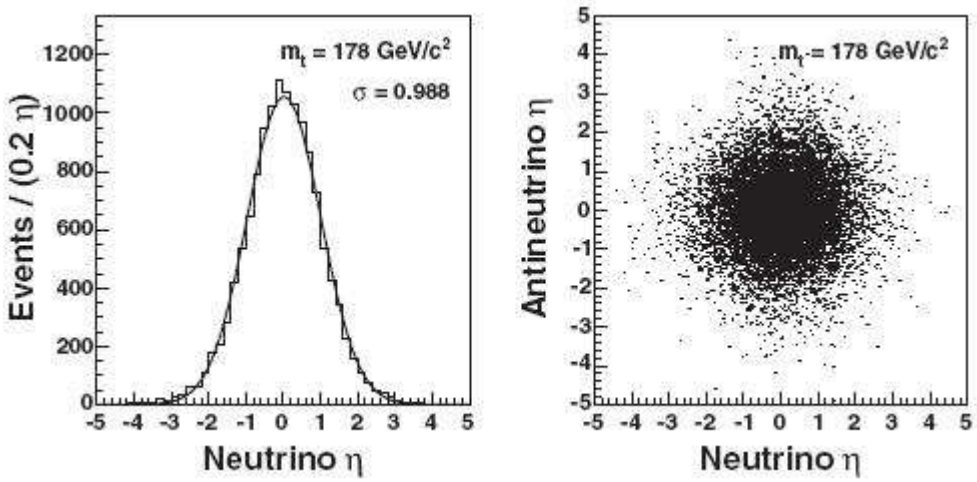

Figure 5.4: Neutrino η distribution with Gaussian fit (left) and neutrino Vs antineutrino η's (right) from a *HERWIG* $t\bar{t}$ sample with $M_t = 178\ GeV/c^2$

The neutrino η width varies little with top quark mass (see Figure 5.5) and a constant width of 0.9888 is assumed.

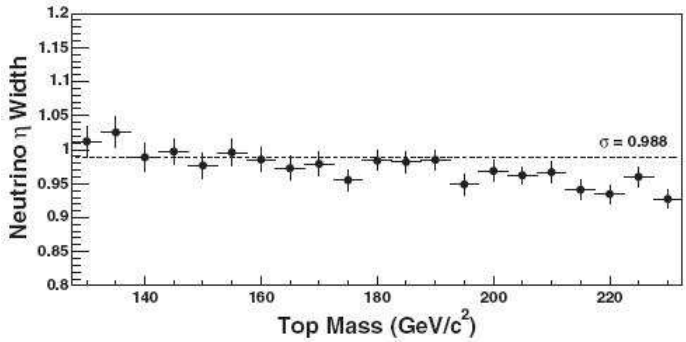

Figure 5.5: η width as a function of generated top quark mass, compared with fit value at $M_t = 178\ GeV/c^2$ (horizontal line)

To ensure symmetry and smoothness, we scan the neutrino η distributions from -3 to 3 in steps of 0.1 and we assign to each step a probability of occurrence $P(\eta_\nu,\ \eta_{\overline{\nu}})$ derived from a zero centered Gaussian with a 0.988 width. Each trial $\eta_\nu,\ \eta_{\overline{\nu}}$ will contribute to the event according to its weight,

w, and probability of occurrence $P(\eta_\nu,\ \eta_{\overline{\nu}})$, as shown in the equation below

$$w(m_t, l-jet) \ = \sum_{\eta_\nu,\ \eta_{\overline{\nu}}} P(\eta_\nu,\ \eta_{\overline{\nu}}) \times w(m_t,\ \eta_\nu,\ \eta_{\overline{\nu}},\ l-jet) \qquad (5.21)$$

- Since no b-tagging is used we sum over the two possible pairings and get the NWA weight function as :

$$W(m_t) \ = \sum_{l^+-jet_1}^{l^+-jet_2} w(m_t, l-jet) \qquad (5.22)$$

We choose as top mass per event the **MPV value of the weight distribution** (see figure 5.6).

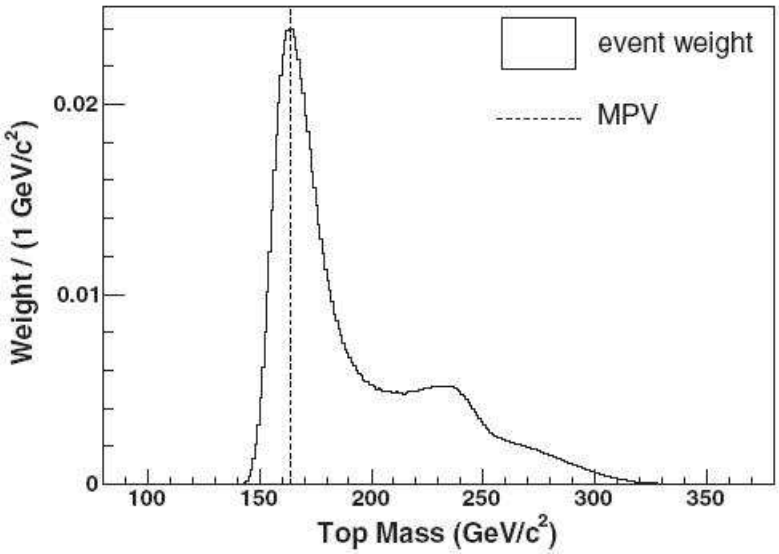

Figure 5.6: NWA weight distribution as a function of the assumed top quark mass (from equation 5.22) for a *HERWIG* Monte Carlo $t\overline{t}$ event with $M_t \ = \ 170\ GeV/c^2$. The vertical line shows that the most probable value (MPV) is close to 170 GeV/c^2.

In the next section we are going to give a brief description on how to measure the top mass in the single lepton and in the all-hadronic channels.

5.2 Lepton plus jets channel

The lepton plus jets channel (see section 1.3) has provided up to now the most precise measurements of the top quark mass. This is because it has the best compromise between background contamination (signal to noise ratio of the order of 1/1, which becomes as high as 20/1 if b-tagging is used) and branching ratio (BR), which is $\approx$ 30%.
In this channel we have in the final state only one neutrino, so we can assign the full missing transverse energy to it. The kinematic equations are over constrained (2-C kinematics) [9]. The only ambiguity not solved by the conservation equations is the sign of the longitudinal component of neutrino momentum (see previous section) and the pairings of jets to light or b quark jets, which can only in part be solved if b-tagging is used. We have respectively 24 combinations if b-tagging algorithm is not used (pretag sample), 12 if only one b-jet quark is tagged and 4 if we have two tagged b-jets [10].

5.2.1 Template method with *in situ* $W \to jj$ *calibration*

Since in the single lepton channel the kinematic is over-constrained and because of two jets attributed to W decays, this method offers the possibility of correcting the jet energy scale to fit at best the W mass. This allows to revise the calibration of the jet energy scale and to reduce the main source of systematic uncertainty that affects the final top mass estimator.

The signal templates $f_s(M_t, \Delta JES)$ are derived from simulations as a function of the top mass M_t and of the correction to the input jet energy scale ΔJES [11]. The background template f_s is independent of M_t and is assumed to be independent of ΔJES.

The final likelihood has the form:

$$\mathscr{L}(M_{top},\ \Delta_{JES}) \ = \ \mathscr{L}_{shape} \times \mathscr{L}^{m_t^{jj}}_{shape} \times \ \mathscr{L}_{backgr} \tag{5.23}$$

where

$$\mathscr{L}^{m_t^{jj}}_{shape} \ \equiv \ \prod_{n=1}^{N\times C_i} \frac{n_s \times f_s(m_{jj,n}|M_{top}, \Delta_{JES}) + n_b \times f_b(m_{jj,n})}{n_s + n_b} \tag{5.24}$$

[9] three more degrees of freedom than in the dilepton channel are constrained

[10] $24 = \frac{2\times 2\times 4!}{2}$ where in the numerator 4! are all the possible jet permutations, a factor 2 is because of neutrino $P-z$ uncertainty and another is because we do not know which is the leptonical decaying W. The factor 2 in the denominator is because a permutation on light jets does not change the likelihood. Combinations for 1-tag and 2-tags are calculated in the same way

[11] the MC samples are generated with fixed M_t and Δ_{JES} (Section 4.1.4)

$m_{jj,n}$ being the djet invariant mass for the nth event and C_i the number of the possible djet combinations. The other factors are given by 5.7, 5.8.
Notice that the likelihood function 5.23 has two free parameters (M_{top}, Δ_{JES}) to be determined by the minimization.

5.2.2 Multivariate template method

This method is very similar to the one already discussed in 5.2.1, since it uses the same sensitive variable M_t^{reco} and the same procedure to reconstruct the top mass per event.
However, it tries a number of improvements in comparison with the previous discussed method:

1. Reducing the statistical uncertainty by estimating the probability of picking up the correct jet permutation up on event-by-event basis and re-weighting the events accordingly.

2. Improving S/N by using a number of kinematic variables to distinguish $t\overline{t}$ events from background ones.

The applied procedure loops over seven different W masses, taking into account the Breit-Wigner distribution for weighting each iteration.
One seeks a set of "good" variables which increase the sensitivity to the top quark mass and improve the discrimination between signal and background. To look for those we compare the probability density distributions for a number of variables in signal and background using $W + 4jets$ sample. It is found that kinematic energy-dependent quantities, above all the transverse energies of the four leading jets, have the best signal to background discrimination.

5.2.3 Template method using decay length technique

This method uses the transverse decay length, L_{xy}, of b-hadrons from top decay in order to reconstruct top mass. Figures 5.7 gives a graphical visualization of L_{xy} in the transverse plane.
In top quark rest frame b-quark momenta $p_b \propto m_t$ for large m_t. For a b hadron of momentum p_b, angle θ with the beam-axis, mass m_b and proper lifetime τ_0, $L_{xy} = \tau_0 \frac{p_b}{m_b} sin(\theta)$ can be directly measured.
This method uses signal and background Monte Carlo distributions as probability density functions from which sets of pseudo-experiments are performed. Mean values of each set of pseudo-experiments are histogrammed and mean and width are extracted accordingly. This procedure is repeated for each generated top mass. Means are then parametrized as function of top mass by using a third degree

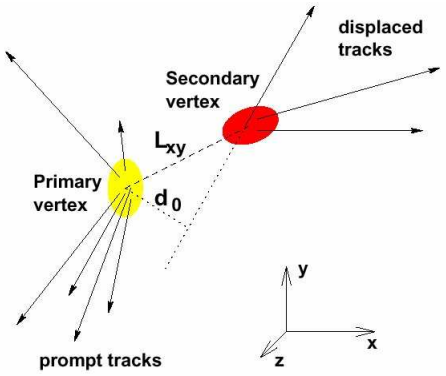

Figure 5.7: A visualization of transverse decay length

polynomial. The fit to the expected distribution of the means as a function of top mass establishes the top mass estimator, once a measured mean L_{xy} is given.

This method, as other methods that suffer from a large statistical error, are relatively new. This means that more studies should be performed in order to optimize these methods.

Despite of the simplicity, these methods are insensitive to large sources of systematic uncertainty (JES for example, see Section 4.1.4) and will be important above all at LHC, where the statistical error will be small relative to the systematic one.

5.3 All-hadronic channel

This channel is characterized by a 6 jets final state topology with little missing energy in the event.
In addition to the small S/N ratio (prior to any cuts $\sim 1/1000$ in the large E_T jet sample) a search in this channel faces multiple ambiguities in the jet-to-partons association. In each event there are $6! = 720$ possible jet permutations (if no b-tagging is applied). However, since a number of them do not change the top and W mass $\frac{720}{2\times2\times2} = 90$ are the possible combinations. If one b-jet is tagged the number of combinations decreases to 30 ($30 = \frac{2\times5!}{8}$), while if both b-jets are tagged we have a total of 6 ($6 = \frac{2\times4!}{8}$) combinations.

At CDF one TM technique has been developed for measuring the top quark mass in the all-hadronic channel. This technique introduces a number of additional kinematical tools relative to template methods in the dilepton and lepton+jets channel.

Despite these difficulties a mass measurement with about a 5 GeV/c^2 overall error has been obtained in a template analysis. A reduction of the error to about 4 GeV/c^2 is in view in a two parameters template analysis where the error on the JES is reduced by constraining to the W mass the jet pairs attributed to the decay of the two W's.

5.4 Comments on the different measurement methods

In this sub-sections we make a number of comments on the merits of different ways to measure the top mass. The full picture of the CDF Run II results can be seen in figure 1.2.

Among the template methods the most sensitive ones are those which use M_t^{reco} as top mass estimator. Naively this is expected, since in the reconstructed top mass much of the kinematical information about the event is included. Other template methods that use less sensitive parameters than M_t^{reco} were developed in order to avoid some systematic uncertainties by which the methods using M_t^{reco} are affected. For example the one using leptons P_T is not affected by jet systematics. Using little sensitive variable, they suffer from a large statistic uncertainty, but the increasing statistic in Run II ($L \approx 2fb^{-1}$) make the statistical error become less important with time. We can expect that in the future these simple methods will give some useful results.

It is appropriate to expand on these considerations in our thesis about the dilepton channel.
Here we have three different methods, KIN, NWA, PHI. They all use M_t^{reco} as sensitive variable, but differ for the variable chosen to constrain the kinematics (see subsection 5.1). All methods give consistent results with comparable errors. The consistency indicates that the systematic errors special to each method have been properly accounted for.

Finally, we comment on why it is important to have many ways to measure top quark mass in different channels and with different methods.
We can mention three advantages

- Measuring the mass with different methods in the same channel, in our case the dilepton one, serves as a cross-check and protects from systematic errors.
- The measurements can be combined to provide a measurement with smaller uncertainty.
- If measurements in different decay channels would significantly differ from one another we might conclude that we are not taking into account some

events that would be background for a channel but not for other ones. This would be hint for physics beyond Standard Model, since all possible kinds of events suggested by the SM have been considered.

Chapter 6

Neutrino ϕ Weighting Method

In this chapter we will focus on the method we are using to measure the top quark mass: the **Neutrino ϕ Weighting Method (PHI)**.

As we wrote in the previous chapter, this method exploits M_t^{reco} as a variable sensitive to the true top mass and reaches the final goal by comparing distributions obtained in Monte Carlo generated events with the distribution obtained in the data.
In 5.1 we discussed briefly two other methods that perform the measure within the same approach. In this chapter we are going to present the full procedure of the PHI method, from the choice of the constrained variables to the top quark estimator. We will provide as many details as possible, and illustrate how we developed and are developing this method step by step. We will also point out the improvement that this analysis represents with respect to analysis performed previously by the Dubna group [44].

6.1 Principles of the Method

6.1.1 Constrained variables

We have: a total of 24 unknowns (t, $\overline{t}$, W^-, W^+, ν, $\overline{\nu}$ 4-momenta) and only 23 equations (see 5.1) available to constrain the kinematics. We must assume at least one of these variables as known, in order to be able to find solutions from the $t\overline{t}$ decay chain.
The PHI method chooses to make assumptions about the 2-D vector $\overline{\mathbf{R}} = (\phi_{\mathbf{1}}, \phi_{\mathbf{2}})$. A net of solutions in the two dimensional $\overline{R}$ plane is built. Because of the symmetry of the solutions for $\phi'_{\nu_1,\nu_2} = \phi_{\nu_1,\nu_2} + \pi$ this net can be limited to the range $(0, \pi) \times (0, \pi)$.

6.1.2 Fitter

The Fitter reconstructs the most likely kinematics of the event in the t-tbar assumption and finds a top quark mass per event.

The χ^2 form

The **Fitter** makes use of all final state observables in the χ^2 below [1], to be minimized in order to find the top mass favoured by the event:

$$\chi^2 \equiv -2ln(\mathscr{P}(\mathbf{x})) \tag{6.1}$$

where $\mathscr{P}$ is the probability density in the variable $\mathbf{x}$: in our case $\mathbf{x}$ means transverse momentum components or a number of invariant masses in the case of jets and leptons, energy in the case of unclustered energy.
We have a contribution to the χ^2 for each observable: we write the total χ^2 in the formula below:

$$\chi^2 = \chi^2_{reso} + \chi^2_{constr} \tag{6.2}$$

$$\chi^2_{reso} = \sum_{l=1}^{2} \frac{(P_T^l - \tilde{P_T^l})^2}{{\sigma^l_{P_T}}^2} + \sum_{j=1}^{2}[-2ln(\mathscr{P}_{tf}(\tilde{P_T^j}|P_T^j))] + \sum_{i=x,y} \frac{(UE^i - \tilde{UE^i})^2}{{\sigma^i_{UE}}^2} \tag{6.3}$$

$$\begin{aligned}\chi^2_{constr} = &-2ln(\mathscr{P}_{BW}(m_{inv}^{l_1,\nu_1}|M_W,\Gamma_{M_W})) - 2ln(\mathscr{P}_{BW}(m_{inv}^{l_2,\nu_2}|M_W,\Gamma_{M_W})) \\ &-2ln(\mathscr{P}_{BW}(m_{inv}^{l_1,\nu_1,j_1}|\tilde{M}_t,\Gamma_{\tilde{M}_t}) - 2ln(\mathscr{P}_{BW}(m_{inv}^{l_2,\nu_2,j_1}|\tilde{M}_t,\Gamma_{\tilde{M}_t}))\end{aligned} \tag{6.4}$$

The upper-marked variables are the output values adjusted by the minimization procedure, whereas P_T and UE (unclustered energy) represent measured values corrected for known detector and physics effects. $\tilde{M}_t$ is the fit parameter giving the reconstructed top mass. Indices BW and tf indicate relativistic Breit-Wigner and transfer function respectively. Since the utilization of Breit-Wigner distributions is new with respect to [44] we will discuss it more in this chapter, while we refer to appendix I for more information on the transfer functions.
The χ^2 has two terms: the first one, χ_{reso}, takes into account the detector uncertainties, whereas the second one constrains the parameters to the known physical constants.
The first sum runs over the primary lepton (tight lepton) and the track lepton

[1] The χ^2 defined below does not refer to a chi-square in the statistical sense but it is still a function to be minimized.

(loose lepton). We use in the denominator the known values for the uncertainty in measuring primary electron and muon P_T respectively. For the loose lepton momentum uncertainty we use the same formula as for the muons:

$$\frac{\sigma^e_{P_T}}{P^e_T} = \sqrt{\frac{a}{P^e_T} + b} \tag{6.5}$$

$$\frac{\sigma^\mu_{P_T}}{P^\mu_T} = c \cdot P^\mu_T \tag{6.6}$$

where $a = 0.135^2$, $b = 0.02^2$, $c = 0.0011$ are taken from the Run I study [43].

The second sum is over the two leading jets. These transverse momenta have been corrected for multiple interactions, underlying event contribution and out-of-cone jet energy loss ("level 4, 6 and 7" corrections in CDF jargon) even though jets corrected for detector defects only ("level 5 corrections") have been considered in selecting the candidate events [2].
The third sum is over the two transverse components of the unclustered energy [3]. The 6.4 terms constrains the invariant masses of the lepton-neutrino pairs and of the lepton-neutrino-leading jet system to the W and to the top mass respectively.

We set $M_W = 80.41\ GeV/c^2$, $\Gamma_{M_W} = 2.06\ GeV/c^2$, as suggested by P.D.G and for the top width we use the function

$$\Gamma_{M_t} = \frac{G_F}{8\sqrt{2}\pi} M_t^3 (1 - \frac{M_W^2}{M_t^2})^2 (1 + 2\frac{M_W^2}{M_t^2}) \tag{6.7}$$

according to the standard model. The dependence of Γ_{M_t} on M_t is shown in figure 6.1.

Allowing for the mass dependence of the top width is something new. We will comment at the end of this section on the improvements obtained by introducing this novelty in the fitter.

[2] see 4.4 for further explanations.

[3] UE is defined as the sum of all unclustered energy in the calorimeter, that is the sum of the towers which are not associated with any of the objects previously considered in the χ^2 formula (tight lepton, track-lepton, two leading jets). Notice that this definition also includes possible additional jets with $E_T > 8\ GeV$ and $|Eta| < 2$ not already taken into account as leading jets.

