Institute of Experimental Physics Slovak Academy of Sciences

Top quark mass measurement in dilepton channel

PhD Thesis

Roman Lysák

Supervisor:
RNDr. Jaroslav Antoš, CSc.

Košice 2007

Abstract

We present a measurement of the top quark mass from events produced in $p\bar{p}$ collisions at a center-of-mass energy of 1.96 TeV, using the Collider Detector at Fermilab. We identify $t\bar{t}$ candidates where both W bosons from the top quarks decay into leptons $(e\nu, \mu\nu, \tau\nu)$ from a data sample of 340 pb^{-1}. The top quark mass is reconstructed in each event separately by the method which draw upon simulated distribution of $t\bar{t}$ longitudinal momentum in order to extract probability distribution for the top quark mass. Representative distributions, or templates, are constructed from simulated samples of signal and background events, and parametrized to form continuous probability density functions. A likelihood fit incorporating these parametrized templates is then performed on the data sample masses in order to derive a final top quark mass. Measured top quark mass is $M_{top} = 169.5^{+7.7}_{-7.2}(stat.) \pm 4.0(syst.)$ GeV/c^2.

To my wife Lucia,
and to my daughter Natalia.

Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my gratitude to all people who made possible this thesis exists.

First, I would like to thank my advisor Dr. Jaroslav Antoš. Without his continuous help, advices and support I would never finish this work.

Another thanks goes to all my colleagues who I worked with at Fermilab. These were people who I worked together on CDF data processing, mainly Miro Siket, Yen-Chu Chen and Steve Wolbers. Other thanks goes to colleagues who I spend the time with, by doing the physics analysis. These were mainly Tuula Mäki, Andy Beretvas and once again Yen-Chu Chen. Special thanks goes to Gong Ping Yeh for his valuable insights into physics and his broader view on what is really important.

I would also like to thank all the Czech and Slovak guys who I met during my stay at Fermilab.

Last, but not least, I would like to thank my wife Lucijka for her support while working on this thesis.

Chapter 1

Introduction

When the top quark was discovered in the 1995 at Tevatron by CDF and D0 collaboration [1, 2], almost 20 years long journey of searching finally ended. Already since the bottom quark discovery [3] in 1977 (also at the Tevatron), the existence of top quark was theoretically expected according to the Standard Model of fundamental particles and forces, so that the bottom quark would have its partner as all the other quarks and leptons have. In fact, after the b-quark was discovered, no one expected it will take until 1995 for top quark to appear, since no one supposed it would have such a huge mass comparing to other quarks (top quark have ∼ 35 times larger mass than the next heaviest quark, the bottom quark).

Standard Model (SM) is the theory of electromagnetic, weak, and strong interactions, i.e. all known interactions which occur in nature except gravity. The basic constituents of SM are 6 quarks and 6 leptons:

$$quarks: \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \qquad leptons: \begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix},$$

which interact between themselves through the interchange of intermediate bosons $\gamma, W^\pm, Z^0$ and gluons. Moreover, there exist the yet undiscovered Higgs boson within SM which cause particles to have a mass.

The top quark would appear to be just another “boring” SM quark with most of the properties well described by the SM. However, the particle masses are not predicted within the SM and need to be measured out. Thus, the top quark mass is the most studied[1] and the interesting property of the top quark and probably also the all high energy physics at the present. There are a few reasons why the top quark and particularly its mass (due to its huge value) is so interesting to study:

---

[1] only at CDF experiment there are ∼ 10 groups which measure top quark mass

- Top quark mass value (~ 170 GeV/c^2) is close to the scale of electro-weak symmetry breaking (EWSB) $v = (\sqrt{2}G_F)^{-1/2} \sim 246$ GeV. Hence, the Yukawa coupling of the top quark to the Higgs boson is intriguingly close to one, $\lambda_t = \sqrt{2}m_{top}/v = m_{top}/173.9$ GeV ~ 1. This raise immediately a questions. Is the top quark mass really generated by the Higgs mechanism as the SM predicts or does the top quark have more fundamental role in the EWSB? Could the non-SM physics first manifest itself in non-standard production and decays of heavy top quark?

- Top quark has such a huge mass that its lifetime is smaller than the QCD time scale, which implies that the top quark as the only quark decay even before it can hadronize, i.e. there are no bounded states of the top quark (e.g. no mesons composed of $t\bar{t}$). Therefore, the information about top quark properties (e.g. spin) is not lost and is directly transfered to its decay products, so we can directly measure them.

- Due to such a huge mass, the QCD corrections for the top quark are important. This is especially true for production of yet undiscovered Higgs boson, the last piece of SM which is still missing. The more precise we will know the top quark mass the more precise we will know the range where the Higgs mass can be and where we should look for it.

Hence, the precise measurement of the top quark mass is very important.

The goal of this thesis is to

- implement, improve and optimize the method of top quark mass measurement [4] (proposed during CDF Run 1 series of data taking) on conditions in new series of data taking CDF Run 2

- perform the measurement of the top quark mass in dilepton channel which will be significantly better than what was achieved before.

A brief overview of the method for the mass reconstruction is as follows. We will use dilepton decay channel for our measurement where both W bosons from $t\bar{t}$ events decay into leptons (for details, see Sec. 2.2.2). Because of 2 neutrinos in the final state, there are not enough measured variables to kinematically constrain the system. Therefore, we make assumption about one more variable (we will see in Sec. 6.1 that it is the longitudinal momentum of $t\bar{t}$ system) to solve the system, i.e. to reconstruct the top mass for particular event (so called, “raw top quark mass”). The raw top quark mass will be our sensitive variable to the original, true top quark mass. The final determination of the top quark mass will be then performed by comparing the distribution of this sensitive variable between data events and the distributions of the same variable derived from simulated events with various input top mass values. The top quark mass whose simulated events, when combined with background, best describe the sensitive variable distribution seen in the data, is our measured top mass.

The thesis is structured as follows. After the introduction in this chapter, we will acquaint the reader with the theoretical basics of the top quark physics in chapter 2. We will then describe the whole apparatus used to collect the data (chapter 3). In the chapter 4, we will present the event selection used to keep the interesting events with top quark produced while keeping the background as low as possible. All the following chapters, will be related to top quark mass measurement. First, we will describe the existing methods for measurement of top quark mass in chapter 5. We will mention there also advantages and disadvantages of different kind of methods. The above mentioned top mass reconstruction method will be described in detail in chapter 6. In chapter 7 are presented the results we obtained on simulated events. The detailed tests of the method on simulated events are presented in chapter 8. There, we will show the method gives appropriate estimates of the top quark mass and also its statistical uncertainty. Chapter 9 deals with systematic errors associated with the top quark mass measurement. Finally, the results which we obtained using CDF data sample are presented in chapter 10 with the conclusion of the work being in chapter 11.

1.1 Unit Convention

For convenience, we will use the Natural Units through the text, that is, $\hbar = c = 1$. Hence, energy, momentum and mass will have the same units - mostly GeV.

Chapter 2

Top Quark Physics

The theoretical description of particles and forces is govern in Standard Model (SM) [5, 6, 7]. Long before its discovery in 1995, the top quark was regarded as an essential ingredient of the Standard Model. Its existence and many of its properties like:

- electric charge $Q^t_{em} = \frac{2}{3}|e|$
- weak isospin partner of b quark: $T^t_3 = \frac{1}{2}$
- color triplet
- spin $\frac{1}{2}$

are determined by requiring the consistency of the Standard Model. We will mention here two particular reasons why top quark should exists.

In order to confirm that the observed top quark is the one predicted by the Standard Model, all of its properties must be experimentally verified by direct observation. We will describe, and wherever it is possible we will also compare between theoretical predictions and current experimental results of some basic top quark properties in Sec. 2.2. There exists many reviews of top quark physics [8, 9, 10, 11, 12], where also the rest of the top quark properties such as charge, spin, $t\bar{t}$ spin correlations, rare decays, etc. are described.

2.1 Why we need the top quark?

There are a few different kind of reasons why the top quark was believed to exist even before its experimental discovery. In next sections we will discuss one purely theoretical reason (consistency of the Standard Model gauge theory) and one coming from requirement of agreement between theory and the experiment (consistency of precision electroweak measurements with the Standard Model).

2.1.1 Anomaly cancellation

One argument which support the existence of a complete third quark generation comes from perturbation theory. In particle physics the terms of a perturbation series are depicted in Feynman diagrams. The first order terms are pictured as tree level diagrams. Higher order terms correspond to loop diagrams. Certain loop diagrams (see Fig. 2.1) are divergent.This divergence is independent of the fermion mass and depends only on the fermion couplings to the gauge bosons. Such divergences can not exist in a physical theory, and must somehow be canceled. These divergences can be overcome by summing up over several divergent terms in a consistent manner and have divergences cancel each other, procedure known as renormalization. It can be shown, that these divergences cancel for each generation of fermions [9]. The cancellation of the gauge anomalies in the Standard Model for the third generation therefore requires that the b quark have a $T_3 = 1/2$ partner with electric charge $Q^t_{em} = 2/3|e|$ and hypercharge $Y^t = Q^t_{em} - T^t_3$. The partner of the b quark is by definition the top quark. Since anomaly cancellation is independent of mass, a priori, the top quark mass could be anything.

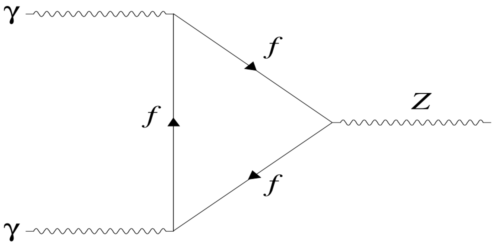


Figure 2.1: An example of a fermion triangle diagram where f could be any but the same of the SM fermions.

2.1.2 Consistency of electroweak measurements with SM

Even through the quark masses are not predicted by the SM, they enter as a parameters in the calculation of radiative corrections to electroweak processes. With high precise measurements at hand, it is therefore possible to indirectly determine the top quark mass from such processes.

Two examples of radiative corrections to the gauge boson self-energies involving the top quark are given in Fig. 2.2. The correction term introduce a quadratic dependence on the top quark mass and therefore the top quark mass plays large role in the

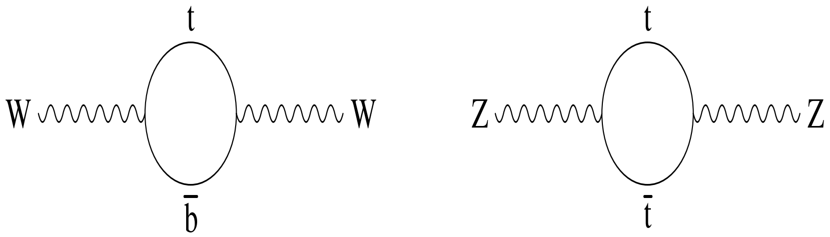


Figure 2.2: Radiative corrections of the top quark to the gauge boson self-energies.

corrections. The effects from other quarks are much smaller due to their much smaller masses. The most precise electroweak measurements are available from e^+e^- colliders operating at the $\sqrt{s} = M_Z$. In the 1990s there were two such particle colliders: the Large Electron Positron collider(LEP) at CERN with four experiments (Aleph, Delphi, L3 and Opal) and the Standford Linear Collider (SLC) with one experiment (SLD).

The LEP and SLD electroweak working groups have performed the constrained fit to the SM using many precisely measured quantities like mass M_Z and width Γ_Z of Z boson, forward-backward asymmetry $A_{FB}^{0,l}$, $(l \in e, \mu, \tau)$ for Z^0 decays into charged leptons and many others, see [13]. Among the results of the fits is an indirect determination of the top mass. If the top mass is left floating, the fit yields a value of $M_{top}^{theory} = 179^{+12}_{-9}$ GeV [13] which is in very good agreement with the latest Tevatron (CDF & D0) combined measured value $M_{top}^{exp} = 172.5 \pm 2.3$ GeV [14].

Such electroweak measurements then not only require top quark, but also have some predictive power over its properties.

2.2 Top quark properties

In this section, we will describe how the top quark is produced and how it decays within Standard Model. Moreover, we will discuss about top quark mass. We will mention possible definitions of mass, what are the current measurements of top mass and why we need to measure the top quark mass as precisely as possible.

2.2.1 Hadronic production of top quark

The top quark can be produced by strong interaction (however, it must be the top-antitop pair production, because the quark flavor is conserved in strong interactions) and also by electro-weak interaction (in this case, single top quark can be produced).

Here, we are interested in production of top quark at hadron colliders, as our

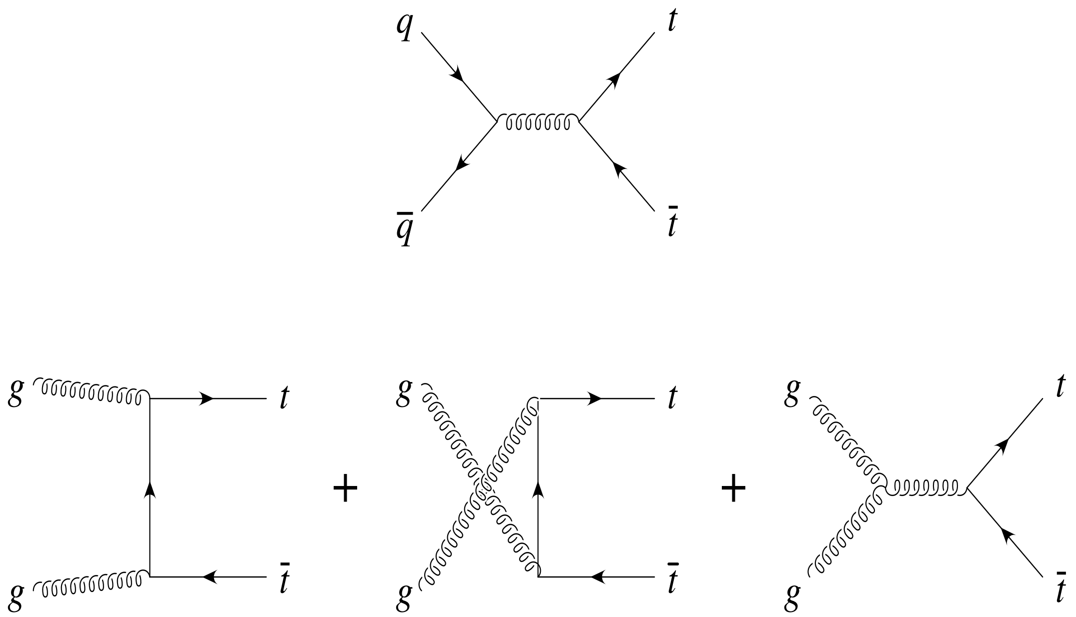


Figure 2.3: Leading order Feynman diagrams contributing to top quark pair production at hadron colliders.

measurement will be performed using $p - \bar{p}$ collider [1]. At hadron colliders, there are two leading-order (LO, i.e. order of $O(\alpha_s^2)$ contributions) subprocesses by which $t\bar{t}$ pairs are produced , the $q\bar{q}$ annihilation ($qq \rightarrow t\bar{t}$) and gluon fusion ($gg \rightarrow t\bar{t}$) as shown in Fig. 2.3. However, about 30% of the contribution to the total $t\bar{t}$ cross-section comes from next-to-leading order (NLO) contributions ($O(\alpha_s^3)$) [10]. Some higher order diagrams for the $q\bar{q}$ annihilation process are shown in Fig. 2.4, the similar is true for gg fusion.

However, we know that quarks can not stay as free particles, rather are confined within hadrons due to asymptotic freedom feature of QCD. Therefore, we can not make beams and accelerate bare quarks and gluons, but only the hadrons which are collections of quarks and gluons. It is possible to treat the quarks and gluons as an essentially free inside hadrons in the limit of large momentum transfers. In such case, each parton carries some fraction of the total momentum of the hadron, $p_{parton} = xp_{hadron}$, where x is the fraction of momentum carried by the parton. We can then look at the interaction between the hadrons as the interaction between free partons, see Fig. 2.5. It can be easily shown, that in order $t\bar{t}$ pair to be produced it is needed

---

[1] Of course, top quark can be, in principal, produced at other types of colliders ($e^+ - e^-$,$e^- - p$,etc.).

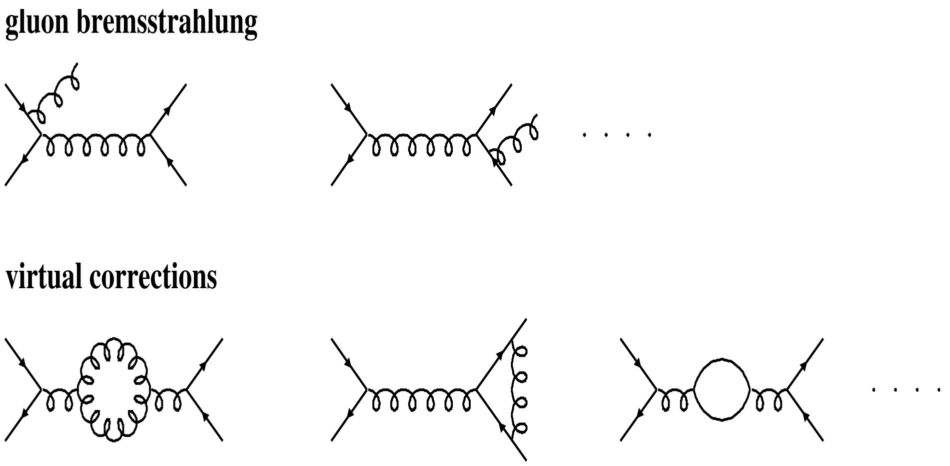


Figure 2.4: Some higher order Feynman diagrams of the $t\bar{t}$ pair production.

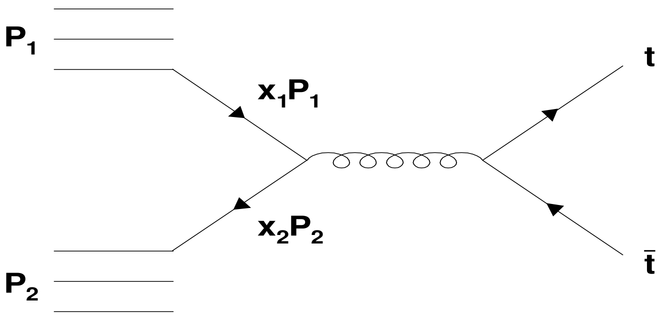


Figure 2.5: Visualization of parton-parton interaction with $t\bar{t}$ production in proton-antiproton collision.

to be fulfilled the following condition

$$x_1 x_2 \geq \frac{4m_t^2}{s}, \tag{2.1}$$

where x_1, x_2 are the fractions of momentum carried by partons coming from initial hadron p_1, p_2, respectively. m_t is the top quark mass and s is the center-of-mass energy of the interaction of initial hadrons. Let's assume that both partons have the same fractions of energy of initial hadrons ($x_1 = x_2 = x$). For Tevatron ($\sqrt{s} = 1.96$ TeV) it is then needed that the fraction of momentum carried by parton $x \geq 0.18$, while at LHC ($\sqrt{s} = 14$ TeV) $x \geq 0.025$.

The total cross-section for interaction $p\bar{p} \to t\bar{t}$ must be then calculated as the sum of the cross-sections for all possible energies of all possible initial partons for a given

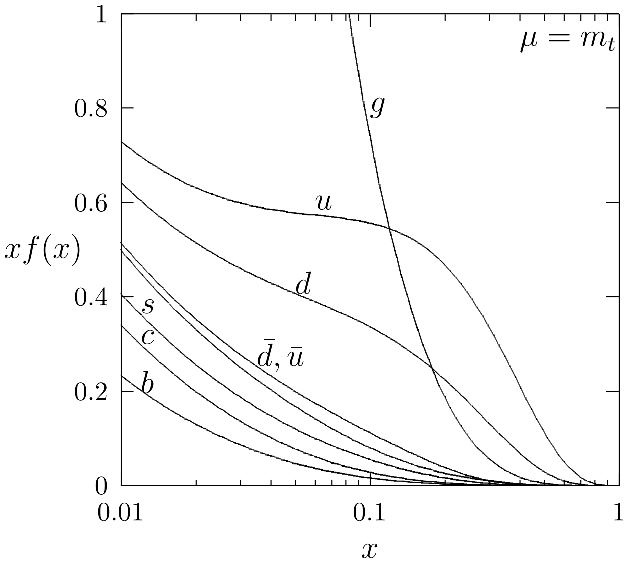


Figure 2.6: Parton distributions functions relevant for $t\bar{t}$ production [12].

energy of initial hadrons

$$\sigma_H(p\bar{p} \to t\bar{t}; s) = \sum_{i,j} \int f_i(x_1) f_j(x_2) \hat{\sigma}_{ij}(q_i q_j \to t\bar{t}; x_1 x_2 s) \mathrm{d}x_1 \mathrm{d}x_2, \tag{2.2}$$

where the hadronic center-of-mass energy is s, the partonic center-of-mass energy is $\hat{s} = x_1 x_2 s$ and the parton level cross section is $\hat{\sigma}_{ij}(s)$. The functions $f_i(x)$ which give us the probability of obtaining the parton i with a fraction energy of x of a parent proton are called Parton Distribution Functions (PDF), see Fig. 2.6. These functions have not yet been calculated from the first principles of QCD and have to be extracted from the available experimental data. As we can see, for Tevatron energy ($x \geq 0.18$) the quarks are dominant within proton, while for LHC ($x \geq 0.025$) the gluons clearly dominate. Therefore, as the partonic cross-sections $qq \to t\bar{t}$ and $gg \to t\bar{t}$ are both of the same order $O(\alpha_s^2)$, at Tevatron qq annihilation dominates over gg fusion (85% vs. 15%), while at LHC the dominant production will be through gg fusion (90% gg, 10% qq) [10].

At Tevatron energy ($\sqrt{s} = 1.96$ TeV), the theoretical cross-section is estimated by two different groups of theorists. The estimate from [15] is $\sigma = 7.2^{+0.8}_{-1.0}$ pb while $\sigma = 7.3^{+0.6}_{-0.6}$ pb according [16], where the theoretical predictions in these papers are interpolated to a top quark mass value $m_{top} = 172.5$ GeV which was at the time of writing this thesis the preliminary world top mass average [14].