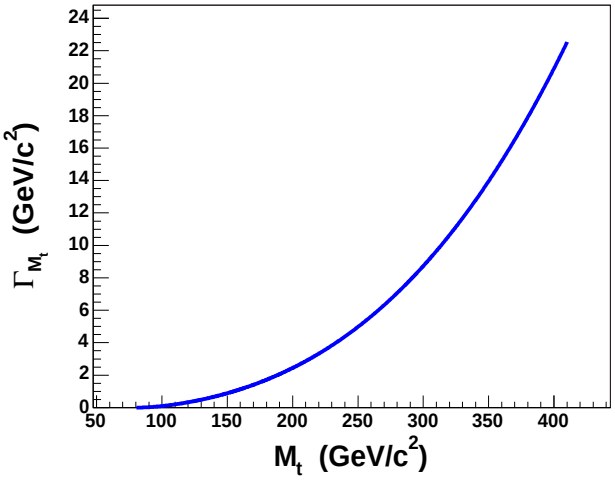

Figure 6.1: Top Width Vs top mass

Scanning the $(\phi_1,\ \phi_2)$ plane

As in the "KIN" and "NWA" methods we cover the entire field spanned by the constraining variables. We choose a step and we scan the $(0,\pi)\times(0,\pi)$ (ϕ_1,ϕ_2) net accordingly.
For each point of the net we can write the following linear system:

$$\begin{cases} P_T^{\nu_1} cos(\phi_{\nu_1}) \ + \ P_T^{\nu_2} cos(\phi_{\nu_2}) \ = \ \not{E}_{T_x} \\ P_T^{\nu_1} sin(\phi_{\nu_1}) \ + \ P_T^{\nu_2} sin(\phi_{\nu_2}) \ = \ \not{E}_{T_y} \end{cases} \tag{6.8}$$

that is solved by:

$$\begin{cases} P_x^{\nu 1} \equiv P_T^{\nu 1} \cdot cos(\phi_{\nu 1}) = \frac{\not{E}_x \cdot sin(\phi_{\nu 2}) - \not{E}_y \cdot cos(\phi_{\nu 2})}{sin(\phi_{\nu 2} - \phi_{\nu 1})} \cdot cos(\phi_{\nu 1}) \\ P_y^{\nu 1} \equiv P_T^{\nu 1} \cdot sin(\phi_{\nu 1}) = \frac{\not{E}_x \cdot sin(\phi_{\nu 2}) - \not{E}_y \cdot cos(\phi_{\nu 2})}{sin(\phi_{\nu 2} - \phi_{\nu 1})} \cdot sin(\phi_{\nu 1}) \\ P_x^{\nu 2} \equiv P_T^{\nu 2} \cdot cos(\phi_{\nu 2}) = \frac{\not{E}_x \cdot sin(\phi_{\nu 1}) - \not{E}_y \cdot cos(\phi_{\nu 1})}{sin(\phi_{\nu 1} - \phi_{\nu 2})} \cdot cos(\phi_{\nu 2}) \\ P_y^{\nu 2} \equiv P_T^{\nu 2} \cdot sin(\phi_{\nu 2}) = \frac{\not{E}_x \cdot sin(\phi_{\nu 1}) - \not{E}_y \cdot cos(\phi_{\nu 1})}{sin(\phi_{\nu 1} - \phi_{\nu 2})} \cdot sin(\phi_{\nu 2}) \end{cases} \tag{6.9}$$

Since we have constrained two degrees of freedom, we can perform a *1C* minimization of the χ^2, given by 6.2. This is done for every point of the net [4].

[4]Net points satisfying the equation $\phi_1 - \phi_2 = 0, \pi$ are conveniently redefined in order to avoid a singularity (6.9).

We note that we would have the same components of the neutrino's momentum $P_{x,y}^{\nu_1,\nu_2}$ for $\phi'_{\nu 1,\nu 2} = \phi_{\nu 1,\nu 2} + \pi$ (see 6.9) and we would include three unphysical solutions ($P_T^{\nu 1} < 0$ and/or $P_T^{\nu 2} < 0$) by considering the remaining three regions $(0,\pi) \times (\pi, 2\pi)$, $(\pi, 2\pi) \times (0,\pi)$, $(\pi, 2\pi) \times (\pi, 2\pi)$ of the $(0, 2\pi) \times (0, 2\pi)$ (ϕ_1, ϕ_2) space, as shown by the equation below:

$$\phi'_{\nu 1,\nu 2} = \phi_{\nu 1,\nu 2} + \pi \implies P'_{x,y}{}^{\nu 1,\nu 2} = P_{x,y}^{\nu 1,\nu 2} \text{ and } P'_T{}^{\nu 1,\nu 2} = -P_T^{\nu 1,\nu 2} \quad (6.10)$$

We avoid this inconvenience by choosing corresponding physical solutions according to 6.10 when unphysical solutions are found.

The net is chosen to have 12x12 points. We have checked that choosing a finer net would not give any significant improvement (see figure 6.5), while each solution requires a long CPU time.
We will deal again with this topic in section 6.4.

With 8 solutions per net point we have to do 1152 $1C$ minimizations, each of which returns a value of M_{ijk}^{reco} and χ^2_{ijk} ($i = 1, \ldots, 12$; $j = 1, \ldots, 12$; $k = 1, \ldots, 8$).

We select the lowest χ^2 solution for each point of the (ϕ_1, ϕ_2) net, thereby reducing the number of obtained masses to 144 per event. All these values carry information on the mass favoured by the event: *the problem is how to weight the information carried by each one of them.*

Weighting the solutions

In section 6.1.2 we introduced the Breit-Wigner functions in the χ^2 formula 6.2 to take into account how invariant masses are distributed in W or t decay. The adopted relativistic Breit-Wigner formula is:

$$\mathscr{P}_{BW}(m_{inv};\ m, \Gamma) \ \equiv \ \frac{\Gamma^2 \cdot m^2}{(m_{inv}^2 \ - \ m^2)^2 \ + \ m^2\Gamma^2} \quad (6.11)$$

where m and Γ are the mass and the width of top or W, m_{inv} refers to the invariant mass, calculated with the appropriate information from lepton and neutrino, in the case of W decay, or from lepton, neutrino and leading jet, in the case of t decay.

In the case of $W \to l\nu$, formula 6.11 has a constant numerator. This plays no role in the χ^2 minimization. In the case of top decay Γ_{M_t} depends on one of the parameters for the minimization of the χ^2. We must allow for this dependence, but we are free to use any Breit-Wigner functions, like the following one, as long as they differ from 6.11 only by a normalization factor:

$$\mathscr{P}_{BW}(m_{inv};\ m, \Gamma) \ \equiv \ \frac{\Gamma \cdot m^2}{(m_{inv}^2 \ - \ m^2)^2 \ + \ m^2\Gamma^2} \quad (6.12)$$

6.11, 6.12 are plotted in figure 6.2 assuming three different top masses.

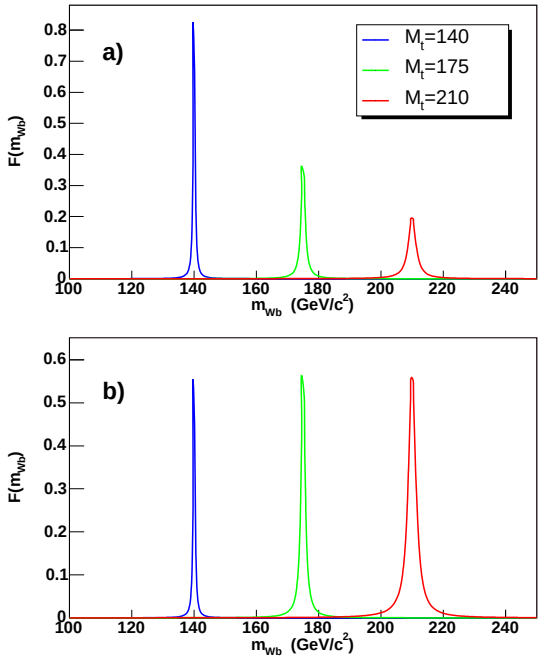

Figure 6.2: Two possible relativistic Breit-Wigner functions. a), 6.12; b), 6.11

On the other hand, while weighting the 144 solutions of an event, the normalization of the adopted BW is important. The expression for the weight is given below:

$$w_{ij} \;=\; e^{-\frac{\chi^2_{ij}}{2}}; \qquad i,j = 1,\cdots,12 \tag{6.13}$$

this formula is obtained by inverting 6.1.

While expression 6.11 was used in the minimization, we adopted expression 6.12 to associate a weight to a solution, since simulations showed that this would minimize the statistical error. We note for completeness that the two χ^2 are related as follows:

$$w_{ij} \;\propto\; \Gamma^{-2}_{M_t} e^{-\frac{\chi^2_{ij,minimiz}}{2}} \tag{6.14}$$

where, obviously, $\chi^2_{minimiz}$ refers to the χ^2 used in the minimization. This relation shows that the solution with the smallest mass will have the largest weight among solutions with the same $\chi^2_{minimiz}$. In such a case this solution is adopted.

6.1.3 Picking up the solution

Once the weight for each of the 144 masses per event is determined, we define an optimized procedure to obtain the final reconstructed mass per event. The procedure follows the steps below:

1. We build a probability density distribution ("pdd") of reconstructed masses with 144 entries. Each mass has been weighted according to 6.13.

2. We identify the most probable value (MPV).

3. We apply a discriminator level (DL) cut and neglect masses in bins with less than the DL fraction of the MPV bin content.

4. We calculate M_t^{reco} as the mass weighted average.

In this measurement we are using $\mathbf{DL = 0.3}$. This was the optimal choice indicated by a study on simulated events [44]. The discriminator level was varied in the interval from 0.0 to 0.99. Figures 6.3 show the results from MC pseudo-experiments [5]. The points at DL=-0.1 obtained by accepting all 144 masses per event [46]. We observe that any positive DL values give a better resolution (better a-priori statistical error [6]) than DL=-0.1 and that a value around DL=0.3 gives optimal performance.

[5] In pseudo-experiment masses are drawn randomly from signal and background templates, according to expected number of signal and background events. For more details see Chapter 8

[6] See Chapter 8 to understand more about the a-priori statistical error.

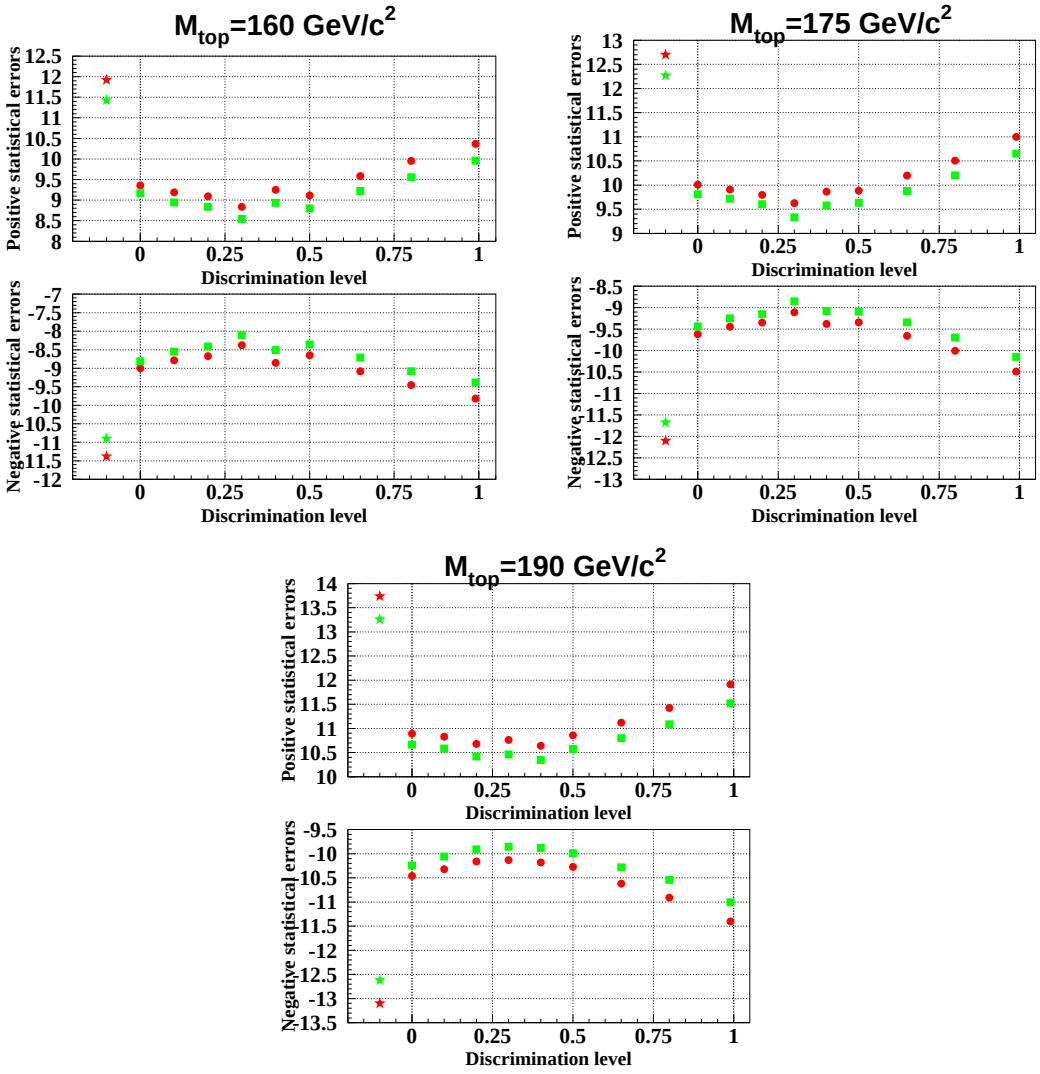

Figure 6.3: Mean (red), median (green) of positive and negative statistical error vs. cut (DL). Top mass 160 (upper left), 175 (upper right), 190 (lower) GeV/c^2

6.2 Templates

The reconstructed top mass might be biased by many sources, also beyond the inaccuracy in the measurements of the involved variables. Examples are theoretical assumptions of some quantities in 6.2 (top decay width, etc...), the possible adoption of a gluon jet as one of the two leading jets in the final state, a wrong lepton-jets combination choice.

In the **Template Method** the top mass is found by comparing the mass distribution in data and in simulated events. In as much as the simulation correctly describes the measurement, since simulated and real data are analyzed in the identical way, biases due to a number of inaccuracies like wrong jet assignments, underlying event and out of cone corrections are automatically accounted for when deriving the top quark mass from the measured mass-sensitive parameter.

Monte Carlo events generation

We generate template events by the Monte Carlo technique. This technique starts from flat distributed variables and, using constraints from theory or experiment-driven models, returns the distributions of the desired variable.

- **Signal samples**

  We generate a large number of $t\bar{t}$ events at parton level with top mass in the range from 155 GeV to 195 GeV in two 2 GeV steps. We simulate particle interactions with the detector, accounting for inefficiencies and for non instrumented zones. Smearing, due to experimental uncertainties, is included as a constrain.

- **Background samples**

  We use the theoretical cross sections and generate the process of interest [7] at parton level. The detector simulation is as for the signal samples.

6.3 Likelihood Form

In the previous sections we gave a prescription to obtain an estimator for the true top mass per event.

The final step [8] is to compare the mass distribution in data events with simulation in Monte Carlo events.

[7] see Chapter 7 for more information on the considered background events

[8] Internal CDF severe rules do not allow to work on data until the analysis method is discussed and approved by the Collaboration.

We perform this comparison by parametrizing with analytic functions M_t^{reco} distributions in Monte Carlo [9], building a M_t^{reco} distribution in data and matching the data to the Montecarlo-derived distributions with a *likelihood unbinned fit and minimization*.

The fit to the likelihood function finds the probability that our M_t^{reco} distribution is described by a mixture of background events and dilepton $t\bar{t}$ events with an assumed top quark mass. The maximum likelihood determines the true top mass.

The likelihood used by the PHI method (6.15) differs from 5.6 because of an additional $\mathscr{L}_{param}$. Below we report the total likelihood.

$$\mathscr{L} \equiv \mathscr{L}_{shape} \cdot \mathscr{L}_{backgr} \cdot \mathscr{L}_{param}; \tag{6.15}$$

where,

$$\mathscr{L}_{shape} \equiv \frac{e^{-(n_s+n_b)} \cdot (n_s+n_b)^N}{N!} \cdot \prod_{n=1}^{N} \frac{n_s \cdot f_s(m_n|M_{top}) + n_b \cdot f_b(m_n)}{n_s+n_b} \tag{6.16}$$

and

$$\mathscr{L}_{backgr} \equiv exp(\frac{-(n_b - n_b^{exp})^2}{2\sigma^2_{n_b^{exp}}}) \tag{6.17}$$

$$\mathscr{L}_{param} \equiv exp\{-0.5[(\vec{\alpha}-\vec{\alpha_0})^T U^{-1}(\vec{\alpha}-\vec{\alpha_0}) + (\vec{\beta}-\vec{\beta_0})^T V^{-1}(\vec{\beta}-\vec{\beta_0})]\}. \tag{6.18}$$

α and β are parameters determined from fitting with analytic functions, the signal and background templates respectively (see Chapter 7 for further details). U and V are the corresponding covariance matrices, returned from the MINUIT [78] minimization.

We assign a shape probability (f_s) that each of the selected "LTRK" events is signal or (f_b) background. These probabilities are weighted according to appropriate signal and background numbers, n_b and n_s, that are self-estimated by the likelihood fit. These two parameters are not completely free. The first one follows a Gaussian distribution (see 6.17) where n_b^{exp} and $\sigma_{n_b^{exp}}$ are taken from "LTRK" Dilepton Cross Section Analysis Group [47], while the $n_s + n_b$ is constrained by a Poisson term $\frac{e^{-(n_s+n_b)} \cdot (n_s+n_b)^N}{N!}$, as shown in 6.16.

It should be observed that in this way the number of signal events estimated by the likelihood is not forced to be distributed around the expected number of signal events calculated by using $t\bar{t}$ cross section info [10].

[9] see Chapter 7

[10] for more info on the cross-section dependent top mass measurement see [50]

The likelihood maximization procedure (we usually minimize $-ln(\mathscr{L})$) returns a **top quark mass estimator $\mathbf{m_t}$**, as the mass value corresponding to $[-ln(\mathscr{L})]_{min}$, and **its statistical uncertainty**, by setting $-ln(\mathscr{L})$ to $[-ln(\mathscr{L})]_{min} + 0.5$. We note that uncertainties in signal and background parametrization are included in the statistical error [11].

6.4 Monte Carlo studies

Studies are required to check whether

- the minimization of the χ^2, defined by 6.2, is able to reproduce the right final particles configuration.
- choosing another segmentation for the $(\phi_1,\ \phi_2)$ net would change the result.

To address the first question we study the kinematics of the dilepton $t\bar{t}$ events [12]. We identify the primary leptons by matching tracks to leptons in the hepg bank. With this technique we also know the neutrinos $(\phi_1,\ \phi_2)$ angles. In parallel we run our standard χ^2 procedure (see 6.1.2).
Examples of χ^2 distributions in the $(\phi_1,\ \phi_2)$ plane are presented in figure 6.4.

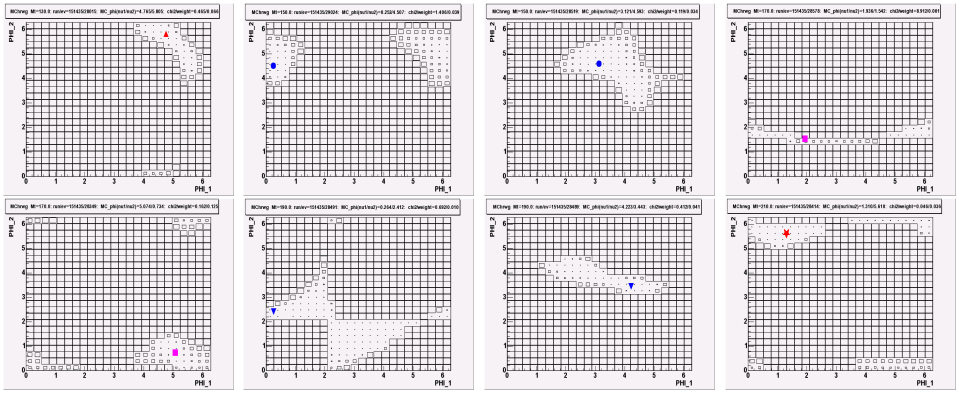

Figure 6.4: χ^2 distributions in the $(\phi_{\nu 1}, \phi_{\nu 2})$ plane for a number of simulated dilepton signal events. The region of neutrino azimuthal angles preferred by the reconstruction can be seen. The genuine neutrino angles are shown as the color marker.

[11] In NWA and KIN methods these contributions are included in the systematic error [45]. Such contribution is called MC statistics error.

[12] we are performing this check using dilepton events passing "DIL" selection (see Appendix B for details)

Each plot corresponds to one particular event. Only one χ^2 out of $\chi^2_{ij}, \quad i,j = 1\cdots,24$ is displayed. We plot the solution up to 2π on both axes since we are interested in which of the four sectors we find the physical solution (see 6.1.1). The white areas (areas with a relative good χ^2 value) surround the genuine point (the one extracted from hepg bank) confirming that our procedure to reconstruct M_t^{reco} is working properly.

To understand the impact of the chosen (ϕ_1, ϕ_2) ***segmentation*** we produced MC templates for three different grids (6x6, 12x12, 24x24), as shown in figure 6.5. The obtained RMS are shown in figure 6.6. While the large mass tail is

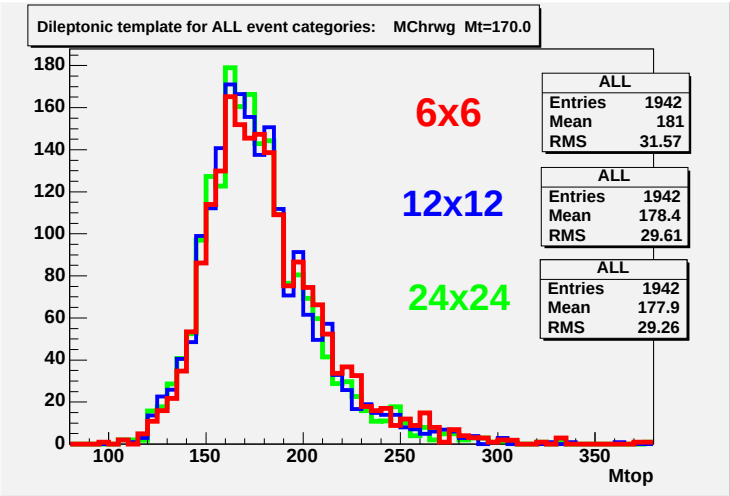

Figure 6.5: MC mass templates for different neutrino ϕ-plane segmentation for top mass 170 GeV/c^2

slightly reduced when the 12x12 segmentation is chosen rather than 6x6, there is no significant progress in increasing the segmentation further to 24x24. Overall the phi segmentation is not influential. A 12x12 grid is an approximately optimized choice.