The theoretical predictions together with the preliminary CDF measurements of $t\bar{t}$

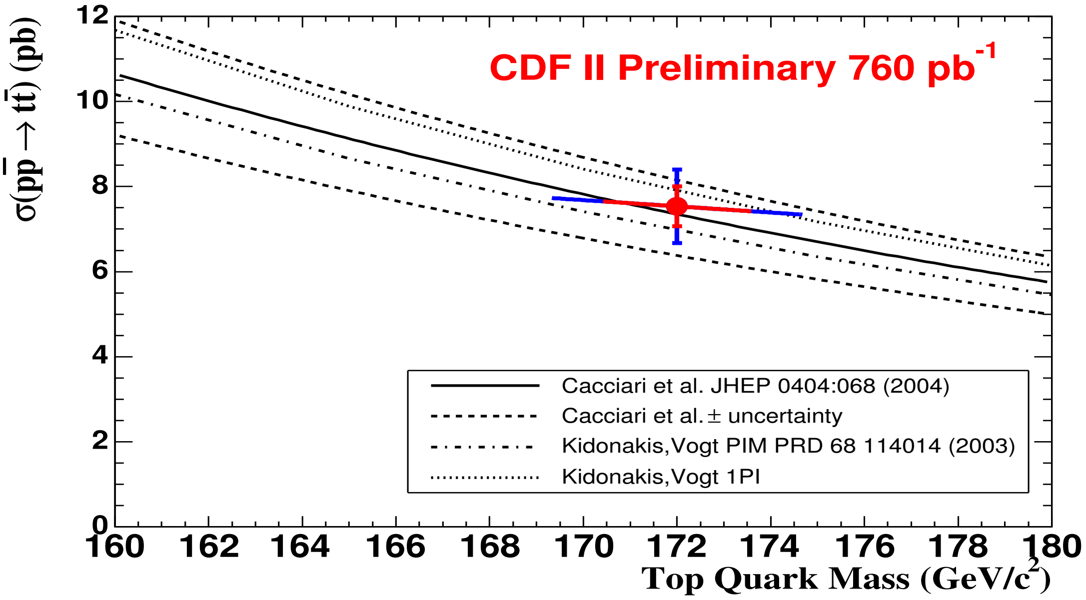


Figure 2.7: Comparison of theoretical and CDF measured $t\bar{t}$ cross-section as a function of top quark mass [17]. CDF measurement (point) is shown together with statistical uncertainties (red bars) and total uncertainties including systematic uncertainties (blue bars). Two different theoretical predictions (shown as curves) together with uncertainties are shown.

production cross section and top quark mass ($\sigma_{t\bar{t}} = 7.5 \pm 0.9$ pb for $m_{top} = 172.0$ GeV) shows good agreement in Fig. 2.7 where there is plotted the dependence of cross-section on top quark mass. The dependence of partonic $t\bar{t}$ cross-section and hadronic $t\bar{t}$ cross-section is shown in Fig. 2.8.

As was pointed out at the beginning of the section, single top quark production is also possible. The weak production of single top quark at Tevatron can occur via the modes depicted in Fig. 2.9:

- the t-channel: a virtual W strikes a b quark. The W boson is space-like ($q^2 = t < 0$, where q is four-momentum of the W)[2]. This mode is also known as *W-gluon fusion*, since the b quark originates from gluon splitting into a $b\bar{b}$ pair.

- the s-channel: this production mode is of Drell-Yan type. A time-like W boson ($q^2 = s \geq 0$) is produced by the fusion of two quarks.

---

[2] the t and later defined s are the Mandelstam variables

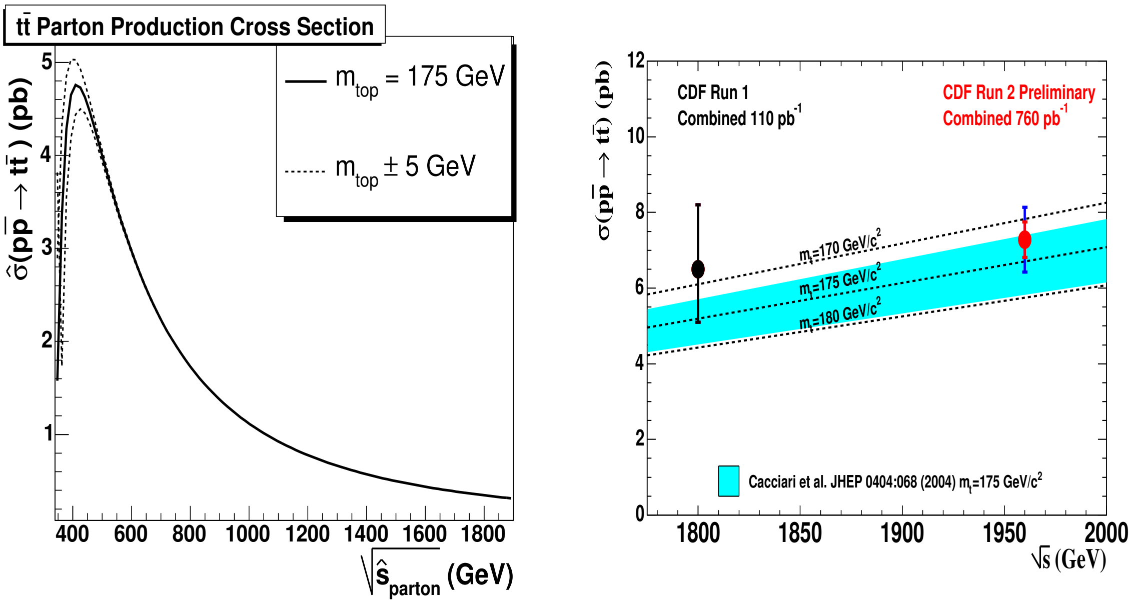


Figure 2.8: Parton production $t\bar{t}$ cross-section as a function of partonic center-of-mass energy $\sqrt{\hat{s}}$ (left) [18]. This is the parton-level cross-section before convolution with the parton distribution function. After the convolution, we get the total cross-section for proton - antiproton interaction (right) [17].

However, there is also possible associated production of W boson and top quark ($b+q \rightarrow b \rightarrow t+W$), but this process is negligible at the Tevatron and is of considerable size only at LHC energies ($\sqrt{s} = 14$ TeV).

Production of single top quark is interesting for a few reasons. First, it can be used to directly measure element of Cabibbo-Kobayashi-Maskawa matrix $|V_{tb}|^2$ which is closely tied to number of quark generations. Second, the spin polarization of single top quark can be used to test V-A structure of weak charged current interactions. Of course, it is background to many SM or non-SM processes (like Higgs boson production) and moreover the deviations from predictions for single top production could lead us to physics beyond SM.

The theoretical expectations for cross-section of single top production are 1.98 ± 0.25 pb for t-channel and 0.88 ± 0.11 pb for s-channel at $\sqrt{s} = 1.96$ TeV [19]. Even the cross-section for single top quark production is only roughly 2.5 times larger than cross-section for $t\bar{t}$ production, it is much harder to observe the single top quark events due to presence of much higher background. Recently, D0 collaboration presented first evidence for production of single top quark [20], where they measured the combined cross-section for s-channel and t-channel to be $\sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 4.9 \pm 1.4$ pb which is almost 1.5 standard deviation higher than expected according Standard

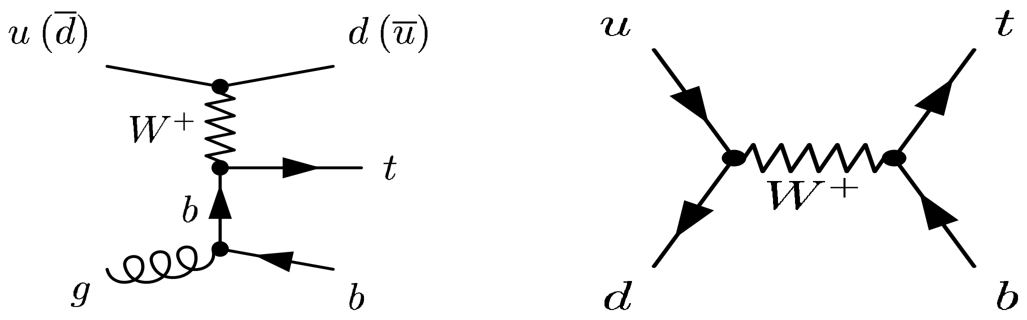


Figure 2.9: Feynman diagrams for t-channel (left) and s-channel (right) weak production of single top quark.

Model.

2.2.2 Top quark decay and $t\bar{t}$ decay modes

Quarks are transformed into their isospin partners via the charged-current weak interaction, but there are also occasional transitions between different generations of quarks. This is described by Cabibbo-Kobayashi-Maskawa (CKM) matrix. The matrix elements V_{tb}, V_{ts} and V_{td} characterize the strength of the transition of the top quark into a bottom, strange and down quark, respectively. These matrix elements are related to each other by $\sum_{j=b,s,d} |V_{tj}|^2 = 1$ assuming CKM matrix to be unitary. Since we know from other measurements that $|V_{ts}|, |V_{td}| \ll 1$ [21] and within SM there are only 3 generations of quarks, it's evident that $|V_{tb}| \approx 1$. So, the dominant decay of the top quark is $t \longrightarrow Wb$ and it can be shown that [22]

$$\Gamma(t \rightarrow bW)/|V_{tb}|^2 \approx \Gamma(t \rightarrow bW) \approx 1.42 \text{ GeV} \Rightarrow \tau = 1/\Gamma \approx 4 \times 10^{-25} s \tag{2.3}$$

The large width is mainly due to strong dependence on large top mass value ($\Gamma_t \sim m_t^3$). Consequently, the top quark lifetime is very short compared to the QCD time scale

$$\tau_{QCD} = 1/\Lambda_{QCD} \approx 1 fm/c = 3 \times 10^{-24} s \tag{2.4}$$

Therefore the top quark decays, as the only quark, before it can form a bound state – hadron.

At the Tevatron, the $t\bar{t}$ pair production is dominant to single top production (see the above Sec. 2.2.1) and moreover it has more distinct experimental signature. Because top quark decays essentially always to b and W, the $t\bar{t}$ decay modes are characterized by the W decays, which can be leptonic or hadronic [21].

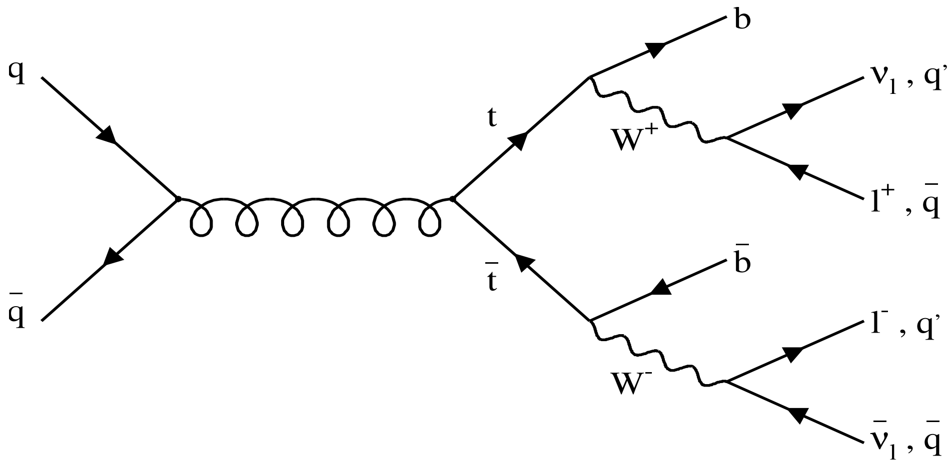


Figure 2.10: $t\bar{t}$ decay chain.

Usually, the final state channels are classified according number of charged leptons [3], see Fig. 2.10:

- *dilepton*: in this case, both W's decay leptonically $W \rightarrow \ell\nu$. This is the channel with the smallest branching ratio (see Tab. 2.1). However, it has the cleanest signal (highest signal to background(S:B) ratio) because of two high energetic, easy detectable and identified leptons and due to two neutrinos in final state which appear as high missing E_t (imbalance of momentum in plane perpendicular to beam direction). Moreover, in the final state there are at least 2 jets coming from hadronization of b quarks. More jets can be produced due to gluon initial and final state radiation (see Fig 2.4).

- $lepton + jet$: in this channel, one of the W decays leptonically while the other one decays to hadrons. This channel is has a reasonable high branching ratio and also the background is quite well understood and reasonable high. It has in the final state at least 4 jets where two of them are b-jets (they can be distinguish from other jets using b-jet identification ("b-tagging") techniques).

- $all - hadronic$: this channel is the one with the highest branching ratio, however the disadvantage of this channel is the huge QCD background, because it has only 6 jets in final state and no clean lepton or high missing E_T.

The branching ratios for all processes belonging to all three channels are summarized in the Tab. 2.1.

---

[3] It is possible to classify events according other variables. For example, at CDF exists top quark mass measurement which selects events with high missing transverse energy, see Sec. 5.1.3

category	decay mode	branching ratio		signal:background
dilepton	$t\bar{t} \rightarrow e\nu be\nu b$	1/81	4/81 (5%)	$\sim 2:1$
	$t\bar{t} \rightarrow \mu\nu b\mu\nu b$	1/81		
	$t\bar{t} \rightarrow e\nu b\mu\nu b$	2/81		
	$t\bar{t} \rightarrow e\nu b\tau\nu b$	2/81		
	$t\bar{t} \rightarrow \mu\nu b\tau\nu b$	2/81		
	$t\bar{t} \rightarrow \tau\nu b\tau\nu b$	1/81		
lepton + jets	$t\bar{t} \rightarrow q\bar{q}be\nu b$	12/81	24/81 (30%)	$\sim 1:1$
	$t\bar{t} \rightarrow q\bar{q}b\mu\nu b$	12/81		
	$t\bar{t} \rightarrow q\bar{q}b\tau\nu b$	12/81		
all-hadronic	$t\bar{t} \rightarrow q\bar{q}bq\bar{q}b$	36/81	36/81 (44%)	$\sim 1:10$

Table 2.1: The $t\bar{t}$ production decay modes with branching ratios and approximate signal to background ratios. The S:B ratios are approximately valid for a selections where no b-jet identification is performed.

Note, as the leptonic decay we consider only decays to e and μ leptons. The τ leptons are hard to identify, because they decay themselves into lighter leptons or even into hadrons [21].

For our top quark mass measurement, we use the *dilepton* channel as we will see later in Sec. 6.1.

2.2.3 Top Quark Mass

One property of top quark which can not be predicted from Standard Model [4], and in fact it was measured and found that it has unexpected high value (comparing to other quarks) is the mass M_{top}.

Unlike leptons, quarks can not live as free particles, so we can not directly measure their masses. Therefore, although one often speaks about quark masses as one would of the mass of electron, any quantitative statement about the value of a quark mass must make reference to theoretical framework which have been used to define it. There are a few different definitions of quark mass:

- *constituent quark masses* - used in non-relativistic quark models of hadrons. These masses are not related to quark mass parameters of the QCD and have only sense in the limited context of particular quark model.
- in the tree-level QCD Lagrangian one finds mass parameters of quarks which are

---

[4] it is the fundamental parameter of SM. That means, it must be measured and then the value is set into the theory *ad hoc*.

however not physical masses

- mass parameters after renormalization, which depend on renormalization scale - *running masses*. Renormalization is procedure that invokes a subtraction scheme to render the physical quantities finite, and requires the introduction of a dimensional scale parameter μ. There are many renormalization schemes possible, where minimum-subtracted scheme ($\overline{MS}$) being the most commonly used. The masses defined using proper scheme can have better desired properties (faster perturbative convergence) than pole mass.

- *bare masses* - the values of quark masses obtained directly in lattice QCD with a lattice spacing a as the ultraviolet cut-off. The bare quark masses can be related to running masses in perturbation theory.

- *current masses* - the quark masses of light quarks obtained in the chiral perturbative theory

- masses (especially of b quark, but also c quark) extracted using Heavy quark effective theory (HQET) . The masses of heavy quarks could not be extracted from lattice QCD simulations up to now due to computing resources limit a^{-1} being in current simulations 1.5 - 2.5 GeV. Therefore effective theory must be used in such a case.

- the *pole quark mass* defined as the real part of a pole in the particle's propagator. For an observable particle such as the electron, the position of the pole in the propagator is the definition of the particle mass. The pole mass and running mass are related, e.g. the top pole mass, m_{pole}, and top $\overline{MS}$ mass evaluated at the $\overline{MS}$ mass scale, $\overline{m}(\overline{m})$, are related by

  $$m_{pole} = \overline{m}(\overline{m}) \left[1 + \frac{4}{3}\frac{\overline{\alpha_s}(\overline{m})}{\pi} + 8.28 \left(\frac{\overline{\alpha_s}(\overline{m})}{\pi} \right)^2 + \dots \right] + O(\Lambda_{QCD}), \qquad (2.5)$$

  and differ roughly by 10 GeV ($\overline{m}(\overline{m}) = 166.5$ for $m_{pole} = 175.6$) [23].

The perturbative propagator of a top quark with four-momentum p has a pole at the complex position $\sqrt{p^2} = m_{pole} - \frac{i}{2}\Gamma$ and yields a peak in the Wb invariant mass distribution when $\sqrt{p^2} \approx m_{pole}$. Given the experimental techniques used to extract the top quark mass (mass is reconstructed as the invariant mass of the products of top quark), these mass values should be therefore taken as representing the pole mass [21]. This will be implicitly used also in this thesis.

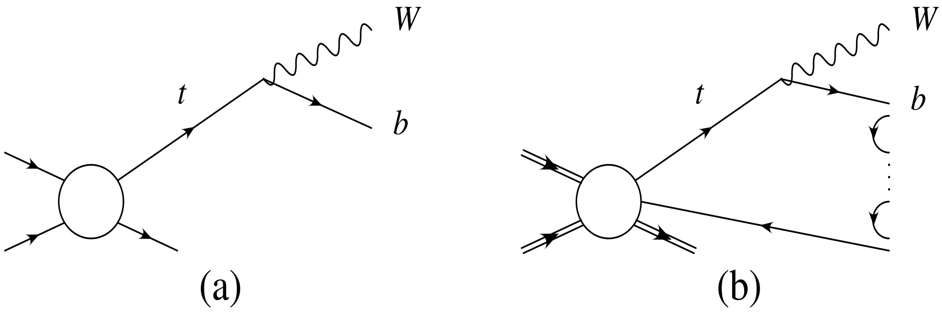


Figure 2.11: The production and decay of a top quark in (a) perturbation theory, and (b) nonperturbatively.

> Note, the pole quark mass can not be used to arbitrarily high accuracy because of non-perturbative effects in QCD. There is an intrinsic ambiguity of order $\Lambda_{QCD} \sim 200$ MeV [21]. For example, let's consider the extraction of top pole mass from the peak in Wb invariant mass distribution . In perturbation theory, the final state is a W and a b quark, as depicted in Fig. 2.11. However, the b quark manifests itself experimentally as a jet of colorless hadrons, due to confinement. At least one of the quarks which resides in these hadrons comes from elsewhere in the diagram, and can not be considered as a decay product of the top quark, as depicted in Fig. 2.11. This leads to irreducible uncertainty of $\sim \Lambda_{QCD}$ in the invariant mass and, hence, an ambiguity in the extracted top pole mass. However, at hadron colliders, experimental systematic uncertainties are limited factor in practice, so the uncertainty on top mass of only about 1 GeV is expected. As was mentioned above, masses in some renormalization schemes can have better intrinsic uncertainty and it is expected that $\overline{MS}$ mass can be measured up to ~ 150 MeV from energy scan of cross-section at linear collider (for details, see Sec. 5.3.2).

As already mentioned above, even if M_{top} is not predicted by the theory, it can be estimated within SM from precise electroweak measurements which give the top quark mass estimate $M_{top} = 179^{+12}_{-9}$ GeV [13].

The top quark mass have been measured at the Tevatron in series of experiments called Run I by both CDF and D0 experiments. The combined Run I average is $M_{top} = 178 \pm 4.3$ GeV. There are already many measurements of top quark mass in Run II (see Fig. 2.12). The latests combined value $M_{top} = 171.4 \pm 2.1$ GeV is in good agreement with above mentioned indirect estimate. The uncertainty on top quark mass at the end of Run II (2009) at Tevatron is expected to be $\Delta M_{top} \leq 1.5$ GeV.

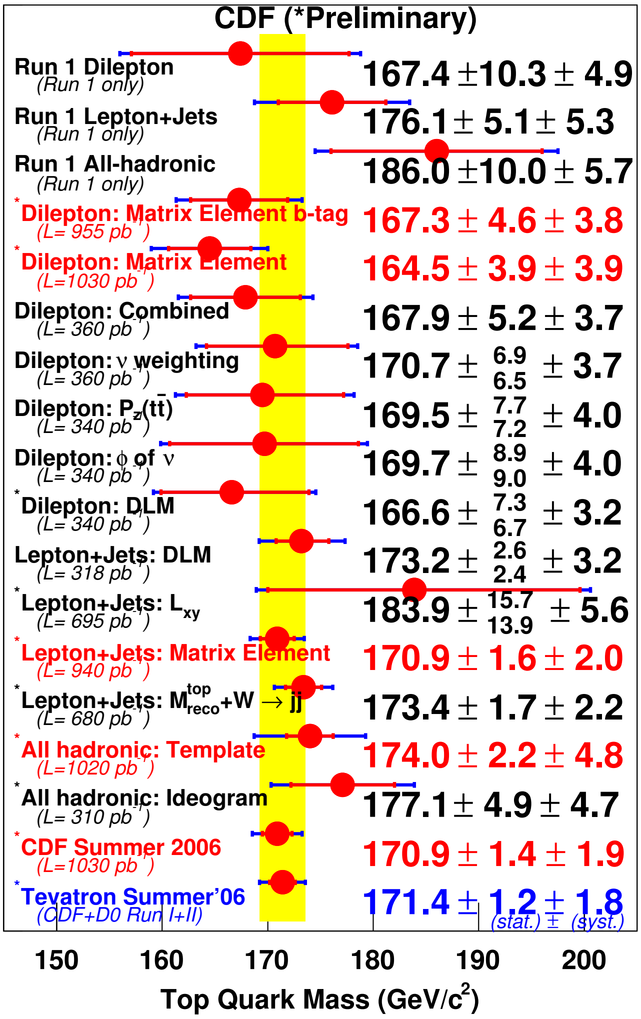


Figure 2.12: Tevatron Run II top quark mass measurements [24]. For each measurement, the red bar corresponds to statistical error of measurement while the blue bar correspond to total uncertainty of measurement (including systematic error).

One of the biggest reasons why the top quark mass is considered to be so important is that it is related to the Higgs boson mass. Fig. 2.13 shows the possible values of the Higgs mass according the measured mass of top quark and W boson. It can be seen that at 1σ confidence level there is already very tiny range for possible Higgs mass according the latest measurements of top quark mass. Knowing precisely what is the top quark mass will therefore tell us where the Higgs mass lies and where we should look for it (which channels to use, etc.).

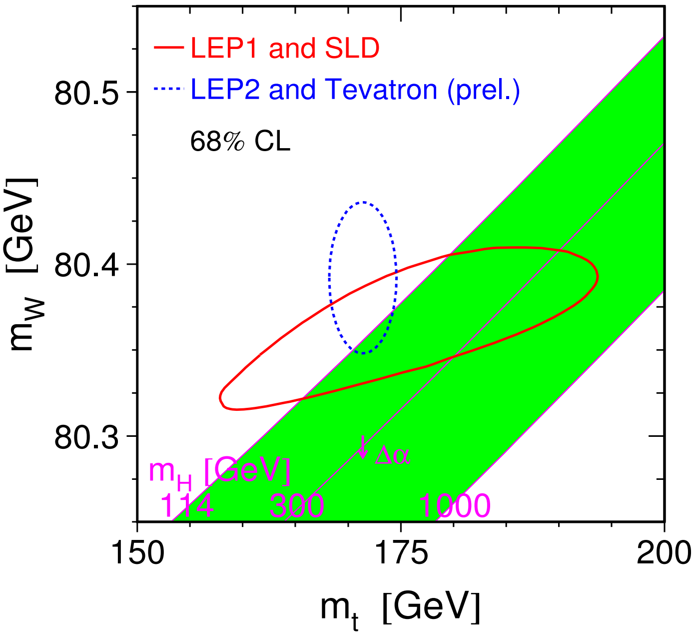


Figure 2.13: Possible range for Higgs boson mass according measured top quark and W boson masses [24]. The dotted curve corresponds to Tevatron measurements and to 68 % confidence level. The green region is the constrain on Higgs mass from direct searches.