More studies which were performed on the method are described in Appendix G.

6.5 Improvements

We introduced two changes relative to [44] in the definition of the χ^2_{constr} (6.4): we switched from the Gaussian functions to represent the top width of the previous

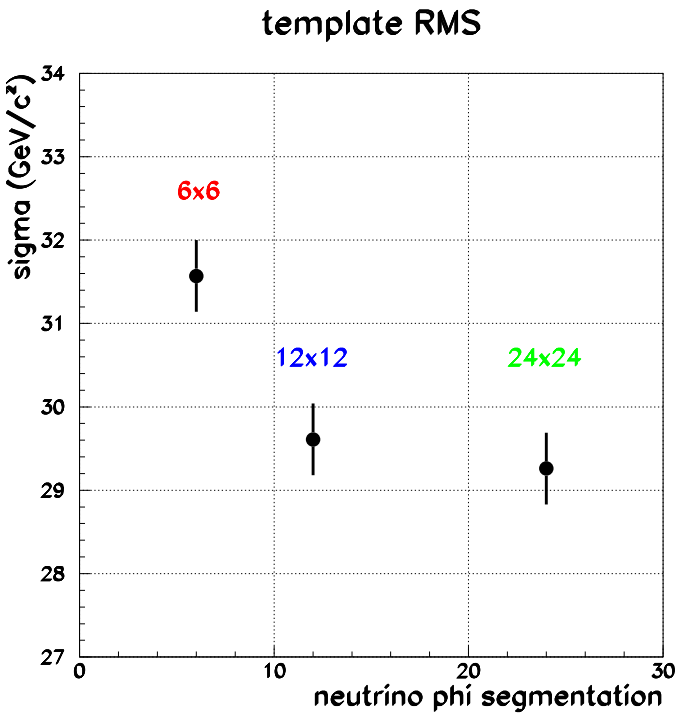

Figure 6.6: RMS of templates vs. neutrino phi plane segmentation

analysis (see appendix F) to the physical Breit-Wigner to schematize more appropriately the t, $\overline{t}$, $W^{\pm}$ decay chains, and we use the M_t dependent rather than a constant Γ_t. We also use an optimized Breit-Wigner (see 6.12) to weight the solutions. A $\sim 20\%$ improvement has been obtained. Details are given in appendix H.

A change in the χ^2_{reso} (6.3) made with respect to the previous analysis is the introduction of the transfer functions [13] $k = (P_T^{part} - P_T^{jet})/P_T^{jet}$, where P_T^{jet} is the transverse momentum of the considered jet and P_T^{part} is the transverse momentum of the parton associated to the jet. Although this term allows for a more correct description of b parton response the improvement in the mass resolution with respect to the previous analysis (see F) is barely significant. The procedure to obtain the transfer functions and the obtained improvement are described in appendix I.

[13] for further details see Appendix I

Chapter 7

Templates

Templates used in this analysis are mass distributions whose entries are $t\bar{t}$ (signal) or background simulated events, selected according to the "LTRK" dilepton selection (see 4.4).

In this Chapter we are going to give some details about Monte-Carlo samples used to create signal (7.1) and background (7.2) templates. In section 7.2.1 the procedure used to combine the background processes and create an overall background template will be described.

7.1 Signal $t\bar{t}$ templates

Pythia [1] $t\bar{t}$ samples of common CDF use [57] with top mass within $155 \div 195\ GeV/c^2$ in $2\ GeV/c^2$ steps have been used [2]. The samples generated with a top mass below $165\ GeV$ and above $185\ GeV$ contain approximately 0.6×10^6 events, while the other ones contain approximately twice as many.

We require that all the events pass the "LTRK" dilepton selection. The number of selected events per each sample is shown in table 7.1. Events generated with higher top masses have a larger probability to be selected because of the larger available space-phase.

For each event passing the selection cuts we reconstruct the invariant mass (M_t^{reco}) and build a mass distribution. We thus obtain a mass distribution (template) for each sample.

[1] Pythia is an event generator for a large number of physics processes.

[2] this range is more or less symmetric around the word average value of the top quark mass ($172.5 \pm 2.7\ GeV/c^2$ [58]). It is chosen wide enough to cover at least 3 standard deviations of the mass value

Sample	Size(10^6 events)	# selected events	Acceptance (%)
ttkt55	0.6	4742	0.79 ± 0.01
ttkt57	0.6	4642	0.77 ± 0.01
ttkt59	0.6	4783	0.80 ± 0.01
ttkt61	0.6	4922	0.82 ± 0.01
ttkt63	0.6	5025	0.84 ± 0.01
ttkt65	1.2	10208	0.85 ± 0.01
ttkt67	1.2	10259	0.86 ± 0.01
ttkt69	1.2	10689	0.89 ± 0.01
ttkt71	1.2	10840	0.90 ± 0.01
ttkt73	1.2	10999	0.92 ± 0.01
ttkt75	4.8	45020	0.94 ± 0.00
ttkt77	1.2	11458	0.96 ± 0.01
ttkt79	1.2	11595	0.97 ± 0.01
ttkt81	1.2	11723	0.98 ± 0.01
ttkt83	1.2	12015	1.00 ± 0.01
ttkt85	1.2	12255	1.02 ± 0.01
ttkt87	0.6	6255	1.04 ± 0.01
ttkt89	0.6	6271	1.05 ± 0.01
ttkt91	0.6	6358	1.06 ± 0.01
ttkt93	0.6	6121	1.02 ± 0.01
ttkt95	0.6	6454	1.08 ± 0.01

Table 7.1: Monte Carlo samples [57], number of generated events (GE), events passing the selection cuts (SE) and acceptances (SE/GE).

In a global fit we parametrize all templates by a top mass M_t-dependent function [3]. We choose a combination of one Landau and two Gaussian functions, as shown in 7.1.

$$\begin{aligned} f_s(M_t^{reco}|M_t) &= p_7(p_6\frac{1}{\sqrt{2\pi}p_2}e^{-0.5(\frac{M_t^{reco}-p_1}{p_2}+e^{-\frac{M_t^{reco}-p_1}{p_2}})}+(1-p_6)\frac{1}{\sqrt{2\pi}p_5}e^{-0.5(\frac{M_t^{reco}-p_4}{p_5})^2})+ \\ &\quad (1-p_7)\frac{1}{\sqrt{2\pi}p_3}e^{-0.5(\frac{M_t^{reco}-p_8}{p_3})^2} \end{aligned} \tag{7.1}$$

In 7.1 the $p_1,\ldots,p_8$ parameters are M_t-dependent as:

$$p_k = \alpha_k + \alpha_{k+8}\cdot(M_t - 175) \quad k = 1,\ldots,8 \tag{7.2}$$

[3]Notice that in this fit M_t is considered a paramater

f_s is a probability density function (p.d.f) that the reconstructed mass M_t^{reco} comes from an event with true top mass M_t.

In figure 7.1 a number of templates are shown. Fitting functions (red lines) are obtained by 7.1.

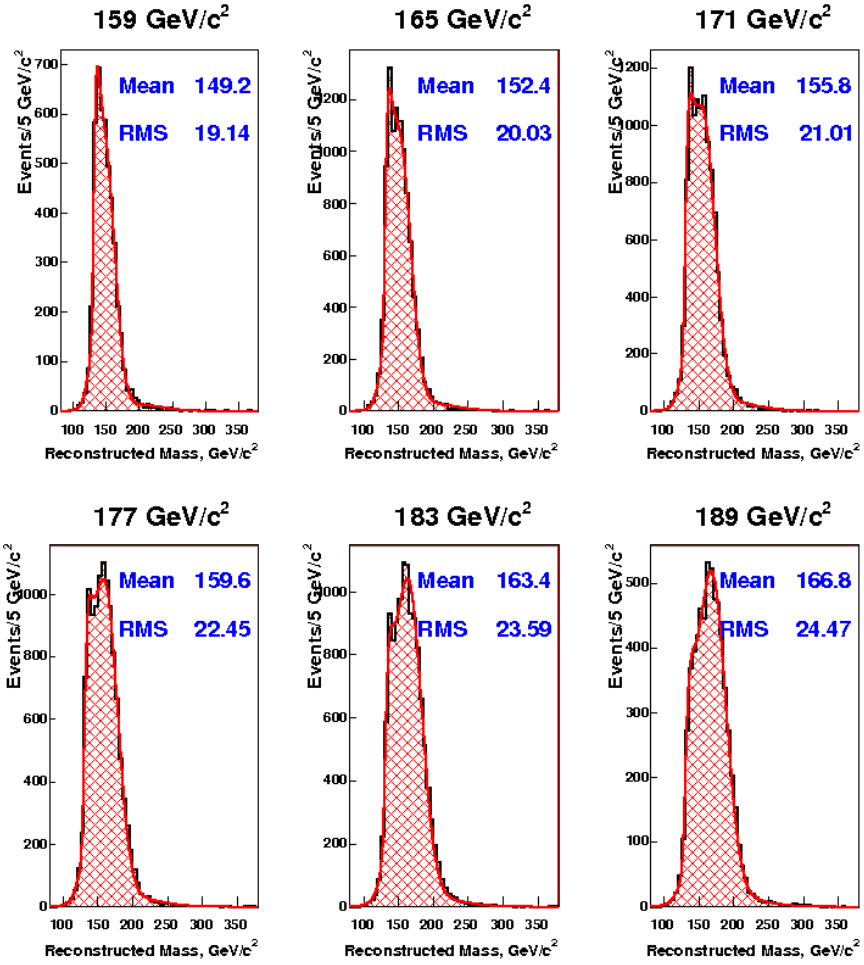

Figure 7.1: Signal templates for a number of generated top masses.

Each template in figure 7.1 has a small but visible narrow peak in the $M_t^{reco} = 140 \div 160\ GeV/c^2$ region, which becomes more important at lower masses. This is a consequence of the fit constraints [4].

7.2 Background templates

From the measured cross section at CDF [47] we expect the dilepton LTRK sample to be composed of about 50% of signal events and about 50% of background events. Therefore it is important that the composition and the overall template of background events be understood.
As already mentioned in 4.3, we include in the background the following processes[5]:

- Diboson
  1. $WW \rightarrow l^+l'^- \ + \ \nu^l\overline{\nu}^{l'}$ [6]
  2. $WZ \rightarrow l^+l^-$
  3. $ZZ \rightarrow l^+l^-$
- Drell-Yan (DY)
  1. $Z/\gamma^* \rightarrow l^+l^-$
- "Fakes": events with a fake lepton (see 4.3 for details)

The used samples [7] are the CDF official MC samples, created with PYTHIA. Samples are listed in table 7.2 [8].

[4] The Breit-Wigner mass distributions, which are narrower at smaller top masses, are biased by the kinematical cut $M_t^{reco} > M_W + M_b$

[5] All these processes may end having in the final state an arbitrary number of jets

[6] $l, l' = e, \mu, \tau$

[7] events in these samples are generated ad hoc for top quark studies. A number of additional large P_T partons are requested at the generator level. This is done in order to reduce the number of useless events for top analysis, given the small acceptance of these events by the selection cuts.

[8] if not specified otherwise, inclusive (on number of jets and type of leptons) processes over all SM-allowed final products are considered.

Category	Process	Sample	Size(10^6 events)	Cross section
Diboson	WW	itopww	2.4	7.70 pb
		wewkbd	3.5	1.27 pb (NLO)
	WZ	itopwz	2.4	2.30 pb
		wewkcd	3.5	0.36 pb (NLO)
	ZZ	itopzz	2.4	2.56 pb
DY ($M_{\gamma^*} = [75, 105]\ GeV/c^2$)	Z(ee)+2p	ztop2p	0.5	3.47 pb
		ztopzb	4.5	3.46 pb
	Z(ee)+3p	ztop3p	0.5	0.55 pb
DY ($M_{\gamma^*} = [20, 75]\ GeV/c^2$)	Z(ee)+2p	xtop2p	0.5	1.61 pb
		xtoppb	4.5	1.60 pb
	Z(ee)+3p	xtop3p	0.5	233 fb
DY ($M_{\gamma^*} = [105, 600]\ GeV/c^2$)	Z(ee)+2p	ytop2p	0.5	117 fb
DY ($M_{\gamma^*} = [75, 105]\ GeV/c^2$)	Z($\mu\mu$)+2p	ztop7p	0.5	3.47 pb
		ztopzt	4.5	3.46 pb
	Z($\mu\mu$)+3p	ztop8p	0.5	0.55 pb
DY ($M_{\gamma^*} = [20, 75]\ GeV/c^2$)	Z($\mu\mu$)+2p	xtop7p	0.5	1.60 pb
		xtoppc	4.5	1.60 pb
	Z($\mu\mu$)+3p	xtop8p	0.5	233 fb
DY ($M_{\gamma^*} = [105, 600]\ GeV/c^2$)	Z(ee)+2p	ytop7p	0.5	117 fb
DY ($M_{\gamma^*} = [75, 105]\ GeV/c^2$)	Z($\tau\tau$)+2p	ztopt2	2.4	4.14 pb
DY ($M_{\gamma^*} = [20, 75]\ GeV/c^2$)	Z($\tau\tau$)+2p	xtopt2	2	1.82 pb
DY ($M_{\gamma^*} = [105, 800]\ GeV/c^2$)	Z($\tau\tau$)+2p	ytop7p	0.25	117 fb
Fakes	W+jets	data	–	–

Table 7.2: Leading order cross section, generated sample and number of accepted events of background processes (first column) [57]. For Drell-Yan the dilepton mass is indicated.

Selected events are reconstructed, as described in Chapter 6. In order to build the $Z/\gamma^* \rightarrow e^+e^-$ (Zee), $Z/\gamma^* \rightarrow \mu^+\mu^-$ (Zmm), $Z/\gamma^* \rightarrow \tau^+\tau^-$(Ztt) templates events are weighted according to σ_s/N_s, where σ_s and N_s are respectively the cross section of the process and the total number of events (see table 7.2) in the sample. WW, WZ, ZZ templates are built with the same procedure. Zee, Zmm, Ztt, WW, WZ, ZZ are normalized to their expected rates [47] for an integrated luminosity of $\sim 2\ fb^{-1}$ (see table 7.3). *Drell-Yan template* is obtained by summing Zee, Zmm and Ztt. *The Diboson template* is the sum of WW, WZ, ZZ. *The Fakes template* is built from data events (see Appendix C for details). Diboson, DY and Fakes templates are shown in figure 7.2.

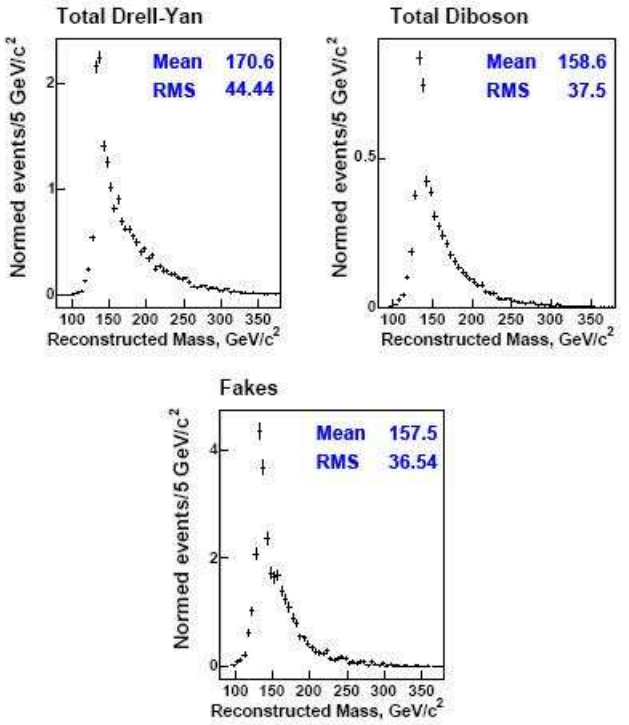

Figure 7.2: Background templates

As already observed in section 7.1, background templates are similar to signal templates around $M_t^{reco} = 140$.

7.2.1 Combined background

The fakes template is normalized to 53.73 events (see table 7.3). Diboson, DY and Fakes templates are then merged into the combined background template (see

figure 7.3).

Process	expected number (1.1 fb^{-1})	rescaled expected numbers (2.0 fb^{-1})
WW	3.90 ± 0.36	7.02 ± 0.65
WZ	1.43 ± 0.38	2.57 ± 0.68
ZZ	0.34 ± 0.02	0.61 ± 0.04
$Z/\gamma^* \to e^+e^-$	7.75 ± 2.24	13.95 ± 4.03
$Z/\gamma^* \to \mu^+\mu^-$	3.40 ± 1.15	6.12 ± 2.07
$Z/\gamma^* \to \tau^+\tau^-$	7.31 ± 0.89	13.16 ± 1.60
Fakes	29.85 ± 5.86	53.73 ± 10.55
Background	53.99 ± 6.60	97.18 ± 11.88

Table 7.3: Expected background events for an integrated luminosity of 1.1 fb^{-1} [47] and 2.0 fb^{-1}. Same notations as in table 7.2.

The fitting function, f_b, for the combined background is reported in 7.3.

$$
\begin{aligned}
f_b(M_t^{reco}) \;=\; & \beta_7\Big(\beta_6 \frac{1}{\sqrt{2\pi}\beta_2} e^{-0.5\left(\frac{M_t^{reco}-\beta_1}{\beta_2} \,+\, e^{-\frac{M_t^{reco}-\beta_1}{\beta_2}}\right)} + \\
& + (1-\beta_6)\frac{1}{\sqrt{2\pi}\beta_5} e^{-0.5\left(\frac{M_t^{reco}-\beta_4}{\beta_5}\right)^2}\Big) \;+\; (1-\beta_7)\frac{1}{\sqrt{2\pi}\beta_3} e^{-0.5\left(\frac{M_t^{reco}-\beta_8}{\beta_3} \,+\, e^{-\frac{M_t^{reco}-\beta_8}{\beta_3}}\right)}
\end{aligned}
\tag{7.3}
$$

Two Landau and one Gaussian functions have been used. The main difference with respect to the signal fitting functions f_s is in the expression of the parameters $\beta_1, \ldots, \beta_8$.

The fit to the combined template is shown as a continuous curve in figure 7.3.

7.3 Concluding comment

In this Chapter the parametrized templates for signal and background events have been presented. These parametrizations will be used for the top mass measurement.

Before applying the method to data, we will apply it to simulated events to test whether it gives reasonable outcomes. These tests will be subject of next Chapter.

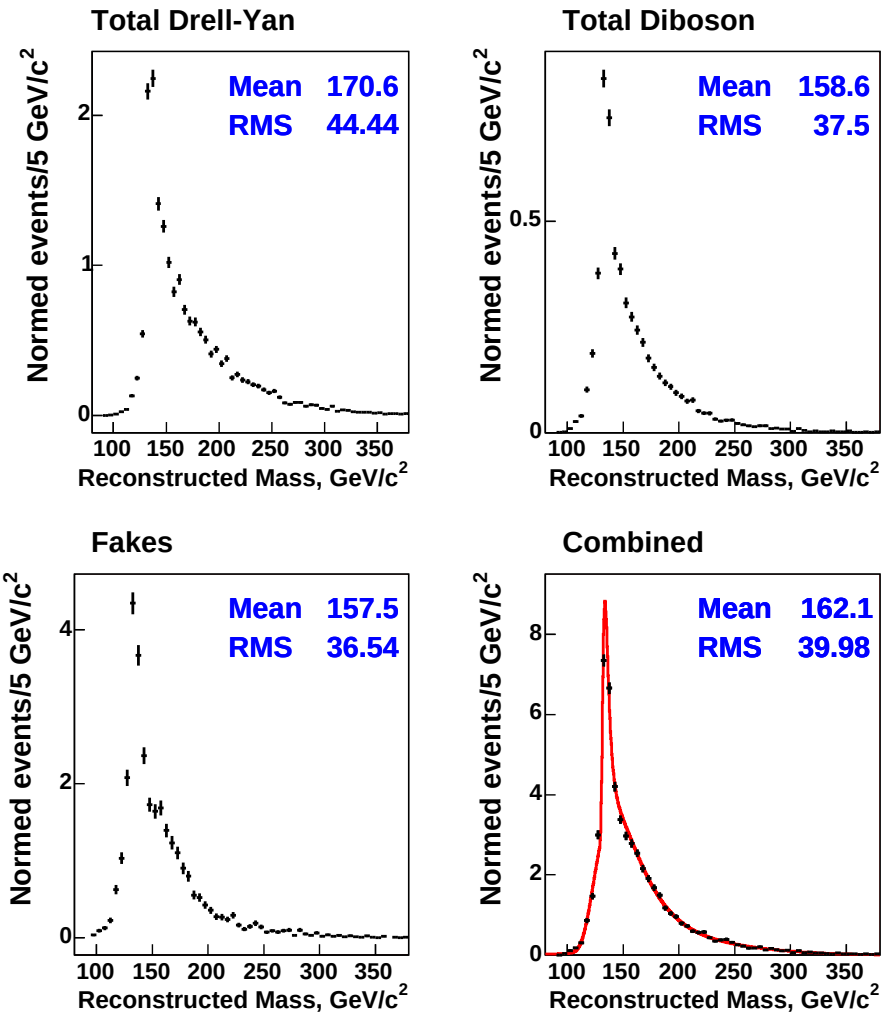

Figure 7.3: Drell-Yan, Fakes, Diboson and Combined background templates. Parametrization (red line) is superimposed to the combined template.