Chapter 3

CDF experiment

The CDF[1] experiment[25] is performed at Fermi National Accelerator Laboratory[26] (familiarly called Fermilab) in Batavia, about 60 km west of downtown Chicago.

Fermilab, originally named the National Accelerator Laboratory, was commissioned in year 1967. Since then, two major components of the Standard Model were discovered at Fermilab: the bottom quark (1977) and the top quark (1995). In July 2000, Fermilab experimenters announced the first direct observation of the tau neutrino [27], the last fundamental fermion (lepton or quark) to be observed within SM. At present, there are 18 experiments running (or just analyzing already taken data) at Fermilab among which CDF and D0 (both being collider experiments) are the biggest ones. About 650 people from 60 institutions from 13 countries are the members of CDF collaboration as of now.

CDF experiment studies the collisions of protons p and anti-protons $\bar{p}$ which are both accelerated to the same energy at the accelerator Tevatron with the total center-of-mass energy of interaction being $\sqrt{s} = 1.96$ TeV.

The series of data taking, called Run I, proceeded between years 1992-1996. The huge success of this series of experiments was already mentioned experimental discovery of the top quark (1995).

In the year 2002, the Run II started taking physics data and is planning to run up to year 2009. The goal of this series of experiments is:

- characterization of the properties of the top quark, e.g. measurement of the top mass with uncertainty $\delta M_{top} \sim 1.5$ GeV (when CDF and D0 results will be combined). The current combined result is $171.4 \pm 1.2(stat.) \pm 1.8(syst.)$ GeV, which corresponds to a total uncertainty of 2.1 GeV [28].

---

[1] **C**ollider **D**etector at **F**ermilab

- a global precision electroweak program, e.g. measurement of W boson mass with uncertainty $\delta M_W \sim 20$ MeV. The current CDF result with 200 pb^{-1} of data is $80413 \pm 34(stat.) \pm 34(syst.)$ MeV $= 80413 \pm 48$ MeV [29], which is the most precise single measurement of W mass up to date.

- direct search for new phenomena, e.g. searches for super-symmetric particles like squarks, leptoquarks, extra dimensions, etc. Up to now, no physics beyond Standard Model was observed. However, many limits on particles masses (e.g. leptoquarks, stop quark, W', b') were set [30, 31, 32, 33].

- tests of perturbative QCD at next-to-leading order and large transfer momentum Q^2. The current CDF results on inclusive jet production using the midpoint jet algorithm (and also Kt algorithm) agree very well with the NLO QCD [34, 35].

- constraint of the CKM matrix with high statistics B decays, e.g. observation of the B_s oscillations. B_s oscillations were already observed at CDF [36].

One special challenge for the Run II is the observation of the Higgs boson. In the Standard Model, all the fundamental fermions gain mass from a field, called the Higgs field, which permeates the universe. It should be possible observe quanta of this field - the long sought Higgs boson. This is the last piece of the Standard Model to be observed, and is also the key to understanding any physics beyond the Standard Model.

As shown in Fig. 3.1, there have been performed already two estimates of the expected sensitivity to observe Higgs boson at Tevatron. The narrow curves in the figure are the estimations from 2003 Higgs Sensitivity Working Group study [37] where systematic uncertainties are excluded, and the thicker curves are the results from 1999 SUSY/Higgs Working Group study [38]. The 2003 study group benefited from more realistic Monte Carlo simulation and first Run II data. Each report includes calculations of the estimated luminosity (in fb^{-1}) required to exclude Higgs boson at 95% confidence level (CL), assuming a Higgs boson is not present, as well as the luminosity requirements for the evidence (at the level of 3 standard deviations) and 5σ discovery. If the mass of the Higgs boson is close to 115 GeV (this is very probable according Fig. 2.13), the Tevatron experiments should be able to observe 3σ excess with about 3 fb^{-1}. The projected total Run II integrated luminosity 8 fb^{-1} should be just enough for the discovery of Higgs boson.

In order to meet all these tough goals, these major components of the experiment are necessary to work:

- accelerator: to speed up the particles to such energies where we are able to produce interesting events

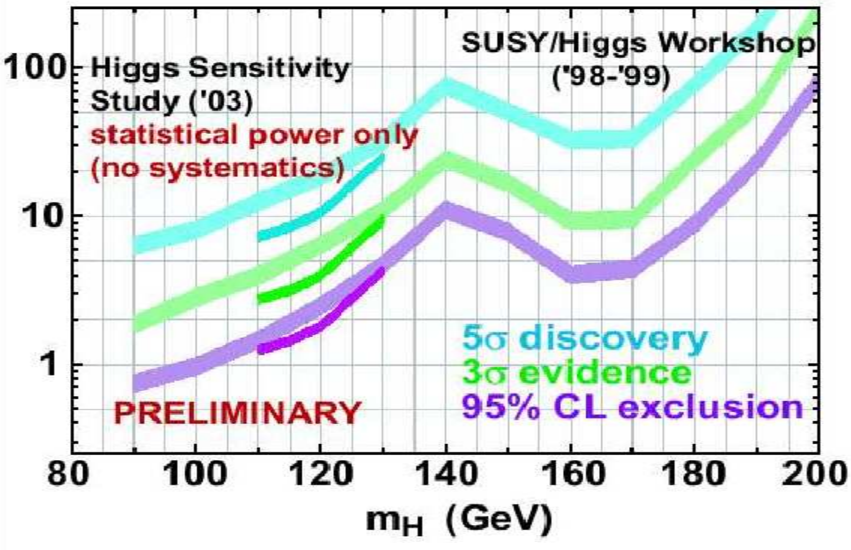


Figure 3.1: Integrated Tevatron luminosities in fb^{-1} per experiment (CDF & D0) needed for 95% confidence level exclusion, 3σ evidence and 5σ discovery of Higgs boson as a function of Higgs mass [37].

- detector: to register products of particles interactions
- offline reconstruction software: to provide users demanding physics quantities used for analyses (momentum, charge,...) out of the detector measured quantities (energy depositions in calorimeters, hits in the track detectors,...) and make these data easily accessible

We will describe them in the following sections.

3.1 Accelerator

The accelerating system consist of a few accelerators (Fig. 3.2) where the particles are gradually accelerated up to final energies.

In the first stage, the 750 keV Cockcroft-Walton pre-accelerator [39, 40] ionizes hydrogen gas and accelerates created negative ions H^- (proton with 2 electrons) using a positive voltage. The ions then pass into the 150 meter long linear accelerator – Linac [41], which uses oscillating electrical fields to accelerate the ions to 400 MeV. At the end of the Linac, the ions pass through a carbon foil, to remove the electrons, and the bare protons then move into the Booster. The Booster [42, 43] is a small

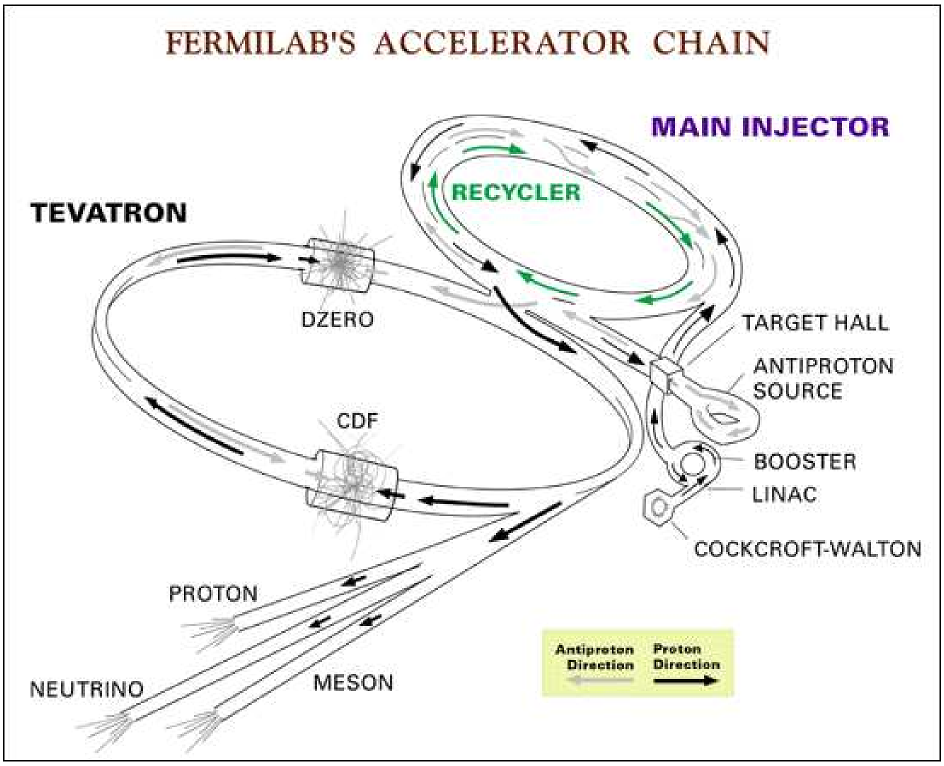


Figure 3.2: Accelerating complex in Fermilab.

(about 75 meters in diameter) circular accelerator (synchrotron) where protons attain an energy of around 8 GeV.

From the Booster, the particles pass into about one kilometer in diameter circular synchrotron - Main Injector [44] where a few things happen. First, the protons are accelerated up to 120 GeV in order to produce the anti-protons. These 120 GeV protons from Main Injector are directed out on the nickel target producing a range of particles including anti-protons which are collected and stored in the accumulator ring (Anti-proton Source). The ring then pass the anti-protons to the Main Injector. Back in the Main Injector the rest of the protons together with anti-protons are then accelerated (in the opposite directions) up to 150 GeV before injected to the Tevatron.

In Main Injector tunnel resides also Antiproton Recycler. During collisions at Tevatron, not all anti-protons are used up. At the end of collider store, there is still significant amount of anti-protons available (~ 75 %). Recycling the unused anti-protons and using them again, significantly reduce the time needed for production of anti-protons (stacking time), which is between 10 to 20 hours. Recycler takes the anti-protons from Tevatron at the end of store, “cool” them (reduce the transverse momentum spread

number of bunches	36
bunch spacing	396 ns
bunch length	50 cm
number of protons in bunch	$24 \cdot 10^{10}$
number of anti-protons in bunch	$6 \cdot 10^{10}$
anti-protons production rate	$1.6 \cdot 10^{11}$ / hour
maximum anti-protons in accumulating ring:	$2.4 \cdot 10^{12}$
beam radius	29(p), 21($\bar{p}$) μm
luminosity lifetime (store)	7 hours (average)
number of interactions per crossing	2.3

Table 3.1: Some parameters of Tevatron in Run II [21, 44].

within a bunch) and reintegrate into the stack, so they can be used in the next Tevatron store. Anti-proton availability is the most limiting factor at Tevatron to obtain large luminosities.

At the synchrotron Tevatron [44] (1 km in radius), the both protons and anti-protons are accelerated up to final energy of 980 GeV. The acceleration is provided, as in a cyclotron, by radio frequency oscillator that supplies energy increment every time a particle crosses an accelerating gap. To hold the particles on track, the Tevatron uses magnetic field that increases with time as the momentum of particles increases. Superconducting dipole magnets, cooled in liquid helium, producing magnetic field up to 4.2 T are used for that. The Tevatron can then sustain both beams for hours ("store"). The "run" is the continuous data taking (recording of interactions) by the experiment within a store, so there could be more runs within one store. Some parameters of Tevatron are summarized in Tab. 3.1.

The two main upgrades of Tevatron from Run 1 to Run 2 operation were increase of total energy of interaction from 1.8 TeV to 1.96 TeV and decrease of bunch spacing from 3.5 μs to 396 ns.

Tevatron is currently the world's highest energy particle accelerator ($\sqrt{s} = 1.96$ TeV) and will remain until Large Hadron Collider (LHC) will start its operation in summer 2007 where proton – proton collisions at center-of-mass energy $\sqrt{s} = 14$ TeV will occur.

The variable which characterize the performance of accelerator is called luminosity $\mathcal{L}$ (also "instantaneous luminosity"). Luminosity is the number of integrating particles per unit area per unit time. The integral of this quantity over the time is integral luminosity L, which together with cross-section σ for a given type of interaction (e.g. production of top-antitop pair) gives the number of produced events $N = L \cdot \sigma$. For

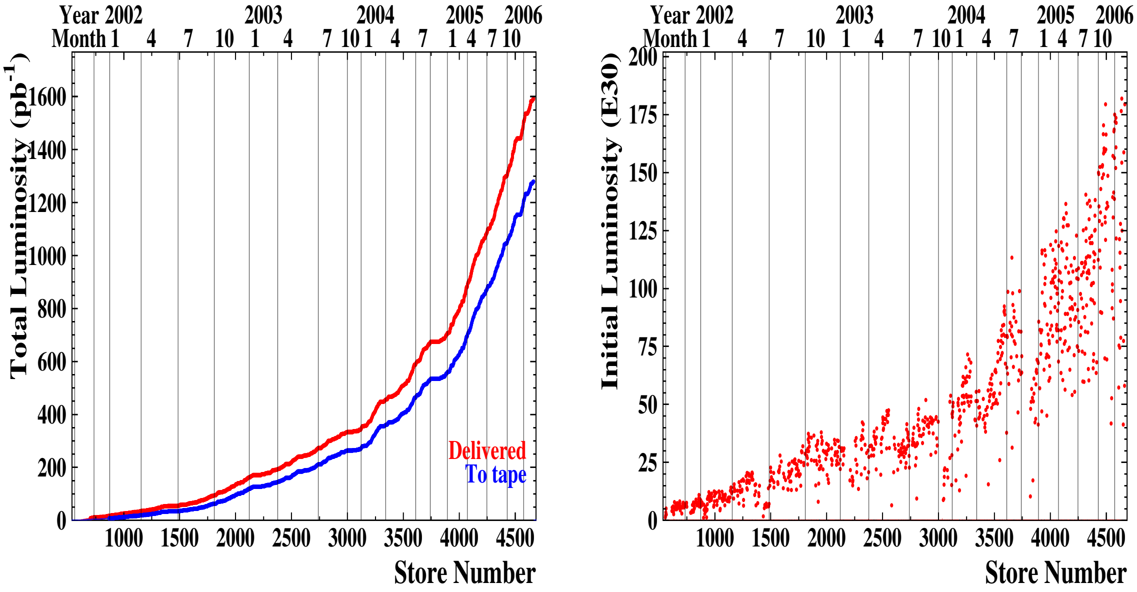


Figure 3.3: Integrated (left) and the instantaneous (right) luminosity at the CDF as a function of store number (axis below) and the time (axis above).

an collider, the luminosity can be defined as:

$$\mathcal{L} = f n_B \frac{N_1 N_2}{A}, \tag{3.1}$$

where f is the revolution frequency, n_B is the number of bunches in one beam in a ring, $N_i, i \in (1,2)$ is the number of particles per bunch and A is the transverse area of the beam. In real calculations, the luminosity is multiplied by a form factor which corrects for the bunch shape and depends on ratio of bunch length and the beta function at the interaction point β^*. The beta function is a measure of the transverse beam width. At the beginning of the collider store, the luminosity is the highest ("initial" or "peak" luminosity) because with the time, the number of particles in a bunch decreases.

The progress of the total integrated luminosity of the CDF in Run II is plotted in Fig. 3.3 (luminosity delivered by Tevatron and actual amount of data registered by CDF detector and stored to tapes).

For our measurement, we use the data collected in Run II from the beginning (march 2002) up to august 2004. This corresponds to the total integrated luminosity 340 pb^{-1} of physics data.

The total expected luminosity for Run II (up to year 2009) will mostly depend on $\bar{p}$ production and is expected to be between 4 fb^{-1} (base goal) and 8 fb^{-1} (projected luminosity). According to the latest performance of the Tevatron, it is expected that the delivered luminosity will be around 6 pb^{-1}.

The instantaneous luminosity of $1.8 \cdot 10^{32}$ cm^{-2}s^{-1} was reached as it is showed in the Fig. 3.3 and is expected to rise up to $\sim 3 \cdot 10^{32}$ cm^{-2}s^{-1} later during the Run II (this is roughly 100 times lower than what is expected at the LHC).

The two detectors were build along the Tevatron ring: CDF and D0. They both register the interactions of protons and anti-protons at the total energy of interaction $\sqrt{s} = 1.96$ TeV in center-of-mass system.

3.2 CDF detector

CDF collaboration detector [45] (Fig 3.4) is multipurpose 5000 tons heavy solenoidal detector complex which registers collisions of protons with anti-protons accelerated at Tevatron. In process of collisions many different types of particles are produced, such as stable particles (electrons, muons[2],...), jets[3], neutrinos, short lived particles (pions, kaons). The goal of the detector is to gain as much information about the interaction as possible.

Because the CDF detector register collisions of the same energetic protons and anti-protons, it was proposed with forward-backward and also azimuthal symmetry.

The detector combines precision charged particle tracking with fast projective calorimetry and fine grained muon detection. The closest to the interaction point are located tracking systems which are contained in 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 (anti-parallel with the direction of incoming protons). Calorimetry and muon systems are all outside the solenoid.

Major differences from Run I to Run II detector include: the replacement of the central tracking system; the replacement of a gas sampling calorimeter in the plug-forward region with scintillating tile calorimeter; preshower detectors; extension of the muon coverage; a Time-of-Flight (TOF) detector and upgrades of trigger, readout electronics, and the data acquisition systems.

The detector systems most important to our analysis are calorimeters (they register and measure energies of jets and electrons) and muon and tracking detectors which register and measure momenta of muons, respectively. The main features of these detectors are summarized below, but first we describe the coordination systems and used variables.

---

[2] muons, of course, decay. However, they live long enough to traverse whole detector.

[3] these are the showers of particles flying in one direction which come out from the hadronization of original quark or gluon

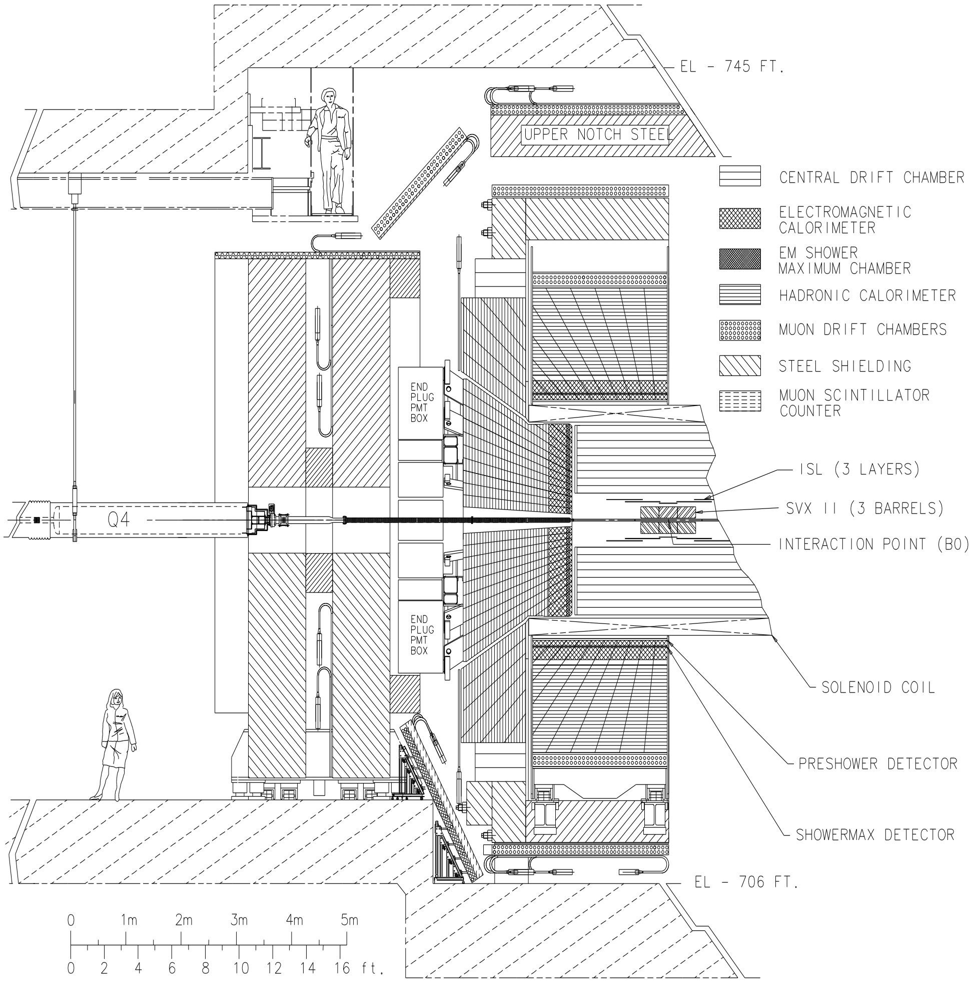


Figure 3.4: Schematic view of the CDF detector.

3.2.1 Definition of coordinate system and used variables

The coordination system associated with the detector is as follow. The positive direction of z axis is parallel with the incoming protons. The $x - y$ plane is perpendicular to z axis. The x axis is horizontal pointing outward from the accelerator ring and y axis is vertical (the positive direction of y axis is upward), so this system is right-handed coordination system.

Other, very common coordinates are r, θ, ϕ and η. The r distance is just the distance in transverse plane $r = \sqrt{x^2 + y^2}$. The angle θ is the polar angle measured from the proton direction and the angle ϕ is the azimuthal angle measured from the x axis. The pseudo-rapidity is defined as $\eta = -\ln(\tan(\theta/2))$ and is useful coordinate, since the difference $\Delta\eta = \eta_1 - \eta_2$ is the Lorentz invariant.

We define the transverse momentum as the momentum in the plane perpendicular to the beam axis, $P_T = P \sin\theta$ and it is measured by the tracking detectors. The transverse energy, $E_T = E \sin\theta$, is essentially the same except the measurement comes from the calorimeters. Therefore, although E_T is referred to as an energy, it is in this definition the 2 dimensional vector quantity in the transverse plane $(E_{T_x} = E_T \cos\phi, E_{T_y} = E_T \sin\phi)$. Since, the measurements of muons momenta come from the tracking detectors we will refer to their momenta with P_T while the measurements of jets and electrons momenta comes from calorimeters, so will be referred to with E_T.

3.2.2 Tracking detectors

The track detectors allow us:

- to calculate particle characteristics such as momentum, sign of the electromagnetic charge, decay length, etc. using measured track parameters of particles (e.g. curvature).
- to identify the particles in case we combine the information from track detectors with information from other detectors, e.g. if the particle has a track in the track detectors and to this track belongs also the hits in the muon detectors, we can identify such particle as the muon.