Chapter 8

Testing the Method

The machinery for deriving the top mass from fits to a mass-sensitive experimental distribution has been presented in the previous chapters.

This chapter describes the tests that have been performed to check that the mass and its statistical uncertainty [1] are correctly estimated. These tests have been performed on simulated samples, where the true top mass was known.

Section 8.1 describes the technique used for the tests. The results will be presented in Section 8.2.

8.1 Pseudo-Experiments

The implementation of a generic Pseudo-Experiment (PE) is as follows.

The expected composition of signal and background events in the LTRK data sample is reported in table 8.1 [2].

	expected numbers ($2.0\ fb^{-1}$)
Signal	108.95 ± 3.38
Background	97.18 ± 11.88

Table 8.1: Expected signal and background events for an integrated luminosity of $2\ fb^{-1}$

[1] The systematic uncertainties will be discussed in chapter 9

[2] The expected numbers of signal and total background events are obtained from the CDF top cross section measurement made with lower statistics ($1.1\ fb^{-1}$) [47] and rescaled to the luminosity of $2fb^{-1}$)

This sample can be simulated from Monte Carlo generated events: an average of 109 events is drawn from a signal template and an average of 97 events from the combined background template. Top mass and statistical error can be estimated by a likelihood fit of this sample, as explained in Section 6.3.

This procedure can be iterated, by splitting the original MC sample in many uncorrelated sub-samples. A mass distribution is built with the result of each PE. This distribution is expected to be centered on the generated top mass.

In principle we can perform as many PE's as we want [3]. However, Monte Carlo samples have a limited statistics. In our case it was not possible to have a number of PE's as large as desired containing ~ 100 signal events in average each [4], when disposing of simulated event samples of only $\sim$ 5000 to 45000 (see table 7.1), without generating a significant correlation among PE's. Generating a much larger Monte Carlo statistics was prohibitive from a computing point of view.

Some correlation can be accepted while still gaining in accuracy: the **re-sampling method**[5] demonstrates that the gain in accuracy can be accomplished up to a number of pseudo experiments satisfying the condition $(N_{sample}/N_{PE}^{ev})^2$, where N_{sample} is the number of entries in the signal template (see 7.1) and N_{PE}^{ev} is the number of events in each PE. Beyond this value the correlation among PE's does not allow any further improvements.

For the measurement an average number of 10000 PE's has been chosen.

To test our method we proceeded as follows.

- We generated $N_{PE}^{t\bar{t}}$ signal events and N_{PE}^{back} background events, according to Poisson distributions with means 108.95 and 97.18 respectively.
- We draw $N_{PE}^{t\bar{t}}$ and N_{PE}^{back} masses from the signal template and from the combined background.
- We perform the likelihood fit and find the top mass (M_t^{fit}) and its statistical error.
- We iterate this procedure on 10000 PE's [6]

[3] The accuracy of these tests increases with increasing number of PE's, which is limited in practice.

[4] we do not repeat the same event in the same PE.

[5] We refer the reader to [59] for more information.

[6] A refinement is in progress by applying a Gaussian smearing to the expected number of signal events.

8.2 Sanity Checks

A number of checks have been performed to prove that the adopted templates parametrization and the likelihood fit are able to return a mass value close to generated mass. In order to do this, top mass, mass error and pull distributions are built. The pull is frequently defined as:

$$\frac{M_t^{fit} - M_t}{\sigma(M_t^{fit})} \tag{8.1}$$

where $\sigma(M_t^{fit}) = \frac{\sigma^+(M_t^{fit}) - \sigma^-(M_t^{fit})}{2}$, $\sigma^{\pm}(M_t^{fit})$ are the positive and negative statistical errors. A slightly different pull definition is used in this thesis:

$$\frac{M_t^{fit} - M_t}{\sigma'(M_t^{fit})} \tag{8.2}$$

where $\sigma'(M_t^{fit}) = \begin{cases} \sigma^+(M_t^{fit}) \ if \ M_t^{fit} < M_t \\ |\sigma^-(M_t^{fit})| \ if \ M_t^{fit} > M_t \end{cases}$

8.2.1 Reconstructed Masses and Statistical Errors

We use the median of the mass distribution as top mass estimate (M_t^{out}). Figure 8.1 shows M_t^{out} versus input mass and the bias, defined as $M_t^{out} - M_t$.

The slope of the fitting straight line in the upper plot is consistent with 1.000. The fit was performed in the mass range [160,190] GeV/c^2 [7]: an average bias of -0.16 ± 0.10 is obtained. Although this can be considered compatible with zero, we apply a shift of -0.16 to the result on data.
The error bars in the figure 8.1 are mainly due to the limited MC statistic [8]. They are calculated with a *Boostrap procedure*:

- The bin contents of signal and background templates (see chapter 7) are used as means of Poisson distributions. New bin contents are generated accordingly and new templates are built.
- Re-sampling is performed by drawing masses from new templates.
- The above steps are repeated 100 times for signal and background.
- Two distributions are built.

[7] this covers 3 σ around the world average top mass

[8] Actually the error bars result from $\sigma_{LS} \oplus \frac{\sigma_{PE}}{\sqrt{N_{PE}}}$ where σ_{LS} is the contribution of the limited statistic and $\frac{\sigma_{PE}}{\sqrt{N_{PE}}}$ is due to the limited number of PE. σ_{PE} is the average RMS of the distribution for PE's and $N_{PE} = 10^4$

- The sigmas from the Gaussian fits to the two distributions are summed in quadrature.

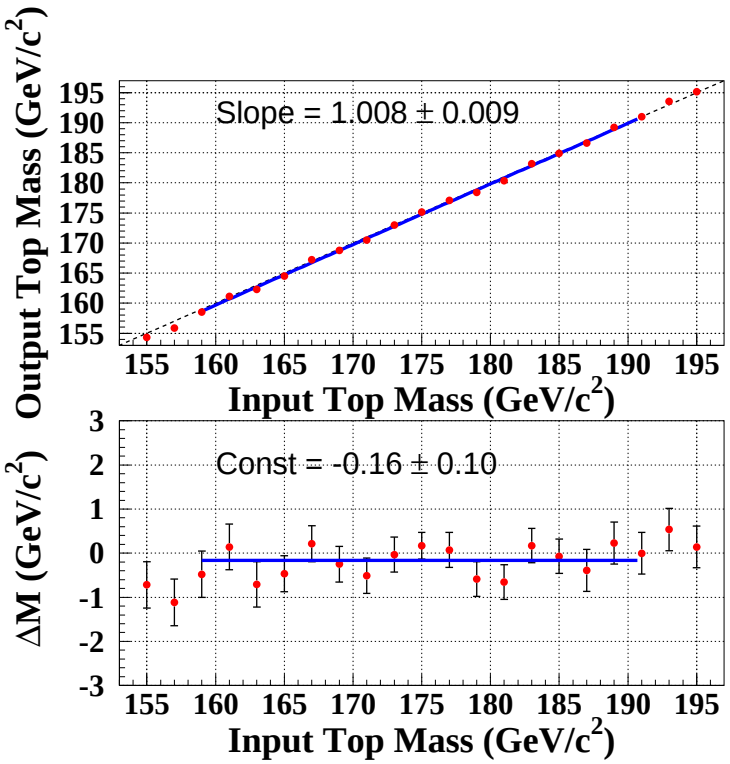

Figure 8.1: Reconstructed versus input masses (upper) and bias (lower)

We used the signal template for $M_t = 175\ GeV/c^2$. The results of this procedure are shown in figure 8.2.
Average mass uncertainty of signal and background templates, associated to bootstrapping are 0.15 and 0.27 respectively. 0.28 is the induced top mass error. The errors on points of figure 8.1 with generated top mass different from 175 GeV/c^2 are calculated by applying a scaling factor [9] which corrects for the different statistic in signal templates (see table 7.1).

[9] The scaling factor is: $\frac{\sqrt{N^{ev}_{175}}}{\sqrt{N^{ev}_{M_t}}}$ where $N^{ev}_{M_t}$ and N^{ev}_{175} are the numbers of selected events in the sample with generated top masses M_t and 175 GeV/c^2 respectively

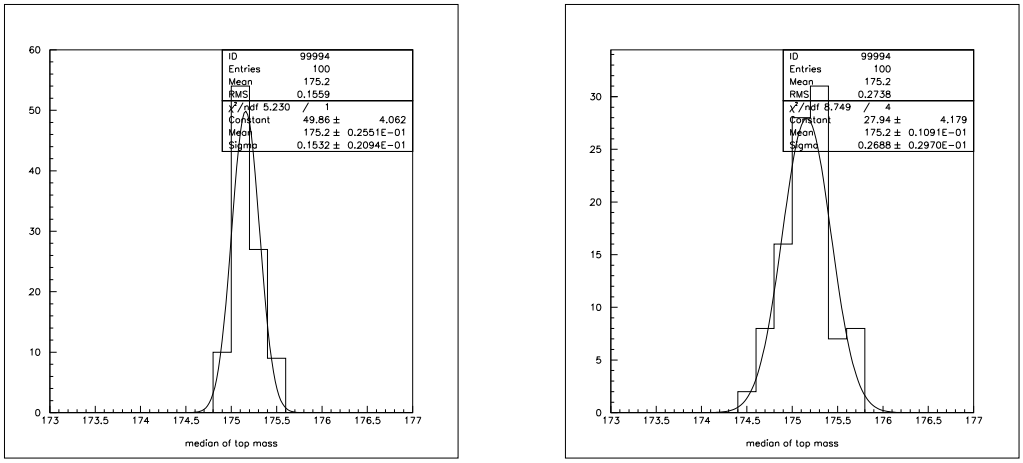

Figure 8.2: Bootstrap results for signal (left) and background (right)

The expected statistical error distribution for an input top mass of 175 GeV/c^2 is shown in figure 8.3.

Its average is 4.3 GeV/c^2. The statistical error dependence on M_t is shown in the figure 8.4. Errors are estimated with the bootstrap procedure mentioned above.

8.2.2 Pull Distributions

Pulls gauge the goodness of mass reconstruction: they express, in units of the statistical error, the bias of our method. Examples of pulls which have been obtained are plotted in figure 8.5.
Pulls have been fitted by using Gaussian functions: the obtained χ^2_{red} shows a good agreement (see figure 8.5).

The mean and the width of pulls versus generated top mass are shown in figure 8.6.
The straight line fit on the upper plot has an offset of -0.04, which is slightly different from the expected value of 0.00. This offsets corresponds to the bias (seen in figure 8.1) for which a correction has been applied.

The average pull width is 1.011 ± 0.005. Accordingly, the statistical error obtained from data will be rescaled by 1.011.

In order to estimate the "a-priori" average statistical error a generate mass of 175 GeV/c^2 has been used as mentioned above (8.2.1). After rescaling by 1.011, it becomes 4.4 GeV/c^2.

Since the machinery (templates, likelihood fit, etc ...) has been tested on the same MC samples that have been used to build signal and background templates,

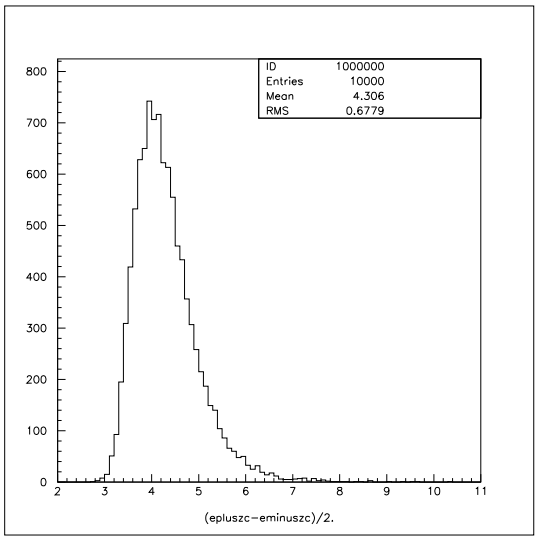

Figure 8.3: Expected statistical error for an input top mass of 175 GeV/c^2.

some independent tests on orthogonal samples are needed in order to validate the method. These are performed by analyzing "blind samples".

8.2.3 Blind samples results

"Blind" samples are provided by the CDF Top Analysis Group Conveners in order to accomplish this goal. Signal samples are generated with a top mass, which is known only to the Conveners. The same distributions as in section 8.2.1 and 8.2.2 have been built.

Figure 8.7 shows examples of results by PE's from 10 blind samples. Bias and mean of the pulls are shown in figure 8.8 .

An average bias of 0.27 ± 0.38 GeV/c^2 and an average width of pulls of 1.029 ± 0.016 have been estimated by a straight line fit. Results are fully consistent with the previously obtained values (see 8.2.1, 8.2.2) and provide the final validation of the method.

It is worth observing that any technical imperfection (e.g. a mathematically inappropriate definition of mass probability, 6.1) would play no role in a template method, in as much as it is applied identically to data and simulations and it is proved to return the correct mass values in blind samples.

The underlying theory to the analysis has a number of limitations (e.g lowest order cross sections used in background estimates) whose impact on the result is accounted for in the estimate of the systematic errors (chapter 9).

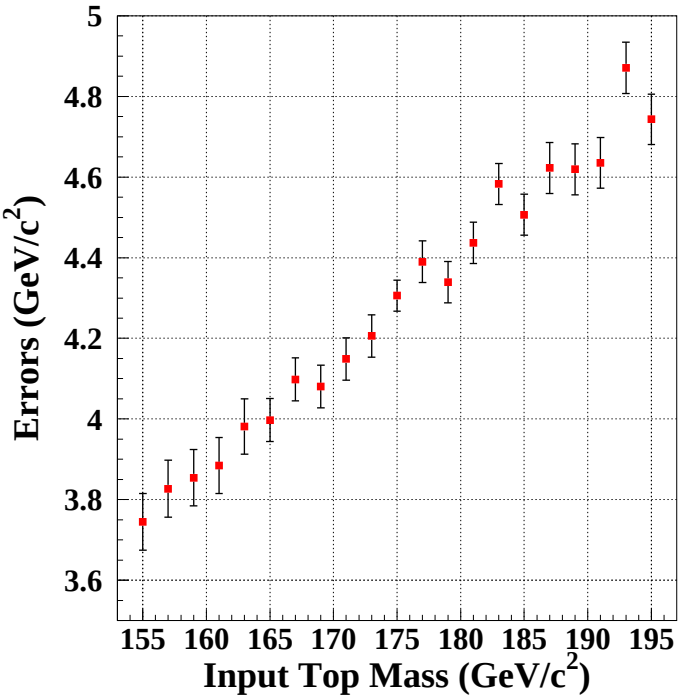

Figure 8.4: Mean of statistical error versus generated top mass.

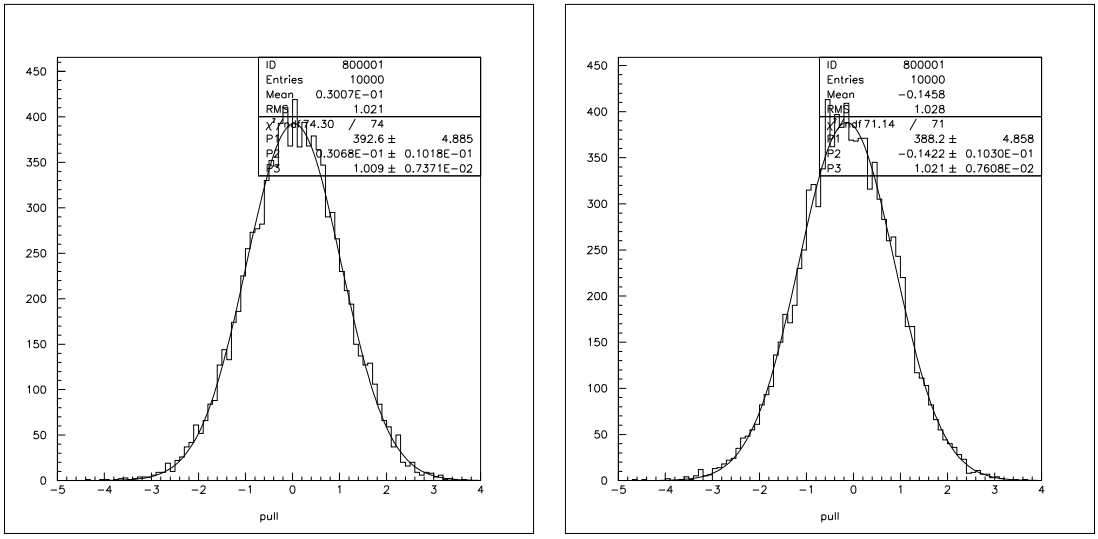

Figure 8.5: Pulls for $M_t = 175\ GeV/c^2$ (upper) and $M_t = 181\ GeV/c^2$ (lower)

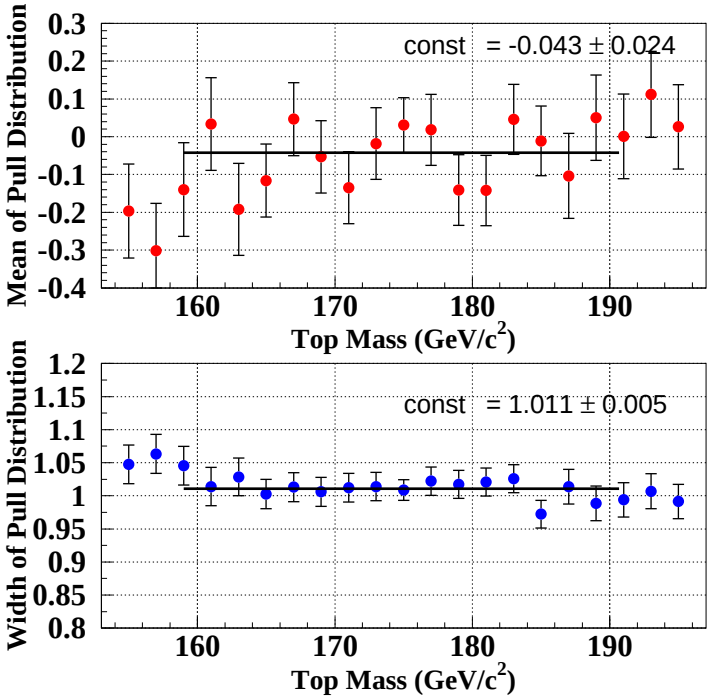

Figure 8.6: Mean (upper) and width (lower) of the pulls versus generated top mass

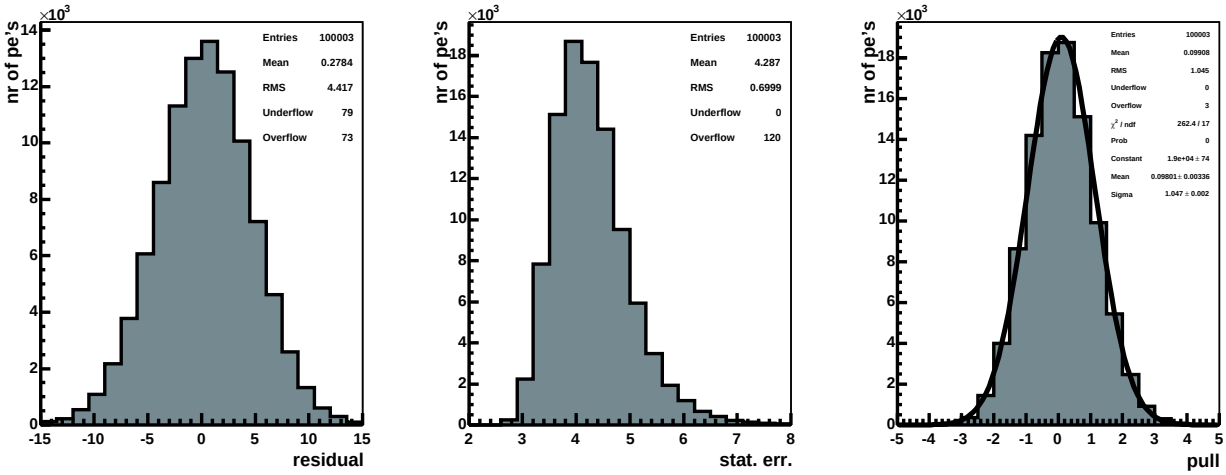

Figure 8.7: Blind samples results: bias (left), statistical error (middle), pulls (right plot)

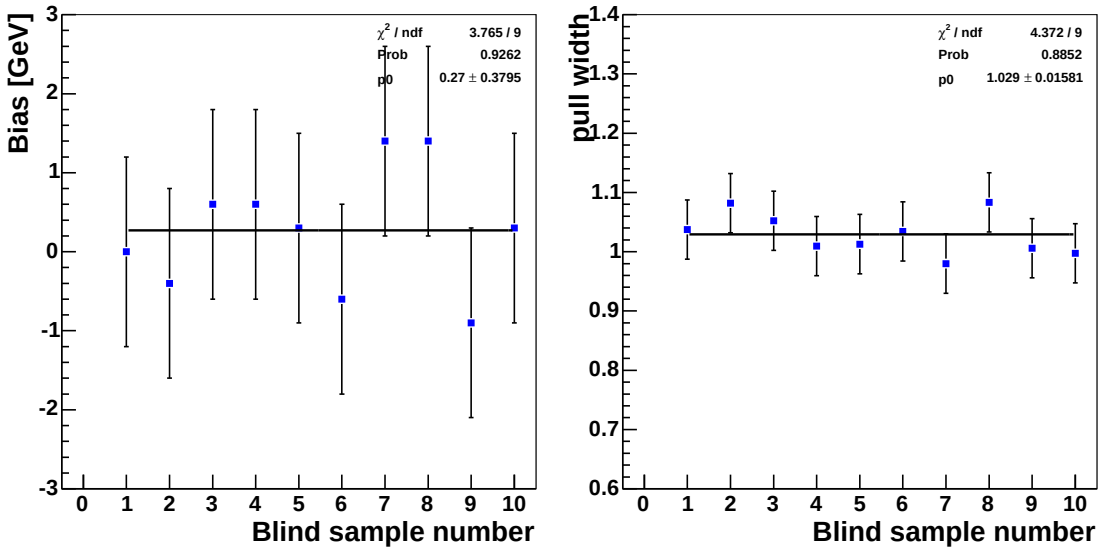

Figure 8.8: Blind samples results: bias (left), width of the pulls (right)

8.2.4 Summary of corrections

The method has been tested according to the standard procedures at CDF. Results from these tests show that outputs of the method has to be corrected according to the following table:

Parameter	correction
Top mass	$-0.16\ GeV/c^2$ (shift)
Statistical error	1.011 (factor)

Table 8.2: Shift applied to the top mass and multiplication factor applied to statistical error.