The uncertainty in the momentum measurement using the tracking detectors can be deduced the following way. In homogeneous magnetic field, tracks of charged particles travel on a helix with the axis of the helix parallel to the magnetic field. In transverse plane, the particle travels on a circle. For each track, we are measuring the bend angle α, which is the angle the momentum vector is rotated by in the magnetic field. This angle is inversely proportional to the radius of the circle and thus to the particle momentum[4], $\alpha \sim 1/P$. The fractional momentum error has a term due to angular error $\mathrm{d}\alpha$ which is proportional to the momentum [5]

$$\mathrm{d}\alpha \sim \frac{\mathrm{d}P}{P^2} \Rightarrow \frac{\mathrm{d}P}{P} \sim (\mathrm{d}\alpha)P = (const.) \cdot P \oplus d, \tag{3.2}$$

where the resolution of the angle α measurement is given by the detector properties and it is constant value for a given detector. There is artificially included in the relative momentum resolution formula the additional term d due to multiply scattering.

---

[4] the particle with the charge Q moving in the magnetic field B with the momentum P transverse on magnetic field, has a radius of circle $r = P/(B \cdot Q)$

[5] $\oplus$ means that the uncertainties are independent and should be added in quadrature to obtain the total uncertainty

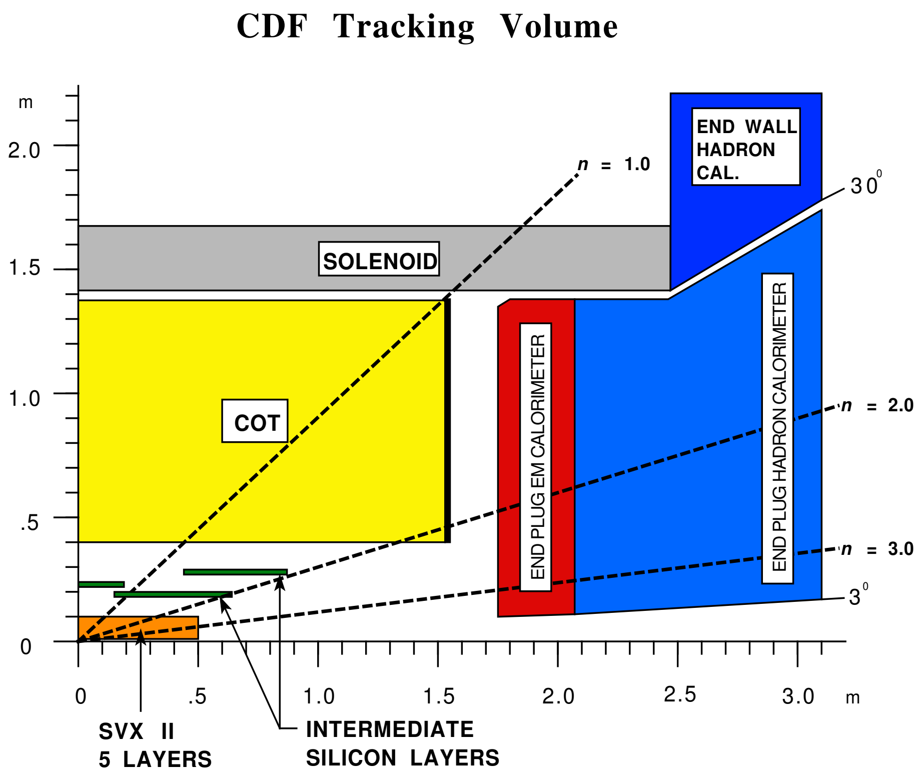


Figure 3.5: A view of one quadrant of CDF tracking systems surrounded by the calorimeters.

The CDF tracking system (Fig. 3.5) consists of silicon micro-strip system [46] and of an open-cell wire drift chamber [47] that surrounds it.

Vertex detector system

At the CDF, we use silicon strip detectors as the vertex detectors. These are semiconductor detectors which principal of working is as follows. Semiconductor detectors are $p - n$ junction diodes which are operated at reverse bias. This creates a sensitive region depleted of mobile charge and sets up an electric field. As charged particles pass through the strips of semiconductor material, they cause small ionization currents which are transfered to electrodes and can be measured. The amount of charge gathered is, to first order, proportional to the path length traversed by the charged particle in the detector material. The advantage of silicon detectors is their much higher resolution in tracking charged particles than older technologies such as cloud chambers or wire chambers.

The CDF silicon micro-strip detector (see Fig. 3.6) consists of seven layers (eight

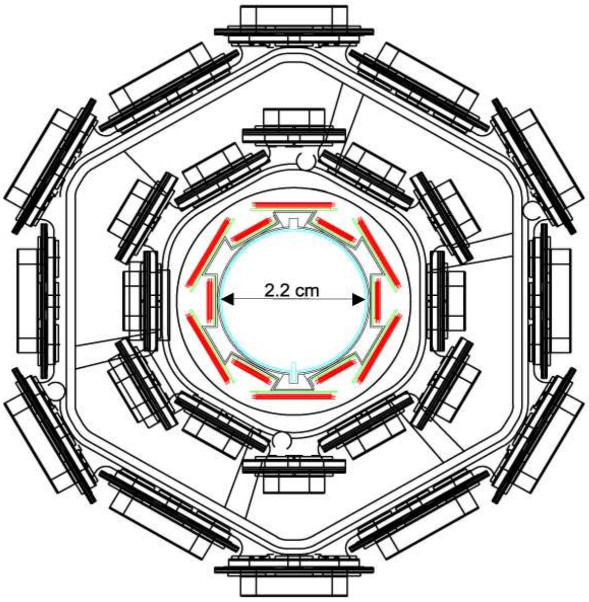


Figure 3.6: End view of the innermost three layers of the CDF Run II silicon system, showing Layer 00 along with the first two layers of the SVX II detector.

layers for $1.0 < |\eta| < 2.0$) in a barrel geometry that extends from radius of r = 1.5 cm from the beam line to r = 28 cm. The layer glued directly on the beam pipe is a radiation-hard, single sided detector called Layer 00. The remaining seven layers are radiation-hard, double-sided detectors. The first five layers after Layer 00 comprise the SVX II system and the two outer layers comprise the ISL system. The Layer00 and SVX II layers are arranged in 12 azimuthal wedges while the ISL has 24 azimuthal wedges. The SVX II sensors are arranged in three barrels with a total length of 96 cm while ISL has a total length of 175 cm. Of the five SVX II double-layers, three layers (1st, 2nd, 4th as counting from the innermost) have 0° and 90° strips with respect to beam axis ("stereo layers"), while the other two (3rd and 5th) have 0° and ±1.2° small angle strips ("1.2° stereo layers"), respectively. ISL has also 2 double-layers with axial and small angle stereo (1.2°) strips. This was designed in order to provide good resolution in locating the z position of secondary vertices and to allow track reconstruction in three dimensions in vertex detector alone.

The spatial resolution in SVX II is 12 μm and in ISL 16μm. The combined resolution of SVX II + ISL system in the forward region $|\eta| > 1.0$ is $\delta p_T/p_T^2 \sim 0.4\%$ for transverse momentum. The impact parameter resolution of the combination of SVX II and ISL is 40 μm allowing excellent identification of secondary vertices.

COT

The principal of drift chamber working is as follows. Charged particles ionize the atoms while passing the volume of the detector filled with gas (at the CDF is used mixture argon:ethane(50% : 50%)+isopropyl). The ionization electrons are eventually collected by anode wire where strong electric field accelerates the electrons to produce secondary ionization and hence the avalanche. Detecting signal on wires (hits), we can follow the track of the particle in the detector. In addition, the energy lost of particle can be obtained from measuring of the signal amplitude at the wire, because the pulse width is related to the total charge deposited on the wire.

In CDF detector, there is 3.1 m long cylindrical drift chamber (COT) which provides coverage for $|\eta| < 1$. In the radial range, it goes from 40 to 137 cm and provides 96 measurement layers, organized into alternating axial and $\pm 2°$ stereo super-layers (total 8 super-layers), see Fig. 3.7. Each superlayer is divided in ϕ angle into "supercells" and each supercell has 12 sense wires and maximum drift distance approximately the same. Therefore, the number of supercells in given superlayer increases with the radius. The supercell is tiled by 35° with respect to radial direction to compensate for the Lorentz angle of the drifting electrons in the magnetic field.

The COT hardware provides, as its output, the drift time for each detected hit on COT wire. This time measurement is converted into a position measurement. The hit position resolution is approximately 140 μm. A reconstructed track provides accurate information in the $r - \phi$ plane for the measurement of the transverse momentum p_T, and substantially less accurate information in the $r - z$ plane for the measurement of pseudo-rapidity η. The momentum resolution for tracks is $\sigma(p_T)/p_T^2 = 0.15\%\ (\mathrm{GeV})^{-1}$.

3.2.3 Calorimeters

At the CDF, segmented electromagnetic and hadronic sampling calorimeters surround the tracking system and measure the energy flow of interacting particles in the pseudo-rapidity range $|\eta| < 3.64$. Electromagnetic calorimeters measure energy of electrons and photons while hadronic are used for the measurement of jets energies.

In general, a sampling calorimeter consists of an active medium which generates signal and a passive medium which is used as an absorber. The active medium may be a scintillator, liquid, a gas chamber, or a semiconductor. The passive medium is usually the material with high density, such as iron, lead, cooper, etc. At CDF, there are used as active medium organic scintillators.

The principle of detection in organic scintillators is as follows. The charged particle

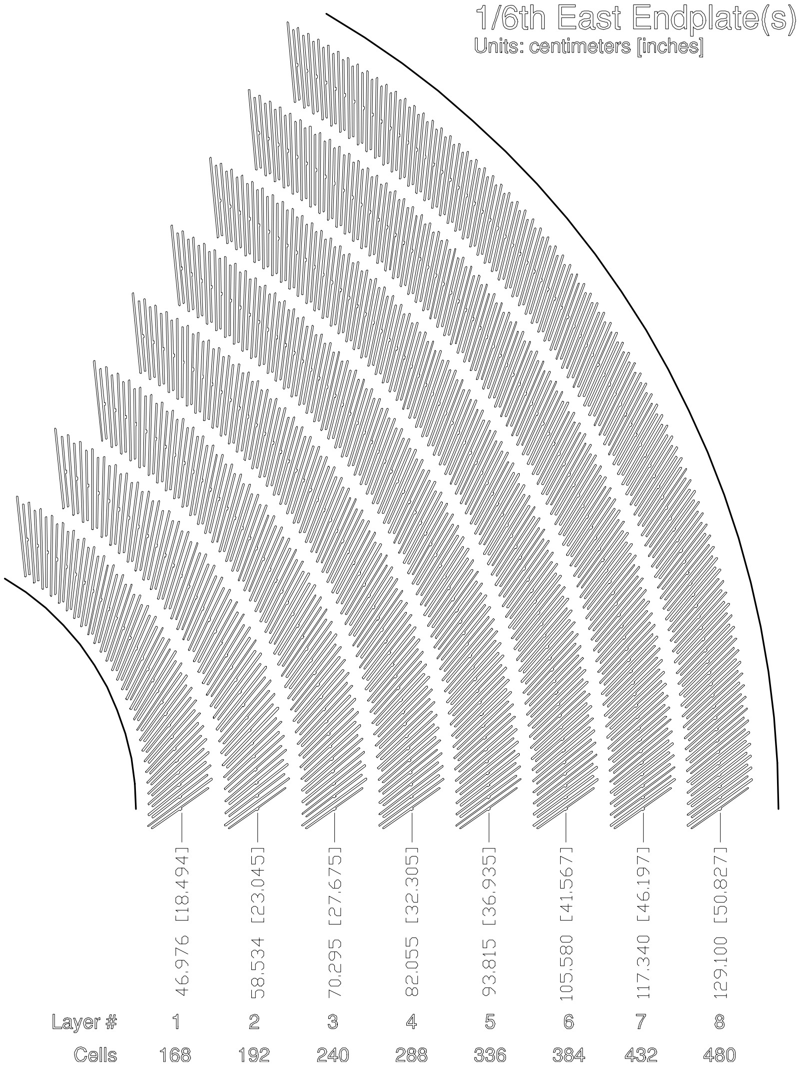


Figure 3.7: 1/6 section of the COT end plate. For each superlayer is given the total number of supercells and the average radius.

which travel the matter leaves behind it excited molecules. The molecules will release some fraction of its energy as optical photons which are then collected by photomultipliers. The amount of produced photons is proportional to the energy deposited by traveled particle.