Chapter 9

Systematic Uncertainties

Besides the statistical uncertainty, there are several sources of systematic uncertainty on the top mass.

Since our method compares findings to Montecarlo expectations discrepancies between simulated and data events originate systematic uncertainties (sections 9.1 to 9.6). Other systematic uncertainties originate from the finite statistic used to build the simulated background templates (section 9.7).

The procedure for estimating systematic uncertainties is as follows. Some parameters used to simulate events (e.g. Jet energy scale, amount of the initial and final state radiation, etc.) are shifted [1] around their central value and shifted templates are generated. Top masses are measured using the shifted templates and the difference from default values is used to estimate the systematic uncertainty.

The sample (ttkt75) with a generated top mass of 175 GeV/c^2 is used to represent $t\bar{t}$ events in the mass range of interest.

9.1 Jet Energy Scale

The calorimeter energy assigned to jets is corrected according to the calibrated detector response to electrons and hadrons and for physics effects. Six levels of jet energy corrections have been considered, each of which have associated uncertainties σ (4.1.4). The procedure to account for these uncertainties is as follows[2]:

- In event selection and reconstruction each level is shifted by the $\pm\ \sigma$

[1] σ is the standard deviation on the shifted parameter. The values of these parameters will be reported later.

[2] Drell-Yan and Diboson samples which account for about 87% of the overall background are used (see table 7.2). Since fakes background has been estimated on data, it is not included in this systematic.

- Shifted signal and background templates ($T^{\pm}$) are obtained.
- Two sets of 10^4 PE's are performed by using $T^{\pm}$ (Chapter 8): shifted top masses ($M^{\pm}$) are estimated.
- $(M^{+} - M^{-})/2$ is taken as JES systematics.

These 6 contributions are finally summed in quadrature in order to obtain the overall *Jet Energy Scale (JES) systematic uncertainty*. Results are reported in table 9.1.

Reconstructed top mass distributions from PE's where level 5 jet energy correction was shifted $\pm\,\sigma$ is shown in figure 9.1.

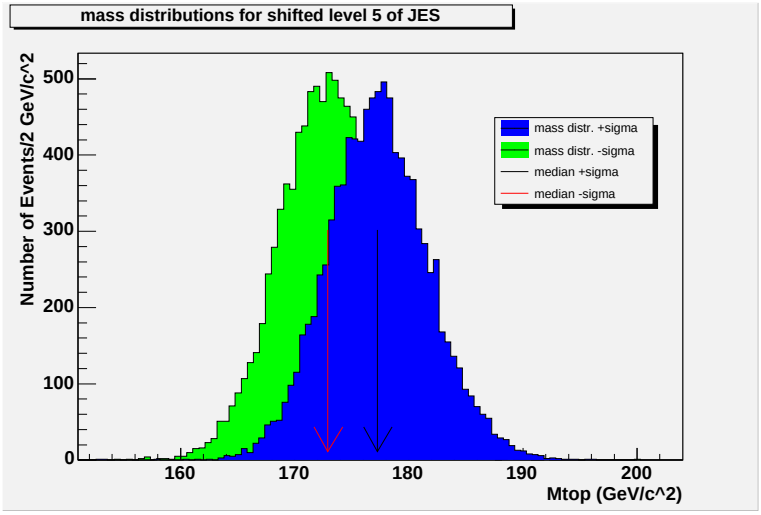

Figure 9.1: Top mass distributions from PE's. T^{-} (green) and T^{+} (blue) for level 5 (see text for symbol significance). Arrows indicate the medians of the distributions.

Level 5 corrects for the absolute calorimeter energy scale. Its uncertainty dominates the overall JES uncertainty.

Systematic uncertainties for each level and the total one are shown in table 9.1.

The total Jet Energy Scale uncertainty is $\mathbf{2.9\ GeV/c^2}$.

Despite of long efforts made to improve the understanding of the energy dependence of calorimeter response this source of uncertainty is still dominating in most of the jet analyses performed in CDF [3].

[3] Sensitivity of top mass to jet energy comes about because the leading-jets originate directly from $t(\bar{t})$ decay.

Level	Source	Mass (GeV/c^2)		Uncertainty (GeV/c^2)
		$+\sigma$	$-\sigma$	$\Delta M/2$
1	η-Dependent	175.80	174.58	0.61 $\pm$0.03
4	multiple interactions	175.20	175.16	0.02 $\pm$0.03
5	Absolute Scale	177.26	172.93	2.17 $\pm$0.03
6	Underlying Event	175.37	175.05	0.16 $\pm$0.03
7	Out-of-cone	176.99	173.36	1.81 $\pm$0.03
4	splash-out	175.49	174.94	0.28 $\pm$0.03
total				**2.9$\pm$0.03**

Table 9.1: Mass shifts when a $\pm 1\sigma$ shift is applied to signal and background (see text). Last row reports the total JES systematic uncertainty.

9.2 b-Jet Energy Scale

JES uncertainties are estimated with studies on samples dominated by light-quark and gluon jets [60]. b-jets (jets initiated by a b quark) differ from light quarks jets because of three features [62]:

1. *Heavy Flavor Fragmentation*: the b-hadron resulting from b-quark fragmentation carries a larger fraction of the parent quark momentum than hadrons born in light quark fragmentation [4].

2. *Colour Flow*: b-jets experience a mechanism of color compensation which is channel specific.

3. *Decay*: A large fraction of the decay products of b-hadrons are leptons and neutrinos [5]

The shift on top mass due to b-JES uncertainty is estimated by applying a $\pm$ 1% shift to b-jet energy in the signal sample. Two shifted masses are then estimated. Table 9.2 shows the results.
According to the standard CDF rules the obtained $\Delta M/2$ is multiplied by 0.6 [6]: the systematic uncertainty due to b-jet energy scale is **0.62 GeV/c^2**.

[4] An intuitive understanding of this effect is that, being much heavier than u,d,s,c, the b-quark is only slightly decelerated when combined with a light quark to form a b-hadron

[5] Since leptons and neutrinos interact in a different way with the calorimeter than the more common hadronic particles, b-jets will have a different response on average than W-jets. Although this difference has been modeled in event reconstruction (see 6.2) uncertainty on leptonic branching fractions cause additional uncertainty.

[6] This factor is obtained from simulations. Simulations have shown that 1 GeV/c^2 shift in top mass is returned by 1% shift in b-JES and that the b-JES systematic is expected to be 0.6 GeV/c^2 for all top mass analysis [61, 62]. A more correct procedure would be to apply a shift

Source		Datasets	Mass GeV/c^2	Mass Shift GeV/c^2	Syst. GeV/c^2
B-JES	$\pm$ 1% shift in JES for b-jets	ttkt75(-) ttkt75(+)	174.15$\pm$0.04 176.24$\pm$0.04	$\Delta M/2$ =1.04$\pm$0.03	0.62

Table 9.2: Results from PE's by applying a $\pm$1% b-jet energy scale shift in signal events (see text).

b-JES systematic uncertainty is the largest systematic after the JES one. Since b quarks are direct top products, the top mass is very sensitive to the uncertainty on their energy scale.

9.3 Lepton energy scale

The top mass uncertainty associated to the uncertainty of lepton energy scale has been studied by applying $\pm$1% shifts to the lepton p_T [7] [72]. Results are shown in table 9.3.

Source		Datasets	Mass GeV/c^2	Mass Shift GeV/c^2	Syst. GeV/c^2
l-JES	$\pm$ 1% shift in ES for leptons	ttkt75(-) ttkt75(+)	174.84$\pm$0.04 175.48$\pm$0.04	$\Delta M/2$ =0.3$\pm$0.03	0.3

Table 9.3: Results from PE's by applying a $\pm$1% shift to the lepton momentum scale shift in signal events (see text).

Half difference between the two shifted results (0.3 GeV/c^2) is assumed as systematic error from lepton energy scale uncertainty.

9.4 MC generators

As mentioned in Chapter 7, by default MC signal samples are generated with the Pythia generator. However, at CDF two different $t\bar{t}$ generators, Pythia and

to the b-JES in the simulated Heavy Flavor Fragmentation, Color Flow and Decay and add the resulting mass shift in quadrature. However, this procedure would be much more time consuming and is not expected to give significantly different estimates of the systematic uncertainty.

[7] In this case leptons are defined as those tracks or electromagnetic objects that match generated leptons

Herwig, are available. These generators use different models for the hadronization, underlying events and multiple interactions [63].

MC generators systematic takes into account differences which arise by using signal samples, generated with Herwig rather than Pythia. Signal sample (top mass 175 GeV/c^2) generated with Herwig (htop75) and Pythia (ttkt75) are compared. Table 9.4 summarizes the results from PE's.

Source		Datasets	Mass GeV/c^2	Mass Shift GeV/c^2	Syst. GeV/c^2
Generator	Pythia	ttkt75	**175.17±0.15**	ΔM =0.46±0.35	0.46
	Herwig	htop75	174.71±0.32		

Table 9.4: Results from PE's for different Monte-Carlo generators. See text for significance of symbols.

0.46 GeV/c^2 is the estimated systematic of MC generators on top mass.

9.5 Initial and final state radiation

Jets radiated from interacting partons in the initial and final state affect the top mass measurement because they can be misidentified as leading jets or can change the kinematics of the final state partons. In order to study the effect on the parametrization of initial (ISR) and final (FSR) state radiation four samples ($+\sigma_{ISR}, -\sigma_{ISR}, +\sigma_{FSR}, -\sigma_{ISR}$) are generated [8]. Top mass reconstruction is performed on these samples: results are summarized in table 9.5.

Source		Datasets	Mass GeV/c^2	Mass Shift GeV/c^2	Syst. GeV/c^2
ISR	$+\sigma_{ISR}$	itoprk	175.00±0.32	ΔM =-0.17±0.34	0.34
	$-\sigma_{ISR}$	itoprl	174.95±0.32	ΔM =-0.22±0.34	
FSR	$+\sigma_{FSR}$	ftoprj	175.05±0.32	$\Delta M/2$ =0.17±0.22	0.22
	$-\sigma_{FSR}$	ftoprl	175.38±0.32		

Table 9.5: Top mass shifts obtained from PE's. See text for significance of symbols.

[8] $\pm\sigma_{ISR/FSR}$ corresponds to an increase or decrease of the ISR/FSR in simulated events. See Appendix J for details on the generation of these samples.

ISR mass shift is below the nominal value (175.17) both for $+\sigma_{ISR}$ and $-\sigma_{ISR}$. However, the shifts are smaller than the errors [9]. According to CDF rules under such circumstances the error, 0.34, is used as ISR contribution to the systematic uncertainty.

0.22 GeV/c^2 is taken as FSR systematic uncertainty.

9.6 Parton distribution functions

Monte Carlo samples are generated by using parton distribution functions (PDF's) named CTEQ5L [70]. Discrepancy in top mass arising from different PDF's must be estimated.

The PDF uncertainty is composed of three sources [73]:

- PDF parametrization.
- PDF choice.
- QCD scale (Λ).

In order to estimate these three sources of uncertainty the following PDF's have been used:

- CTEQ5L: default PDF.
- CTEQ6M: default next-to-leading order PDF from CTEQ group [70].
- MRST72: PDF from MRST group [71] with the same QCD scale as CTEQ5L
- MRST75: Same as MRST72 but with a different Λ [10]

These uncertainties are estimated by the "re-weighting method" [74]: a PDF-dependent relative weight to CTEQ5L [11], is assigned to each event. Templates are re-built and PE's are performed. Results are shown in figure 9.2.

The uncertainty on PDF parametrization is estimated by shifting $\pm$ σ the 20 eigenvectors of CTEQ6M PDF (red and blue dots in figure 9.2). Half differences in the derived mass are added in quadrature.

The difference in top mass between CTEQ5L and MRST72 is used as systematic uncertainty associated to choosing one particular PDF. Difference in top mass

[9] This indicates that our method is insensitive to initial state radiation

[10] MRST75 is generated with $\Lambda = 300\ MeV$, while MRST72 with $\Lambda = 228\ MeV$

[11] This approach is less CPU time consuming than generating MC events according to all PDF's listed above

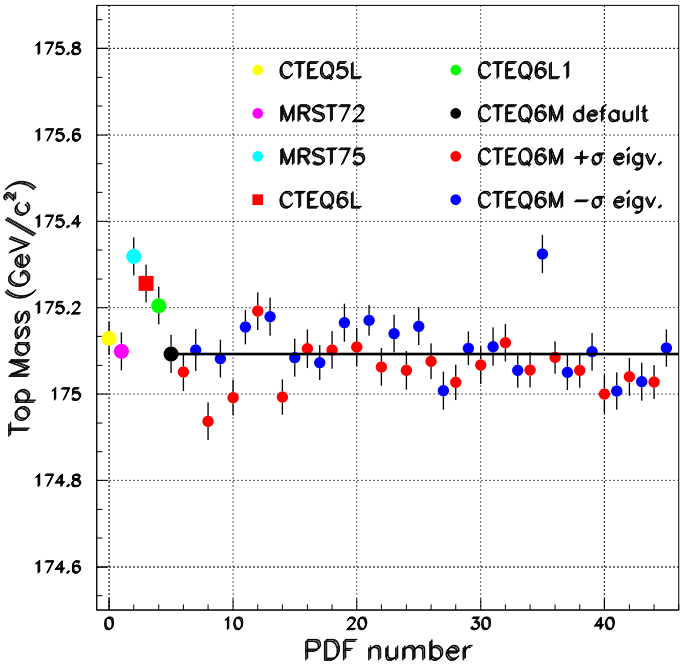

Figure 9.2: top mass reconstruction using different PDF's. Black horizontal line is the nominal value for CTEQ6M PDF.

between MRST72 and MRST75 is the systematic uncertainty for a different QCD scale.

These systematic uncertainties are then added in quadrature. Results are summarized in table 9.6:

Source	Mass Shift GeV/c^2	Syst. GeV/c^2
CTEQ PDFs	Sum($\Delta M/2$)=0.21±0.13	0.21
CTEQ5L vs MRST72	ΔM=0.07±0.06	0.07
Λ (MRST72 vs MRST75)	ΔM=-0.22±0.06	0.22
Total		0.31

Table 9.6: PDF systematic uncertainties. See text for significance of symbols.

0.31 GeV/c^2 is the PDF systematic uncertainty.

9.7 Background shape

Another source of systematic uncertainty is due to the possible mis-modeling of the background shape.

Three independent sources may contribute significantly to the mis-modeling:

1. The background composition
2. The Fakes shape
3. The Drell-Yan shape

9.7.1 Background composition

The expected Diboson, Drell-Yan and Fakes number of events are varied alternatively by $\pm\sigma$ (see table 7.3), without changing the total background number. Three $\pm$ shifted background templates are obtained and used for PE's. Results are presented in Table 9.7.

0.46 GeV/c^2 is the systematic uncertainty for the background composition.

9.7.2 Fake events shape

Fakes represent the largest sources of background (see table 7.3). Effect on top mass coming from uncertainty on fake shape is studied.

New fake templates are built by applying a ± linear E_T-dependent shift in the fake rate matrix (see figure C.1) [12].

The two obtained fake templates are combined with the other background templates, according to the procedure described in section 7.2.1. PE's from these templates are performed. The result is shown in table 9.7: 0.41 GeV/c^2 is the systematic uncertainty related to the Fakes shape.

9.7.3 Drell-Yan events shape

Drell-Yan events can pass the LTRK selection because jet mismeasurements can cause a large missing E_T. Mismodeling of this effect may shift the reconstructed top mass. In order to estimate the sensitivity of our measurement to this effect, the weight in $Z/\gamma^* \to ee,\ Z/\gamma^* \to \mu\mu$ [13] templates has been increased (and decreased) by a factor of 2. Only the contribution of events with tight lepton and track lepton inside Z window [14] has been considered for re-weighting [15]. Two new combined background templates have been obtained and used for PE's. The result is shown in table 9.7. 0.3 GeV/c^2 is the systematic uncertainty related to Drell-Yan events shape.

9.8 Comments on errors

As already mentioned in chapter 8, errors on PE results are estimated according to the following formula:

$$\sigma = \sigma_{LS} \oplus \frac{\sigma_{PE}}{\sqrt{N_{PE}}} \tag{9.1}$$

where σ_{PE} is the average standard deviation from the single PE, $N_{PE} = 10^4$ and σ_{LS} is the uncertainty because of limited statistic (8.2.1).

In order to calculate errors on individual systematic uncertainties (see tables 9.1-9.7) formula 9.1 has been interpreted as follows:

- $\sigma = \frac{\sigma_{PE}}{\sqrt{N_{PE}}}$ if "$\pm\sigma$ shifted" top masses are obtained from the same signal sample. This is the case of *JES*, *B-JES*, *l-ES* and *background shape* systematic uncertainties.

[12] The shift has the following form: $w^i_\pm = w^i_0 \pm [\frac{1}{8} \cdot (2i_{E_T} - 8) \cdot \sigma^i_w];\ i = 0, \ldots, 8$, where w^i_0 is the fake rate in the ith E_T bin, σ^i_w is the uncertainty on w^i_0, and i_{E_T} denotes the E_T bin (see figure C.1)

[13] $Z \to \tau\tau$ events can have real $\not\!\!E_T$ because of τ-decay neutrinos. For this reason they are not considered in the background shape systematic.

[14] see section 4.4 for Z-window definition

[15] For details about the factor used for re-weighting see [51]

Source		Mass GeV/c^2	Mass Shift GeV/c^2	Syst. GeV/c^2
BG	Diboson(-σ)	175.15±0.04	$\Delta M/2$ =-0.02±0.03	0.03
compos-	Diboson(+σ)	175.11±0.04		
ition	DY(-σ)	174.90±0.04	$\Delta M/2$ =0.25±0.03	0.25
	DY(+σ)	175.40±0.04		
	Fakes(-σ)	175.53±0.04	$\Delta M/2$ =-0.39±0.03	0.39
	Fakes(+σ)	174.74±0.04		
				0.46
Fake	-linear E_T-dependent shift	174.80±0.04		
shape	in fake rate matrix		$\Delta M/2$ =0.41±0.03	0.41
	+linear E_T-dependent shift	175.61±0.04		
	in fake matrix			
Drell-Yan	decreased weight	174.88±0.04		
shape	in the Z window		$\Delta M/2$ =0.30±0.03	0.30
	increased weight	175.48±0.04		
	in the Z window			

Table 9.7: Top mass shifts obtained from the PE for different BG composition, Fakes and Drell-Yan shape. See text for more details.

- $\sigma \approx \sigma_{LS}$ otherwise. This is the case of MC *generator*, *ISR* and *FSR* systematic uncertainties [16].

9.9 Summary of the systematic uncertainties

The total systematic error together with the single contributions are shown in table 9.8.
The systematic uncertainty on top mass is 3.1 GeV/c^2. It is obtained by adding in quadrature the single contributions [17].

[16] The bootstrap procedure used to estimate σ_{LS} accounts for the finite statistic signal sample. In table 9.4 the error on top mass from Pythia sample is ~ 0.15, as shown in figure 8.2. All other errors are ~ 0.32 since they are estimated from samples of an about twice smaller statistic.

[17] Fully uncorrelated sources have been assumed.

CDF RunII Preliminary

Source	Uncertainty (GeV/c^2)
Jet Energy Scale	2.9
b-JES	0.6
Initial State Radiation	0.3
Final State Radiation	0.2
Parton Distribution Functions	0.3
Monte-Carlo Generators	0.5
Background composition	0.5
Fakes shape	0.4
DY shape	0.3
l-ES	0.3
Total	**3.1**

Table 9.8: Summary of systematic uncertainties

9.10 Conclusions

Ten sources of systematic uncertainty have been included in the systematic error on top mass. At present (winter 2007) CDF physicists are working towards two goals:

- Decrease the systematic uncertainties with improved analysis techniques on larger statistic.
- Account for minor previously neglected sources of systematic uncertainty (e.g. top width uncertainty) which with increasing statistic are becoming sources of significant errors.

Figure 9.3 is an overview of the estimated dependence of statistical and systematic errors on top mass versus the integrated luminosity by combined CDF and DØ.

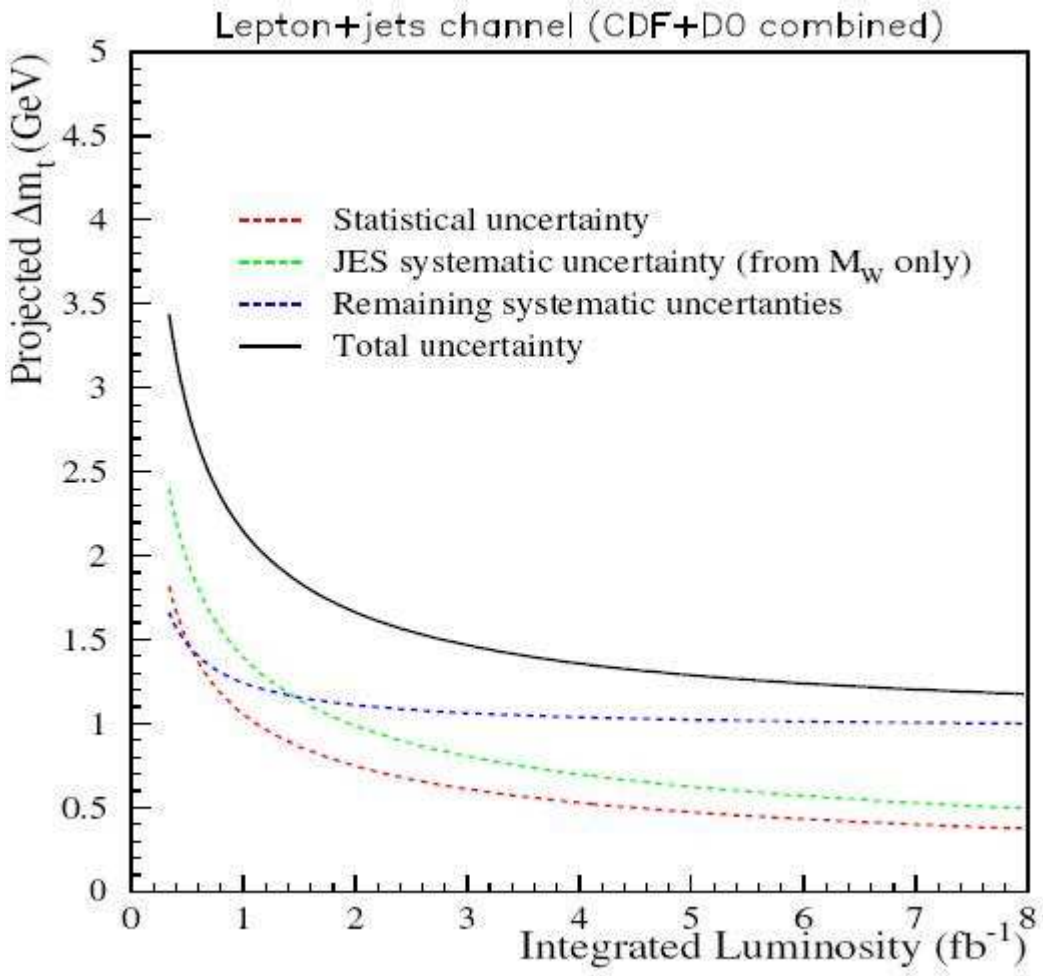

Figure 9.3: Projection of statistic and systematic uncertainties on top mass measurement as a function of integrated luminosity [77]. CDF and DØ combined results in the single lepton channel are assumed. Only the residual JES calibration error after imposing the W mass to jet pairs associated to the W is accounted for.

Chapter 10

Results

After having tested the method (chapter 8) it is possible to apply it on data events and measure the top mass. Section 10.4 reports our result.

10.1 Data events

The data sample corresponds to an integrated luminosity of 2.1 fb^{-1} [1]. Accordingly, the expected signal and total background events presented in table 8.1 have been obtained by rescaling to this luminosity (see table 10.1).

	expected events (2.1 fb^{-1})
Signal	118.6 ± 3.7
Background	105.8 ± 12.9

Table 10.1: Expected signal and background events for an integrated luminosity of 2.1 fb^{-1}

236 LTRK candidates have been found in the data.

10.2 Fit Results

LTRK events were reconstructed and an experimental mass distribution was built. A background constrained likelihood fit (see 6.15) [2] was performed and the following estimates of top mass and of the positive and negative statistical errors were

[1] The MC samples were tuned to data up to $\int Ldt = 1.2 fb^{-1}$, and the validity of the extrapolation method was tested for data collected up to 2.0 fb^{-1}.

[2] Constraining parameters are $n_b^{exp} = 105.8$ and $\sigma_b^{exp} = 12.9$

obtained:

$$M_{top} \;=\; 167.57^{+4.1}_{-4.0} \tag{10.1}$$

The experimental top mass distribution is shown in figure 10.1.

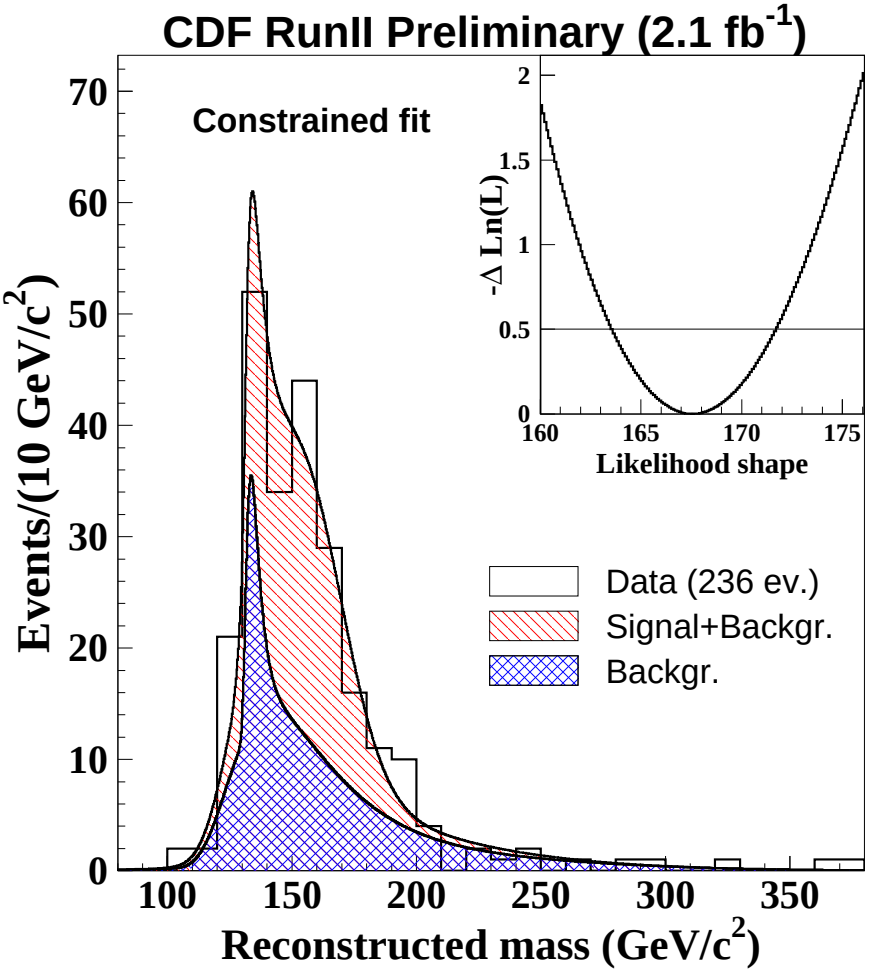

Figure 10.1: Two-component constrained fit to the LTRK sample. The blue shaded area is the background returned by the fit and the red shaded area is the sum of background and signal. The insert shows the fitted mass-dependent negative log-likelihood function.

The fitted function shows the background (blue) and signal contributions to the 236 events. Signal and background numbers of events are reported in table 10.2.

	Constrained Fit returned numbers ($2.1\ fb^{-1}$)
Signal	$126.1^{+18.0}_{-17.4}$
Background	$108.3^{+11.6}_{-11.5}$

Table 10.2: Signal and background events from the constrained fit.

The observed rates (table 10.2) are in agreement with expectations (table 10.1).

As a check, a fit leaving free the background events has been performed [3]. This fit returns:

$$M_{top} \;=\; 167.59^{+4.4}_{-4.2} \tag{10.2}$$

with $117.82^{+26.48}_{-26.59}$ signal events and $118.18^{+27.53}_{-25.52}$ background events. Top mass estimate, signal and background from unconstrained and constrained fit are in agreement.

Figure 10.2 shows the top mass distribution with the unconstrained best-fit function superimposed.

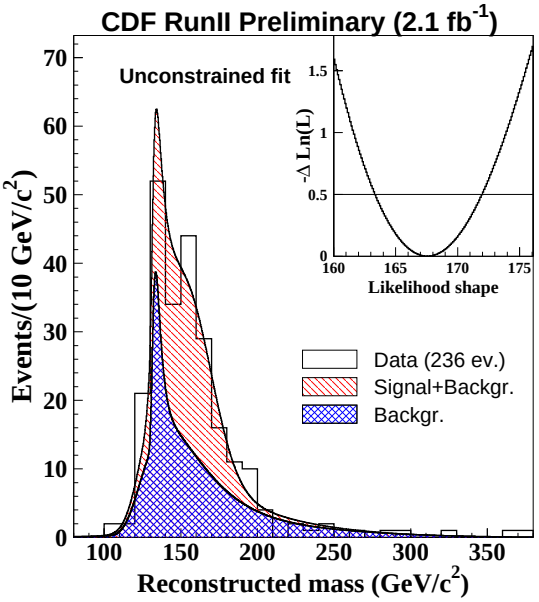

Figure 10.2: Two-component unconstrained fit to the l+trk sample. The blue shaded area is to the background returned by the fit and the red line-shaded area is the sum of background and signal. The insert shows the fitted mass-dependent negative log-likelihood function.

[3] 6.15 without the term $\mathscr{L}_{backgr}$

10.3 Comments on the statistical errors

The expected error distribution for a top mass of 167 GeV/c^2 is shown in figure 10.3. The red arrows show the negative and positive errors found in the experiment.

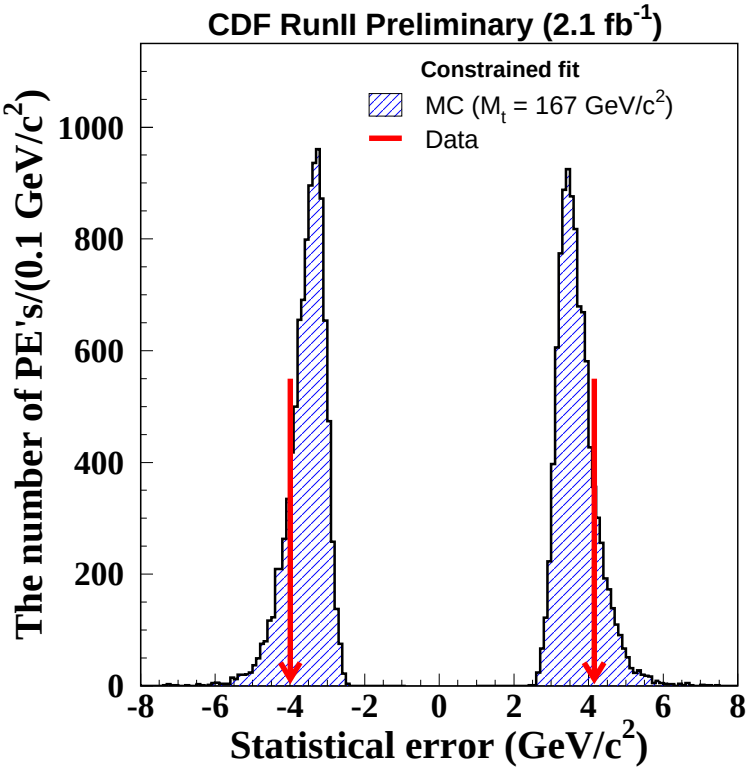

Figure 10.3: Expected statistical errors for a top mass of 167 GeV/c^2. The arrows indicate the errors found in this measurement.

The probability for a better accuracy is 83%.

10.4 Results

The top mass and its statistical errors as output of the background constrained fit are corrected by the expected systematic -0.16 GeV/c^2 shift and 1.011 factor (see table 8.2) respectively. The resulting final value of the top mass is:

$$M_{top} \;=\; 167.7\,^{+4.2}_{-4.0}(stat.) \pm 3.1(syst)\ GeV/c^2$$

Conclusions and Prospects

We have measured the mass of the top quark in the first 2.1 fb^{-1} data collected by the CDF II detector, analyzing top pair events in the dilepton channel. The result is:

$$M_{top} \;=\; 167.7\ ^{+4.2}_{-4.0}(stat.) \pm 3.1(syst)\ GeV/c^2 \tag{10.3}$$

consistent with the Tevatron combined mass ($M_{top} \;=\; 170.9 \pm 1.8\ GeV/c^2$) using all top mass measurements of the CDF and $D\emptyset$ Collaborations up to 2007.
This measurement is also in agreement with the previous result obtained with the same method on a 340 pb^{-1} data sample [45]:

$$M_{top} \;=\; 169.7\ ^{+8.9}_{-9.0}(stat.) + 4.0(syst)\ GeV/c^2 \tag{10.4}$$

A significant improvement in resolution has been obtained. We summarize below the novelties which made this progress possible.

- A larger available statistics. In addition to using a larger data sample, this was because candidates were selected with the LTRK selection (4.4) which is less restrictive than the previous DIL.
- A new algorithm to estimate the "Fakes" contribution to the background (Appendix C) was implemented.
- In the event reconstruction
  - Relativistic Breit-Wigner distribution functions and a M_{top}-dependent top width, as suggested by the Standard Model, were introduced. This allowed a smaller statistical error to be returned by the mass fit.
  - Transfer functions (appendix I) were used for a more correct description of b parton response.

At the moment this measurement has the second best accuracy among mass measurements in the dilepton channel. The same machinery as described in this thesis is expected to result in a statistical uncertainty of $\sim 2.4\ GeV/c^2$, by assuming with about 6 fb^{-1} data as expected to be available in 2010

However, we note that this measurement can be further improved. We are planning to include the following features in the next analysis:

- Select events based on optimized neural networks for the mass measurement, in order to increase significantly the signal to background ratio [81].
- Exploit b-flavor tagging. Tagged and non-tagged events can be analyzed separately and the results can be combined. This technique was proven to lead to an improvement of about 20% in the statistical error [45, 51]

At the moment (Winter 2007) the CDF internal procedure for the publication of this measurement has been started.

Appendix A

Unknown in the dilepton Channel

Particle	P_x	P_y	P_z	E	# Unknown
t,$\overline{t}$	NM	NM	NM	NM	8
ν,$\overline{\nu}$	NM	NM	NM	NM	8
W^+, W^-	NM	NM	NM	NM	8
b,$\overline{b}$	M	M	M	M	0
l^-, l^+	M	M	M	M	0
l^-, l^+	M	M	M	M	0
Underlying event	M	M	NM	NM	2

Table A.1: Measured (M) and non-measured (NM) parameters in the dilepton channel.

Appendix B

DIL Selection

The DIL selections requires two well-identified high-P_T leptons, a large missing E_T and at least two energetic jets in the events. The guidelines of the DIL which was adopted in the past are going to be described below in order to point out the differences with the LTRK selection.

B.1 Requirements on the Physical Objects

The same definition as in the LTRK selection is adopted in DIL for the **first (tight) lepton** (see 4.4.2) [1].

The **second (loose) lepton** must pass the same identification cuts as for the tight lepton with the following exceptions:

- if CEM no isolation is required;
- If muon the muon chamber hit requirement is relaxed [2].

Jets are required to have $E_T^{corr} > 15\ GeV$, where E_T^{corr} is corrected up to L5 (see 4.1.4), and $|\eta| < 2.5$. These jets are called tight jets.

The corrected $\not\!\!E_{\mathbf{T}}$ [3] is required to be larger than 25 GeV.

B.2 Vetoes

- Vetoes Cosmic and conversion removal: see 4.4.3
- Z veto cut: if the invariant mass of the two leptons is in the Z window ([76, 106] GeV/c^2) the following requirements must be satisfied:

[1] No PHX (Chapter 4) electrons used.

[2] More details about loose muon DIL cuts are available at [48, 49]

[3] Corrections applied are much the same as in LTRK. See [48, 49] for further details.

1. $JetSig \equiv \frac{\not\!\!E_T}{\sqrt{\sum_i E_{T_i}} \hat{\not\!\!E}_T} > \sqrt{8}\ GeV$, where the sum is over all tight jets in the same hemisphere as the $\not\!\!E_T$.
2. $dPhiJet > 10^\circ$, where dPhiJet is the minimum azimuthal angle between $\not\!\!E_T$ and the jets.

- L cut: $\not\!\!E_T > 50\ GeV$ is required if the minimum azimuthal angle between missing E_T and any lepton or jet is smaller that 10°.
- $H_T \equiv \sum_i E_{T_i} > 200\ GeV$ cut, where the sum runs over $\not\!\!E_T$, tight and loose leptons and tight jets.
- OS cut: see 4.4.3

More details about DIL selection are described in [48, 49].

Appendix C

Fake Events ("Fakes")

The definition of fake event was given in 4.3.2. A fake LTRK event can be due to a jet faking either the tight lepton or the track lepton. Since the former phenomenon has a very small rate [1] only the latter is taken into account in this analysis.

Fakes are estimated as follows.

1. A tailored W+jets selection on the high-P_T data sample listed in table E.1 is performed. Main steps of the W+jets are the following [2]:
   - Identification of a tight lepton and a large missing E_T (corrected $\not\!\!E_T > 25\ GeV$).
   - Cosmic and conversion removal, $\Delta\phi(\not\!\!E_T, tl)$, $\Delta\phi(Jet - \not\!\!E_T)$ cuts
   - Veto on the LTRK events
2. A collection ("djets") of fakeable objects composed of jets and tracks is created. Jets or tracks matching the tight lepton are not taken into account.
3. Each djet is tested as a trkl. The following cuts are then applied.
   - $\Delta\phi(\not\!\!E_T, trkl)$
   - Z veto/$\not\!\!E_T$
   - OS (if available [3]).
4. Two or more jets not matching the tl and the trkl with $E_T > 20\ GeV$ and $|\eta| < 2$ are required

[1] The tight lepton identification cuts are much more severe than for the track lepton.
[2] Notation refers to 4.4
[3] No charge information is available for jets

It should be noticed that more fakes per event may be signaled by this procedure.

At the end to each returned event a weight is attached according to a fake rate matrix, which expresses the probability that a djet fakes the track lepton. The fake rate matrix (see figure C.1) is function of the djet E_T [4].

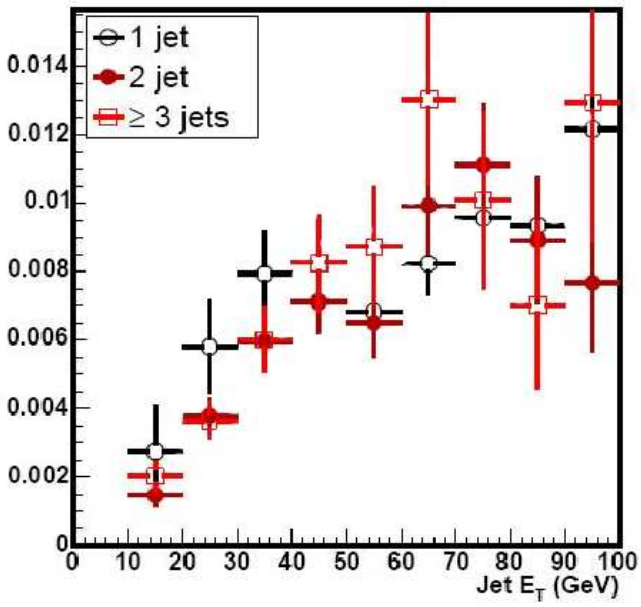

Figure C.1: The track lepton fake rate matrix from $\gamma + jets$ data, using a photon with energy more that 80 GeV, and Z+1jet data [47]. Contributions for a different number of jets are shown.

[4]The fake rate matrix is quite independent of η. See [47] for details.

Appendix D

Requirements of the Exploited Trigger Paths

A trigger path, composed of three levels, is characterized by a sequence of requirements on the physical objects. Trigger paths used in this analysis are the following[1]:

- ELECTRON_CENTRAL_18:
  - L1: one central tower with $E_T > 8\ GeV$, $P_T^{XFT} > 8.34\ GeV/c$, where P_T^{XFT} is the transverse momentum of the associated XFT track
  - L2: one cluster with $E_T^{em} > 16\ GeV$, $|\eta^{XFT}| < 1.317$, $P_T^{XFT} > 8.34\ GeV$
  - L3: one CEM electron of $E_T > 18 GeV$ and $P_T^{COT} > 9\ GeV/c$, where P_T^{COT} is the transverse momentum of the COT track
- MUON_CMUP18
  - L1: one CMU stub with $P_T > 6\ GeV$, associated CMP stubs and an XFT track with $P_T^{XFT} > 4.09\ GeV$
  - L2: one CMUP muon with $P_T^{XFT} > 8.34\ GeV$
  - L3: one CMUP muon of $P_T > 18\ GeV$
- MUON_CMX18
  - L1: one CMX stub with $P_T > 6\ GeV$, trigger signal in CSX and an associated XFT track with $P_T^{XFT} > 8.34\ GeV$

[1]The notation is as defined in the Chapter 3. The requirement on EM objects assumes the constraint $E_T^{had}/E_T^{em} < 0.125$

- L2: one CMX muon of $P_T^{XFT} > 8.34\ GeV$
- L3: one CMX muon of $P_T > 18\ GeV$

- MET_PEM:
  - L1: one CEM or PEM tower with $E_T > 8\ GeV$ and a raw $\not\!\!E_T > 15\ GeV$
  - L2: one PEM cluster with $E_T^{em} > 20\ GeV$
  - L3: one PEM electron of $E_T > 20 GeV$ and an offline $\not\!\!E_T > 15\ GeV$

Appendix E

Exploited Samples and Good Run Lists

High-P_T data samples and good runs [1] used in this analysis are listed in the tables E.1 and E.2.

Triggered lepton type	Sample
High-P_T Central Electrons	bhel0d bhel0h bhel0i bhel0j
High-P_T Central Muons	bhmu0d bhmu0h bhmu0i bhmu0j
High-P_T Plug Muons	bpel0d bpel0h bpel0i bpel0j

Table E.1: Data samples labelled according to the triggered lepton type.

[1] see 4.4.1 for the meaning of the good run lists

Tight lepton type	Good Run List	Version
CEM	em_nosi	v13, v17, v18
CMUP	mu_nosi	v13, v17, v18
CMX	mu_nosi	v13, v17, v18
PHX	em_si	v13, v17, v18

Table E.2: Good Run lists for each tight lepton sub-sample. The conditions to be satisfied by good runs have changed with increased luminosity (versions 17 and 18).

Appendix F

χ^2 form and solution weights in [44]

The χ^2 defined in the analysis [44] is:

$$\chi^2 = \chi^2_{reso} + \chi^2_{constr} \tag{F.1}$$

$$\chi^2_{reso} = \sum_{l=1}^{2} \frac{(P_T^l - \tilde{P}_T^l)^2}{{\sigma^l_{P_T}}^2} + \sum_{j=1}^{2} \frac{(P_T^j - \widetilde{P_T^j})^2}{{\sigma^j_{P_T}}^2} + \sum_{i=x,y} \frac{(UE^i - U\tilde{E}^i)^2}{{\sigma^i_{UE}}^2} \tag{F.2}$$

$$\chi^2_{constr} = \frac{(M_{l_1\nu_1} - M_W)^2}{\Gamma^2_{M_W}} + \frac{(M_{l_2\nu_2} - M_W)^2}{\Gamma^2_{M_W}} + \frac{(M_{j_1 l_1 \nu_1} - \widetilde{m_t})^2}{\Gamma^2_{M_t}} + \frac{(M_{j_2 l_2 \nu_2} - \widetilde{m_t})^2}{\Gamma^2_{M_t}} \tag{F.3}$$

Most notations have been already defined in section 6.1.2. More information is given below.

The second term in F.2 sums over the transverse momenta P_T^j of the two leading jets. These momenta has been corrected for underlying event and out-of-cone energy (see 4.1.4), and have a P_T and η dependent detector resolution $\sigma^j_{P_T}$ derived by simulation [1]. F.3 takes into account Gaussian constraints for the W and top decay chains. The top width is assumed to be 2.5 GeV/c^2.

[1] further details available in [45]

The solutions per each ϕ_1, ϕ_2 point of the net are weighted according to $w_{ij} = e^{-\frac{\chi^2_{ij}}{2}}$ (see 6.1.2).

Appendix G

Impact of ϕ_1-ϕ_2 correlation

As described in Chapter 6, the Neutrino ϕ Weighting Method chooses to make assumptions about the azimuthal angles, $\phi_{\nu 1}, \phi_{\nu 2}$, of the neutrinos. In this Chapter we show that the method is not affected by the correlation (see figure G.1) between $\phi_{\nu 1}, \phi_{\nu 2}$.

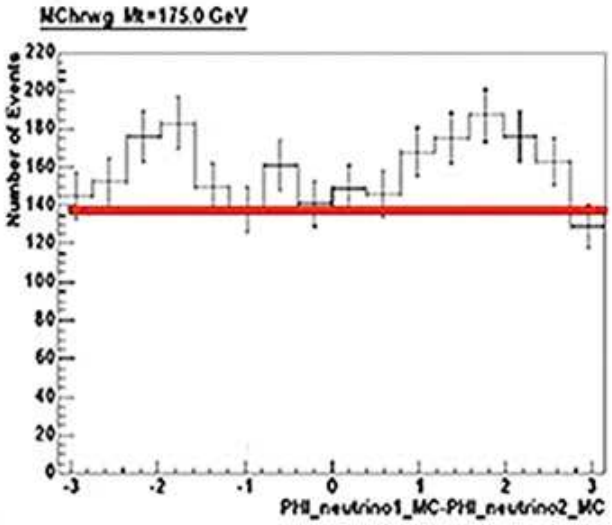

Figure G.1: Difference between the azimuthal angles of the neutrinos expected from simulation. A dilepton DIL (see Appendix B) sample has been used with top mass of 175 GeV/c^2

To check that this is true we draw bi-dimensional distributions as a function of $\phi_{\nu_1}^{MC} - \phi_{\nu_1}^{i}$ and $\phi_{\nu_2}^{MC} - \phi_{\nu_2}^{j} \quad i, j = 1, \ldots, 24$, where $\phi_{\nu_1}^{MC}$ and $\phi_{\nu_2}^{MC}$ are the azimuthal angles extracted directly from hepg bank, while the other two sets of values are

those over which we iterate as described in section 6.1.
The plots below, (fig. G.2), are obtained from Monte Carlo samples generated with different M_t and selected with the standard DIL selection (see Appendix B for details on DIL selection).

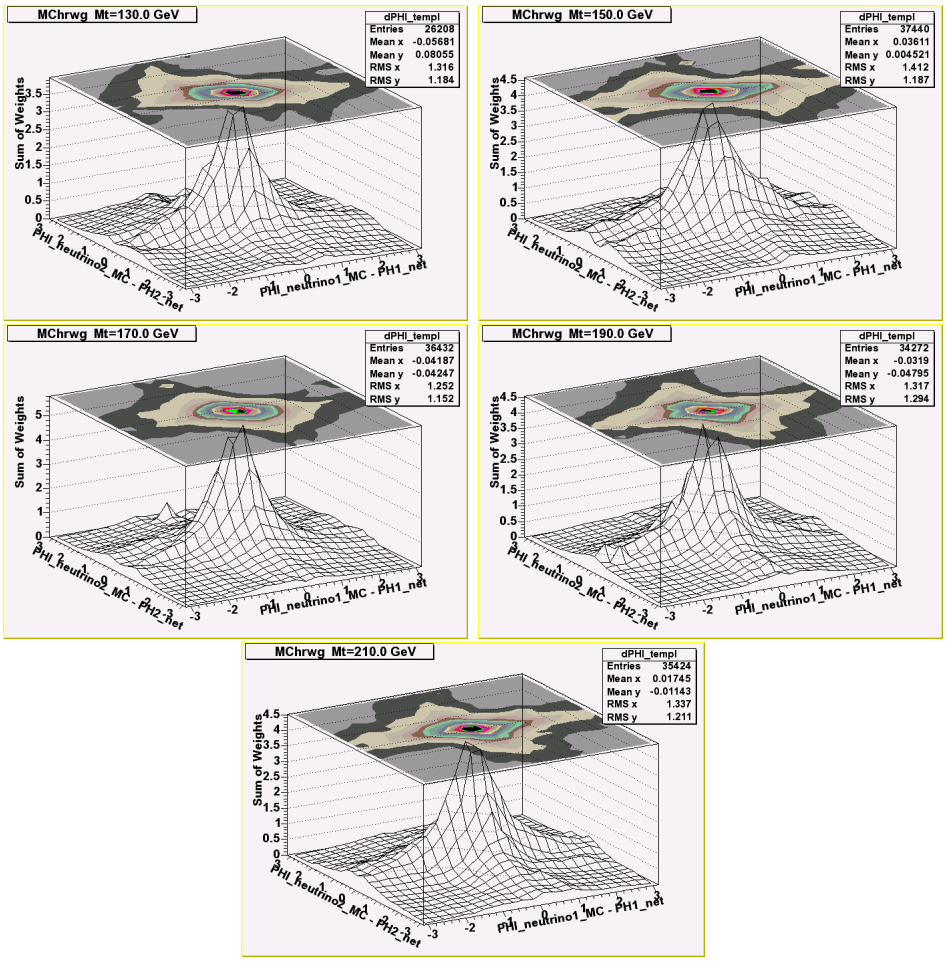

Figure G.2: 2D distributions of $\phi_{\nu 1}^{MC} - \phi_{\nu 1}^{i}$, $\phi_{\nu 2}^{MC} - \phi_{\nu 2}^{j}$ with weights w_{ij}.

The content of each bin in the plots G.2 is defined according to the standard weight $w_{ij} = e^{-\chi_{ij}^2}$, where χ_{ij}^2 is the lowest χ^2 for the $\phi_{\nu 1}^{i}, \phi_{\nu 2}^{j}$ net point.

Plots G.2 show that $\phi_{\nu 1}^{MC} - \phi_{\nu 1}^{i}$, $\phi_{\nu 2}^{MC} - \phi_{\nu 2}^{j}$ are uncorrelated because the curve levels are almost circular around zero [1]: this means that ϕ_1, ϕ_2 have the same correlation as ϕ_1^{MC}, ϕ_2^{MC} and that we can safely choose randomly the azimuthal neutrino angles to constrain the kinematics.

[1] As described in Chapter 6 large value of χ^2 are not taken into account in event reconstruction.

Appendix H

Breit-Wigner functions and M_t-dependent Γ_t in event reconstruction

We are going to discuss two parameters which were refined with respect to the previous analysis [1]. New features are: a m_t dependent rather than constant top width and optimal choices of Breit-Wigner functions for χ^2 and weight (see Chapter 6).

Two definitions of Breit-Wigner functions (BW) were taken into account (6.11, 6.12). Considering both the χ^2 (6.2) and the weight definition (6.13) we have 4 possible BW algorithms for reconstructing an event. Table H.1 summarizes the cases which were considered.

Case	BW in χ^2	BW in weight
Case A	$\frac{\Gamma^2 \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$	$\frac{\Gamma^2 \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$
Case B	$\frac{\Gamma \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$	$\frac{\Gamma \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$
Case C	$\frac{\Gamma^2 \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$	$\frac{\Gamma \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$
Case D	$\frac{\Gamma \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$	$\frac{\Gamma^2 \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$

Table H.1: Breit-Wigner functions considered for adoption in the χ^2 and weight definition

In figure H.1 [2] expected statistical errors versus generated top mass have been plotted for case A, B, C (see H.1) and "0", which is defined according to the χ^2

[1] Previous and current definition of these parameters are given in F and 6.1.2 respectively

[2] This plot has been performed for dilepton DIL samples (see Appendix B for details about DIL selection). The Gen5 framework was used.

and weight of the previous analysis (see appendix F). These errors were obtained in pseudo-experiments [3], for $n_s = 17.2$ [75]. Only signal contribution has been taken into account [4].

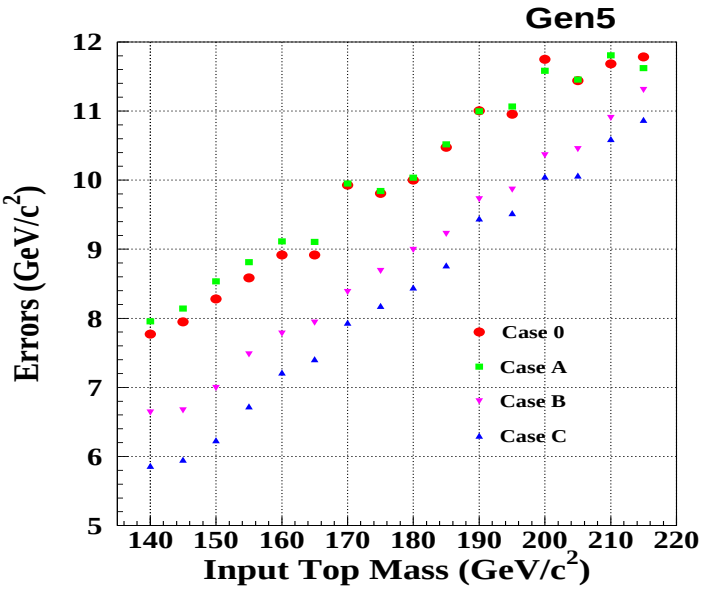

Figure H.1: Top mass dependence of statistical errors by using a mass dependent width, two different definition of Breit-Wigner functions and two different weights for the 144 solutions. (see text for significance of symbols)

The error bars in figure H.1 are too small to be seen, since the boostrap procedure (see Chapter 8) has not been applied [5].

Case D has not been considered since [6]

1. test at hepg level show that reconstructed mass by using $\frac{\Gamma \cdot m^2}{(m^2_{inv} - m^2)^2 + m^2\Gamma^2}$ in the χ^2 has a larger spread than $\frac{\Gamma^2 \cdot m^2}{(m^2_{inv} - m^2)^2 + m^2\Gamma^2}$ (see figure H.2)

2. using 6.11 for the weight definition was found to return a larger statistical error than 6.12 (see figure H.1)

[3] see Chapter 8 for more details on PE's

[4] The background contribution to the statistical error is not expected to depend on the cases reported in table H.1 since background does not depend on top mass.

[5] In figure H.1 the same MC samples have been used at each mass value. Therefore errors on MC limited statistic must not be considered

[6] More details are given in [76].

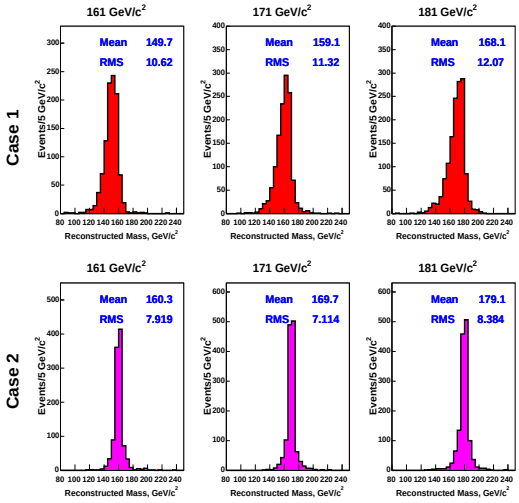

Figure H.2: top mass reconstruction at hepg level by using $\frac{\Gamma \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$ (red) or $\frac{\Gamma^2 \cdot m^2}{(m_{inv}^2 - m^2)^2 + m^2\Gamma^2}$ (purple). Results from three samples with generated top masses of 161, 171, 181 GeV/c^2

Case C (see table H.1) is the case with best mass resolution. The expected improvement with respect to the previous analysis is $\sim$ **20**%

Appendix I

Transfer functions

Figure (I.1), show the transfer functions, fitted in different $(|\eta|, P_T^{jet})$ regions.

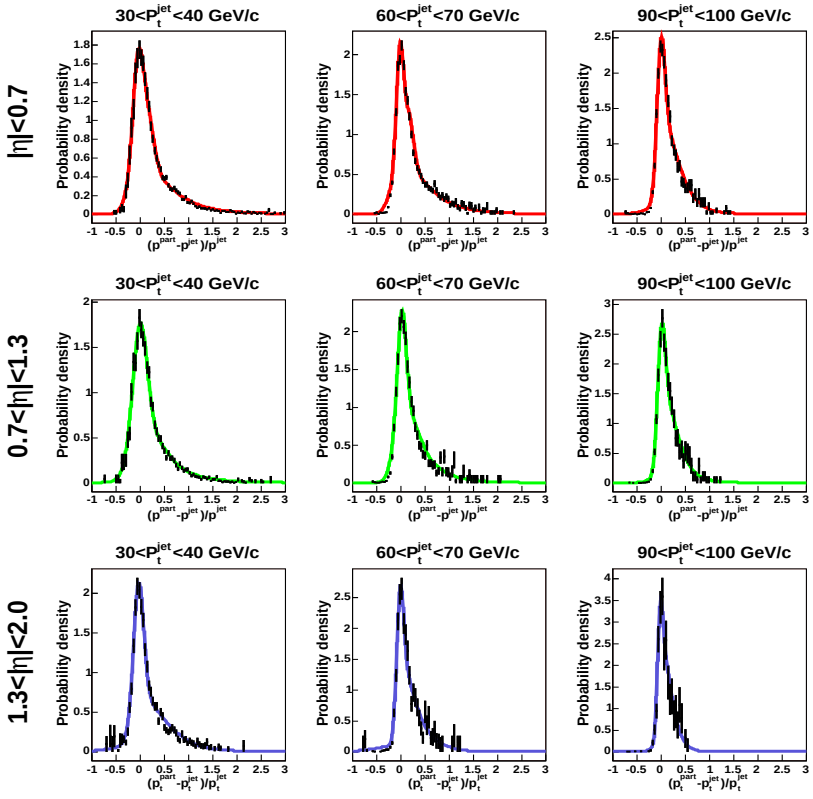

Figure I.1: Transfer functions

In building the transfer functions we used Monte Carlo $t\bar{t}$ events with top masses

175, $175 \pm 0.5\ GeV/c^2$. However, these functions can be used throughout the full analysis because we make them M_t independent by construction, by compensating the b parton P_T dependence on M_t by adding an appropriate weight.

Only a very small improvement with respect to the previous analysis is obtained by adopting these functions, as seen in the figure below:

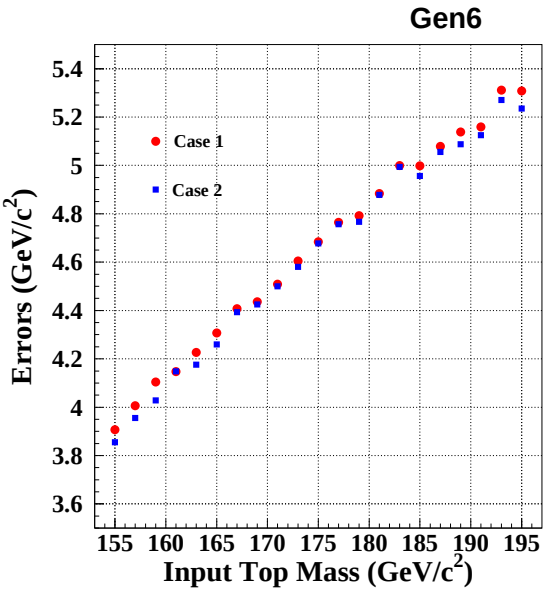

Figure I.2: Statistical errors. Blue dots refer to the χ^2 upgraded with transfer functions, red dots refer to the previous analysis (see appendix F).

Appendix J

Initial and Final State Radiation

Initial (ISR) and final state radiation (FSR) is parametrized by the DGLAP ([64]-[65]) evolution equation that expresses the probability for a parton to branch into two.

For incoming quarks the level of ISR can be measured as a function of the squared invariant mass of the two leptons in Drell-Yan events [1] (see figure J.1). Two simulated samples ($\pm\sigma_{ISR}$) are produced by varying the related QCD parameters in the DGLAP equation in order to cover the region given by uncertainties in measured $< P_T >$ [2] (see figure J.1).

The ISR effect is thus extrapolated to energies of interest.

Since ISR and FSR shower algorithms are the same, the same variation in QCD parameters as in the ISR study is applied in the shower equations for final state partons: $+\sigma_{FSR}$ and $-\sigma_{FSR}$ samples are obtained.

[1] Drell-Yan events have no FSR and are produced via the $q\bar{q}$ annihilation, as most (85%) $t\bar{t}$ pairs are.

[2] $< P_T >$ corresponds to the average amount of ISR in Drell-Yan events

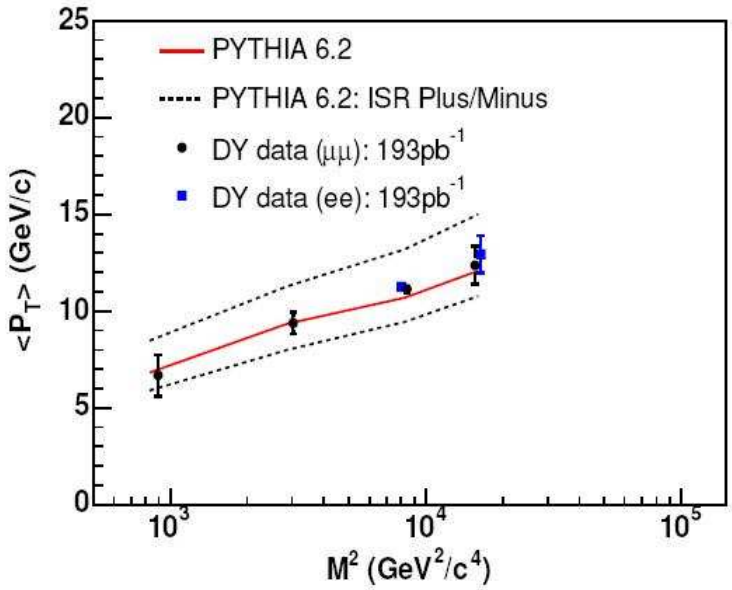

Figure J.1: The average P_T of the dilepton system in Drell-Yan events. A logarithmic dependence on M^2 (squared dilepton invariant mass) can be seen. Data are compared with the predictions of samples simulated with default ISR (PITHYA 6.2) and samples generated with $\pm\sigma_{ISR}$ [69]

Appendix K

Data Validation

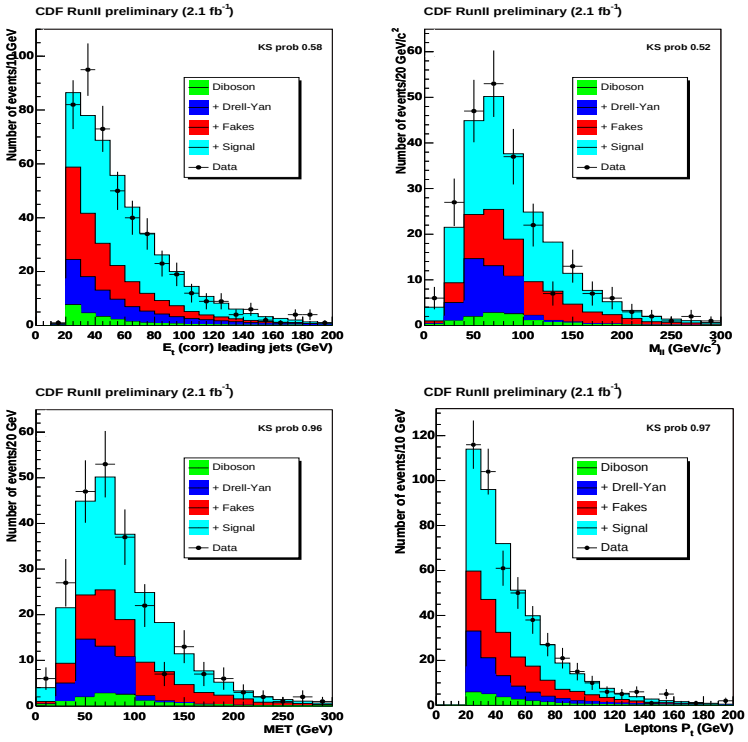

Figure K.1: Leading jet E_T (upper left), tight and track lepton invariant mass (upper right), $\not\!\!E_T$ (lower left), tight and track lepton P_T (lower right) for predicted and observed events (dots) in the LTRK sample. The prediction is scaled so that the predicted number of events is equal to the observed one. Results from the Kolmogorov-Smirnov test are shown. See text for significance of symbols.

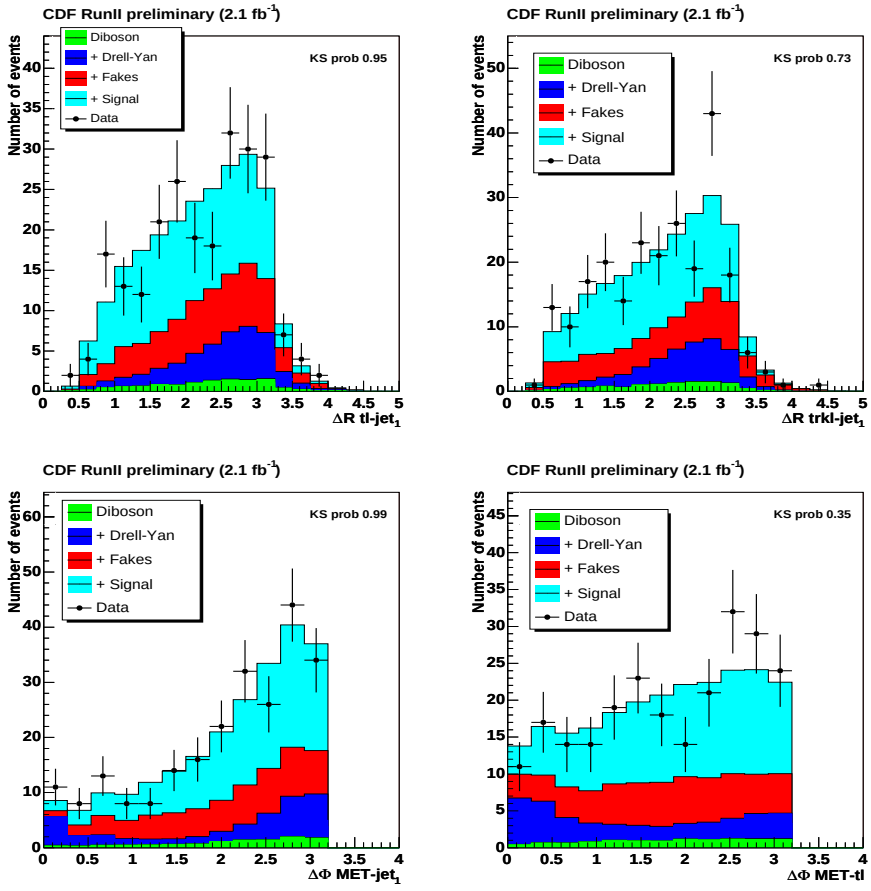

Figure K.2: $\Delta R(tl$, highest-E_T $jet)$ (upper left), $\Delta R(trkl$, highest-E_T $jet)$ (upper right), $\Delta\phi(\not\!\!E_T$, highest-E_T $jet)$ (lower left), $\Delta\phi(\not\!\!E_T,\ tl)$ (lower right) for predicted and observed events (dots) in the LTRK sample. The prediction is scaled so that the predicted number of events is equal to the observed. Results from the Kolmogorov-Smirnov test are shown. See text for significance of symbols.

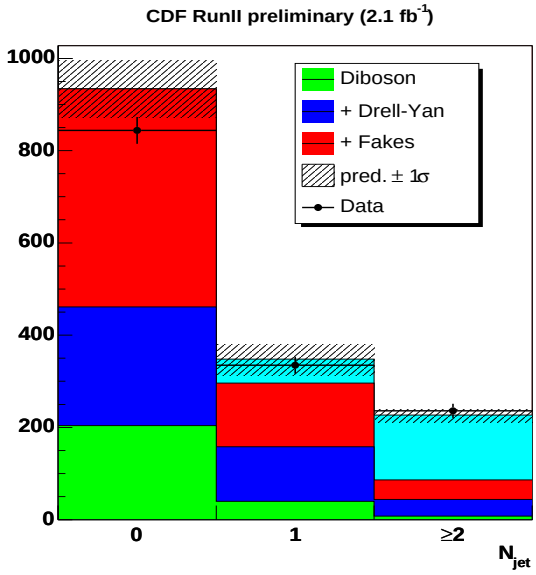

Figure K.3: Number of predicted events compared to observation. The shaded areas show the uncertainties on the predicted numbers.

Bibliography

[1] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories*
Phys. Rev., D **8** 3633 (1973)

[2] H. D. Politzer, *Asymptotic freedom: An approach to strong interactions*
Phys. Rev. Lett., **14** 129 (1974)

[3] S. L. Glashow, *Partial symmetries of weak interactions*
Nucl. Phys. **22** (1961)

[4] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*
Phys. Lett., **13** (1964)

[5] S. Weinberg, *A model of leptons*
Phys. Rev. Lett., **19** 1264 (1967)

[6] I. J. R. Aitchinson & A.J.G. Hey, *Gauge Theories in Particle Physics (second edition)*

[7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*
Phys. Rev. Lett., **13** 508 (1964)

[8] LEP Electroweak Working Group, *status of August 2005*
`http://lepewwg.web.cern.ch/LEPEWWG`

[9] F. Abe et al. *Evidence for Top Quark Production in ppbar Collisions at s=1.8 TeV*
Phys. Rev. Letters 73 225 (1994)

[10] F. Abe et al., The CDF Collaboration, *Kinematical Evidence for Top Pair Production with Multijet Events in ppbar Collisions at s=1.8 TeV*
Phys. Rev. D **51** 4623 (1995).

[11] F. Abe et al. *Observation of Top Quark Production in ppbar Collisions with the Collider Detector at Fermilab*
Phys. Rev. Letters **74**, 2626 (1995)

[12] M. Cacciari et al. *The tt cross-section at 1.8 and 1.96 TeV: A study of the systematics due to parton densities and scale dependence*
JHEP **0404** 68 (2004)

[13] D. Chakraborty, J. Konigsberg, and D.L. Rainwater
Ann. Rev. Nucl. Part. Sci. **53** 301 (2003) [$arXiv : hep - ph/0303092$]

[14] C. S. Hill, The CDF Collaboration, *Operational experience and performance of the CDFII silicon detector*
Nucl. Instrum A **530**, 1 (2004)

[15] T. K. Nelson, The CDF II Collaboration, *The CDF Layer 00 Detector*
Fermilab preprint, FERMILAB-CONF-01-357-E (2001)

[16] A. Affolder et al., The CDF Collaboration, *The intermediate silicon layers detector at CDFII: Design and progress*
Nucl. Instrum. Meth. A **435**, 44 (1999)

[17] A. Affolder et al., The CDF Collaboration, *The CDF intermediate silicon layers detector*
Nuovo Cim. **112A**, 1351 (1999)

[18] T. Affolder et al., The CDF Collaboration, *CDF central outer tracker*
Nucl. Instrum. Meth. A **526**, 249 (2004)

[19] Viktor Veszpremi *Search for the Standard Model Higgs Boson in Proton-Antiproton Collisions at a Center-of-Mass Energy of 1.96 TeV*
PhD Thesis at Purdue University (2007)

[20] A. Artikov et al. *CDF Central Preshower and Crack Detector Upgrade.*
FNAL-PUB-07-023-E (2007)

[21] M. Albrow et al. *The CDF plug upgrade electromagnetic calorimeter: test beam results.*
Nucl. Instr. Meth. A **480**, 524 (2002)

[22] G. Apollinari et al. *Shower maximum detector for the CDF plug upgrade calorimeter*
Nucl. Instr. Meth. A **412**, 515 (1998)

[23] C. M. Ginsburg for the CDF COllaboration, *CDF Run 2 muon system*
Eur. Phys. J. C **33**, S1002 (2004).

[24] T. Dorigo, The CDF Collaboration, *The muon system upgrade for the CDF II experiment*
Nucl. Instrum. Meth. A **461**, 560 (2001).

[25] K. Goulianos, S. Lami, *Performance of a prototype position senssitive towerless calorimeter*
Nucl. Instr. Meth. A **430** 34 (1999)

[26] H. Bachacou et al., *Combining the SecVtx B-Tagging Scale Factors for 5.3.3 Analysis*
CDF note 7480 (2005)

[27] P. Azzi et al., *Histogram tracking in the COT*
CDF note 5562 (2001)

[28] K. Bloom, W. M. Yao, *"Outside-In" Silicon Tracking at CDF*
CDF note 5991 (2002)

[29] Y. Huang, C. Hays, A. Kotwal: *"Inside-Out Tracking"*
CDF note 6707 (2003)

[30] Ulysses, Grundler, Lubomir, Lovas, Anyes, Taffard, *High-Pt muons recommended cuts and efficiencies for Winter 2007*
CDF note 8618 (2007)

[31] G. C. Blazey et al., *Run II jet physics*
$[hep-ex/0005012]$

[32] J. E. Huth et al., *in Proceedings of Research Directions For the Decade: Snowmass 1990*
July 1990, edited by E. L. Berger () p. 134

[33] CDF internal web page
`www-cdf.fnal.gov/internal/dqm/goodrun/good.html`

[34] CDF internal web page
`www-cdf.fnal.gov/tiki/`
tiki-index.php?page=ltrackDileptonAcceptanceChallenge1InvFb

[35] A.Taffard, *Run II Cosmic Ray tagger*
CDF note 6100 (2003)

[36] A.A. Bhatti, K. Hatakeyama, *Relative jet energy corrections using missing E_T projection fraction and djet balancing*
CDF note 6854 (2004)

[37] A.A. Bhatti, K. Hatakeyama, *Update on the relative jet energy corrections*
CDF note 7354 (2005)

[38] B. Cooper, M. D'Onofrio, G. Flanagan, *Multiple interaction corrections*
CDF note 7365 (2005)

[39] A. Bhatti et al., *determination of the jet energy scale at the Collider Detector at Fermilab*
Nucl. Instr. Meth. A **566** 375 (2006)

[40] T.Spreitzer et al., *Electron ID Efficiency and Scale with TopNtuple 6.1.4 for Summer 2006*
CDF note 8274 (2006)

[41] U.Grunder et al., *High-Pt muons recommended cuts and efficiencies for Summer 2006*
CDF note 8262 (2006)

[42] S. Willenbrock
$[arXiv : hep - ph/0211067]$

[43] D. Acosta et al., *The CDFII Detector Technical Design Report, Fermilab Report No. Fermilab-Pub-96/390-E*
Phys. Rev. D **71**, 032001 (2005)

[44] G. Bellettini, J. Budagov, G. Chlachidze, V. Glagolev, F. Prakoshyn, A. Sissakian, I. Suslov, G. Velev, *Top Mass Measurement in Dilepton Events using Neutrino ϕ Weighting Method*
CDF Note 7641 (2005)

[45] A. Abulencia et al., The CDF Collaboration, Phys. Rev. D **73** 112006 (2006)

[46] G. Bellettini, J. Budagov, G. Chlachidze, V. Glagolev, F. Prakoshyn, A. Sissakian, I. Suslov, G. Velev, *Measurement of the Top Quark Mass using the Minuit Fitter in Dilepton Events at CDF*
CDF Note 7239 (2004)

[47] J. Thom, J. Incandela, C. Mills, P. Savard, T. Spreitzer, *A Measurement of the Top Dilepton Cross-Section using the* 1.1 fb^{-1} *Tihgt Lepton and Isolated Track Sample*
CDF Note 8696 (2007)

[48] M.Tecchio, D. Mietl, D. Amidei, T. Maki, V. Giakomopoulou, Y. Takeuchi, N. Kimura, C.-S. Moon, S.Carron, *Top Dilepton Cross Sectionin 1.2 fb^{-1} using the DIL selection*
CDF Note 8741 (2007)

[49] CDF internal web page
`www-cdf.fnal.gov/~tecchio/internal/topntuple/DIL_topntuple.html`

[50] T. Maki, J. Antos, A. Beretvas, Y.C. Chen, R. Lysak, *Cross-section Dependent Top Mass Measurement in Dileptonic Channel using Template Method*
CDF Note 8823 (2007)

[51] T. Mäki, J. Antos, A. Beretvas, Y.C. Chen, R. Lysák, *Top Mass Measurement in Dileptonic Channel using Template Method in* 1.2 fb^{-1}
CDF Note 8727 (2007)

[52] B. Abbott et al., The D0 Collaboration, *Measurement of the top quark mass in the dilepton channel*
Phys. Rev. D **60** 052001 (1999)

[53] T. Affolder et al., The CDF Collaboration, *A Measurement of the top quark mass with the collider detector at Fermilab*
Phys. Rev. D **63** 032003 (2001)

[54] G.V.Velev, *A determination of the top mass from the dilepton events using the modified lepton+jets MINUIT fitter*
CDF note 4607 (1998)

[55] A. Abulencia et al., The CDF Collaboration, *Measurement of the $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\ TeV$ in the all Hadronic Decay Mode*
Phys. Rev. D74, 072005 (2006)

[56] A. Castro, F. Margaroli, *Measurement of the top mass in the all-hadronic channel using the tmt method (1.02 fb^{-1})*
CDF Note 8420 (2006)

[57] CDF internal web page
`www-cdf.fnal.gov/internal/physics/top/RunIIMC/topmc6/`

[58] Particle Data Group (PDG)
`http://pdg.lbl.gov/`

[59] R. Barlow, *SLUO lecture on statistics and numerical methods in HEP, Lecture 6: Resampling and bootstrap*
`http://www.hep.man.ac.uk/u/roger/`(2006)

[60] A. Abulencia et al., The CDF Collaboration, *Top quark mass measurement using the template method in the lepton+jets channel at CDF II*
Phys. Rev. D **73** 032003 (2006)

[61] J. F. Arguin, P. K. Sinervo, S. Xie, *b-jets Energy Scale Uncertainty From Existing Experimental Constraints*
CDF Note 7252 (2004)

[62] J. F. Arguin, P. K. Sinervo, S. Xie, *Combined Measurement of the Top Quark Mass and the Jet Energy Scale in the Lepton+Jets Channel*
CDF Note 7505 (2005)

[63] T. Sjostrand, Comput. Phys. Commun. **82**, 74 (1994); G. Marchesini et al., Comput. Phys. Commun. **67**, 465 (1992).

[64] Y. L. Dokshitzer, Sov. Phys. JETP 46, 641 (1977).

[65] V. N. Gribov and L. N. Lipatov, Yad. Fiz. 15, 1218 (1972).

[66] V. N. Gribov and L. N. Lipatov, Yad. Fiz. 15, 781 (1972).

[67] L. N. Lipatov, Sov. J. Nucl. Phys. 20, 94 (1975).

[68] G. Altarelli and G. Parisi, Nucl. Phys. B126, 298 (1977).

[69] A.Abulencia et al., The CDF Collaboration
Phys. ReV. **D 73** (2006) 032003 [arXiv:hep-ex/0510048]

[70] H.L. Lai et al. Eur. Phys. J. C **12**, 375 (2000)

[71] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, Eur. Phys. J. C **4**, 463 (1998). *Implementation from: PDFLIB: The Parton Density Function Library, Version 8.04 MRST* set 75, CERN

[72] CDF internal web page
`www-cdf.fnal.gov/internal/physics/top/run2mass/sys6.html#pdf`

[73] CDF internal web page
`www-cdf.fnal.gov/internal/physics/joint_physics/instructions/`
PDFUncertainties/pdf.html

[74] CDF internal web page
`www-cdf.fnal.gov/internal/physics/joint_physics/pdfs/pdf_acceptance.html`

[75] *Top Dilepton Cross Section Measurement using the GEN5 DIL Selection*, CDF note 7851, 2005

[76] *Top Quark Mass Measurement in the 2 fb^{-1} Tight Lepton and Isolated Track Sample using Neutrino ϕ Weighting Method using the GEN5 DIL Selection*
CDF Note 7851 (2005)

[77] CDF internal web page
`www-cdf.fnal.gov/physics/new/top/top.html`

[78] F. James, *MINUIT: Function Minimization and Error Analysis*
CERN Program Library, D506.

[79] J.F Arguin, *Measurement of the Top Quark Mass with In Situ Jet Energy Scale Calibration at CDF-II*
CDF Note 7968 (2005)

[80] P. Mehtälä, T. Aaltonen, N. van Remortel *Measurement of the top quark mass in the all hadronic channel using the Ideogram method*
CDF Note 8233 (2006)

[81] B. Jayatilaka, D. Whiteson, R. Shekhar, A. Kotwal, M. Tecchio, *Measurement of the Top Quark Mass in the Dilepton Channel using a Matrix Element Method and Neuroev*
CDF Note 9098 (2007)

Acknowledgements

Writing the thesis represents the end of a 12 month long path of contrasting feelings. The generosity of people who were "no matter what" available to help or comfort me is hidden in each line of this thesis.

Above all I am enormously grateful to *Igor Suslov* and *Corinne Mills* who never hesitated to answer my questions, even if sometimes ingenuous. I hope to meet Corinne one day.

I would like to thank *Vladimir* and *Oleg* from Dubna who helped me with the analysis and cheered me up. I am so grateful to my supervisor *Gueorgui Velev* who leads me with his big experience, my advisor *Giorgio Bellettini* for his kindness and patience in giving me private lessons at the beginning and then correcting my funny mistakes in the thesis and *Guram Chlachidze* who introduced me to the world of high particle physics.

I do not want to forget *the conveners* of the CDF top group and sub-groups who provided me with the framework for the selection code and listened to me in the several presentations I gave.

Many thanks to *Lucas Rogondino*, who transmitted to me a part of his joyful way of life and introduced me to the typical fauna of Fermilab, *Ludovico Cavedon* with his painstaking answers and *Nicola Pozzobon* who shared with me three months of Fermilab life, giving me useful suggestions, and the addiction for Cici's pizza.

Chronologically last but not least *Sandra Leone*, *Giuseppe Latino*, *Michele Giunta* and *Giorgio Chiarelli* who supported me in the last troubled period in Pisa.

In nine months of staying in US I met people also outside the Fermilab laboratory area. I have no words to describe how family *Ewald* welcomed me. My "goofriend" *Kristin* patiently waited for me when I was busy, *Bonnie* talked to me

even if I was shy and sometimes unable to speak English fluently and *Kourtney* made me laugh with her outspokenness. Last, but not least, I want to thank *Kiley*, an amazing 13 year old girl always busy in watching movies and never filled up ("I'm hongry"): she enlightened my heart despite of what she can understand or think.

How I can forget of my family always supporting me in my decisions! I am the lucky beneficiary of their loving encouragements. My brother *Fabio* who passed with merits the Ph.D test in the "Scuola Normale Superiore" in Pisa after a year of sufferings, my mother Pia ("*Piña*") who was all the time by my side and my father with his precious suggestions. I would like only to remember my cousin *Sara*, living in Argentina for 6 months and my grandmother Sofia with her funny way of speaking.

My friends *Serena*, *Floriana*, *Chiara*, *Enrico*, *Elena* with whom I exchanged so many e-mails. A special thank to *Nino* ("*Lino*" as the barber of my town), my best friend, who simply accepted me as I am and my ex-roommate *Marco* with whom I had so much fan.

And I want to remember Maria, Vito, Donovan (the "dog"), Francesco, Federico, Christian, Natalino, Luca, Gian Piero, Anna, Ilja, Gabriele (I owe him a beer as well as Nicola), Sergey and "sgarbish", a south-American friend of mine.

Thanks to you, guys, I've made it!!!