The uncertainty in the energy measurement with the calorimeters can be deduced in following way. The energy deposited in calorimeter is proportional to the number of particles left in detector, $E \sim N$. As the number of particles left in detector is stochastic, it follows Poisson distribution, the uncertainty in number of particles is $\sigma(N) = \sqrt{N}$. Therefore, the uncertainty in the energy measurement is $\sigma(E) \sim$

$\sigma(N) \sim \sqrt{N} \sim \sqrt{E}$. The relative uncertainty of energy is thus inverse proportional to the square root of energy deposited in calorimeter $\sigma(E)/E \sim 1/\sqrt{E}$. In practice, there are additional terms in uncertainty relation:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E}, \tag{3.3}$$

where the constant term b is due to detector non-uniformity and calibration uncertainty. The term c is due to electronic noise, in absolute value it is just the constant.

Central calorimeters

The central calorimeters (and the end-wall hadronic calorimeter) cover the pseudo-rapidity range $|\eta| < 1.1(1.3)$.

The central electromagnetic calorimeter [48] (CEM) uses lead sheets (13 mm thick) interspersed with polystyrene scintillator (5 mm thick layers) as the active medium. The total depth of CEM is $18X_0$ [6]. It has tower segmentation, where each tower is 15 degrees in azimuth (in total, 24 towers in azimuth) and about 0.11 in pseudo-rapidity (in total, 10 towers in pseudo-rapidity). Its energy resolution is $13.5\%\sqrt{E_T} \oplus 2\%$. In the depth $5.9X_0$, there are strip chambers (CES) which determine shower position and its transverse development at shower maximum by measuring the charge deposition on orthogonal strips and wires.

The central (CHA) and end-wall hadronic calorimeter (WHA) [49] use steel absorber (2.5 cm and 5.0 cm thick layers for CHA and WHA, respectively) interspersed with acrylic scintillator (1 cm thick layers in both CHA and WHA) as the active medium. The CHA has exactly same tower segmentation as the central electromagnetic calorimeter. Its relative energy resolution is $75\%\sqrt{E_T} \oplus 3\%$.

Plug calorimeters

The plug calorimeters are once again sampling calorimeters. The top half of one plug is shown in Fig. 3.8. There is an electromagnetic section followed by hadronic calorimeter where both cover the pseudo-rapidity region $1.1 < |\eta| < 3.6$. The segmentation in pseudo-rapidity η and azimuthal angle ϕ from central calorimeters is followed also here.

The plug electromagnetic calorimeter (PEM) is a sandwich of 4.5 mm lead and 4 mm scintillator layers. The total depth of PEM is $20X_0$, where in the depth $5.4X_0$ is placed shower-max position detector (PES), similarly as in central electromagnetic

---

[6] X_0 is radiation length, which is defined as the distance at which the particle loose all but $1/e$ of its energy

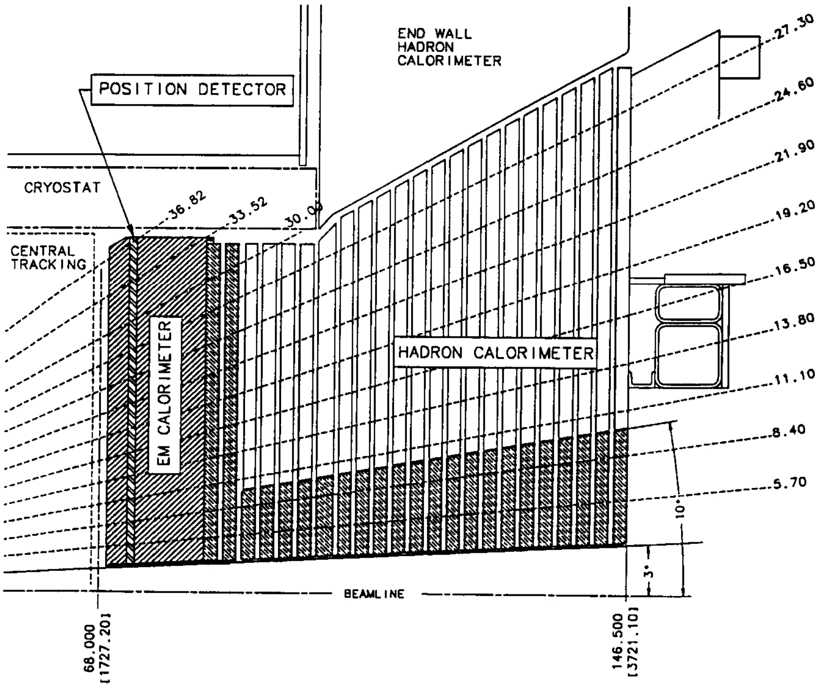


Figure 3.8: Cross-section of upper part of plug calorimeter.

calorimeter. It consists from two layers of scintillator strips with a 45° crossing angle. These are referred to as "U" and "V" strips.

The hadronic calorimeter (PHA) has layers of 5.1 cm iron and 6 mm scintillator. In both, PEM and PHA, the scintillator is polystyrene.

The relative energy resolution of PEM is $16\%\sqrt{E} \oplus 1\%$ [50] while for PHA it is $74\%\sqrt{E} \oplus 4\%$.

3.2.4 Muon detectors

The muon detectors are used for identification of the muons. At CDF, muon detectors consist of drift chambers and scintillators. For description of principals of drift chamber and scintillator working, see Sec. 3.2.2 and 3.2.3.

The four chamber muon system (CMU, CMP, CMX, BMU) measures drift time which is converted to a drift distance, i.e. a distance from the wire to a location that muon had occupied in its flight, in the plane perpendicular to the chamber sense wire.

Scintillator counters are needed for association of muon track to $p\bar{p}$ interaction that produced it, since the typical drift time is larger than bunch crossing time.

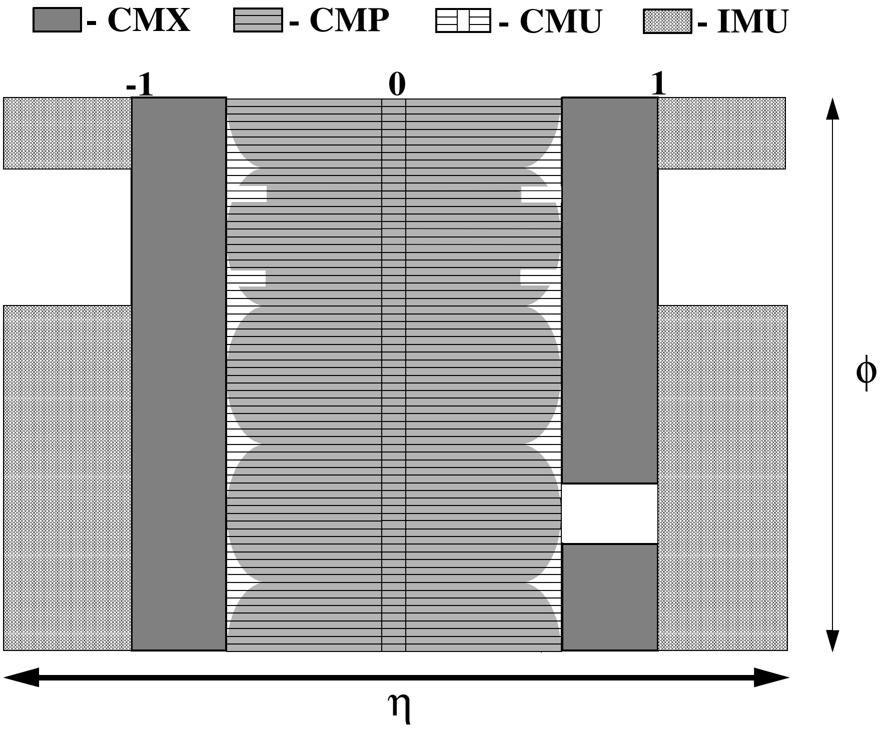


Figure 3.9: Location of muon detector systems in azimuthal angle ϕ (across whole range from zero to 2π) and pseudo-rapidity η.

The muon detection system resides beyond the calorimetry. There are 4 different detectors for muon identification. The muon detectors coverage in the plane of azimuthal angle ϕ and pseudo-rapidity η is shown in Fig. 3.9.

The Central MUon Detector (CMU) is embedded in the central calorimeter wedges at their outer radius. It has four layers of planar drift chambers. The CMU can detect muons with the momentum $p_T > 1.4$ GeV.

An additional four layers of planar drift chambers, Central Muon Upgrade (CMP), are situated behind the 0.6 m of steel. The part of this detector are also scintillation counters (CSP). CMP provides confirmation for CMU tracks but with reduced background. The CMP can detect muons with at least transverse momentum of 2.2 GeV. The CMP chambers are of fixed length in z and form a box around the central detector. Therefore, the pseudo-rapidity coverage varies with azimuth as shown in Fig. 3.9. Both, CMU and CMP chambers provide coverage in the pseudo-rapidity range $|\eta| \leq\sim 0.6$.

The Central Muon eXtension (CMX) consists of drift tubes (CMX) and scintillation counters (CSX) covering pseudo-rapidity $\sim 0.6 \leq |\eta| \leq\sim 1.0$. The muons with transverse momentum greater than 1.4 GeV can be detected there.

The Intermediate MUon chambers (IMU) are extending the coverage to the region $\sim 1.0 < |\eta| <\sim 1.5$. They can identify muons with $p_T \in (1.4, 2.0)$ GeV depending on the pseudo-rapidity of particle due to different amount of detectors residing before IMU. The IMU system consists of 4-layer barrel of drift tubes (BMU) and scintillators (WSU, BSU, TSU). However, in this analysis we don't use muons from IMU chambers.

3.2.5 Trigger systems

The trigger and data acquisition systems are designed to accommodate the high rates and large data volume of Run II. Roughly 1.7 Million events are produced every second. Average stored event have size $\sim 200\ kB$ which would mean $\sim$ 300 GB of data per second to be stored. This is impossible at the present and moreover it would be big waste of resources since the majority of the events are not interesting. Therefore, the events must be selected in order to have at the end the rate $\sim$ 20 MB/s of data which can be stored. This is accomplished by trigger system which selects only interesting events according decisions based on some preliminary, quick information from the detectors.

CDF implements three level trigger system, see Fig. 3.10. A set of requirements that an event must fulfill at Level 1, Level 2 and Level 3 make a trigger path.

Level 1

The first level, Level-1 (L1) uses custom designed hardware which restrict the event rate based on preliminary information from COT, central calorimeters and the muon systems to select sufficiently interesting collisions. Because the original design for occurrence of interactions was every 132 ns (however, interactions really occurs every 396 ns) and there exist 42 events pipeline, there is total time of 5.5 μs available for processing the event.

L1 makes a decision based on simple counting of the physics objects, e.g one 10 GeV electron or two 8 GeV muons.

The improvement comparing to Run I comes from including the track information to the L1 trigger which have been achieved by eXtremely Fast Tracker (XFT) device. XFT allows track reconstruction in transverse plane with track momentum resolution $\Delta p_T / p_T^2 < 2\%$ for tracks with $p_T > 1.5$ GeV. Using XFT in L1 trigger, it allows a track to be matched to an electromagnetic calorimeter cluster for improved electron identification and a track to be matched to a stub in the muon system for better muon identification and muon resolution.

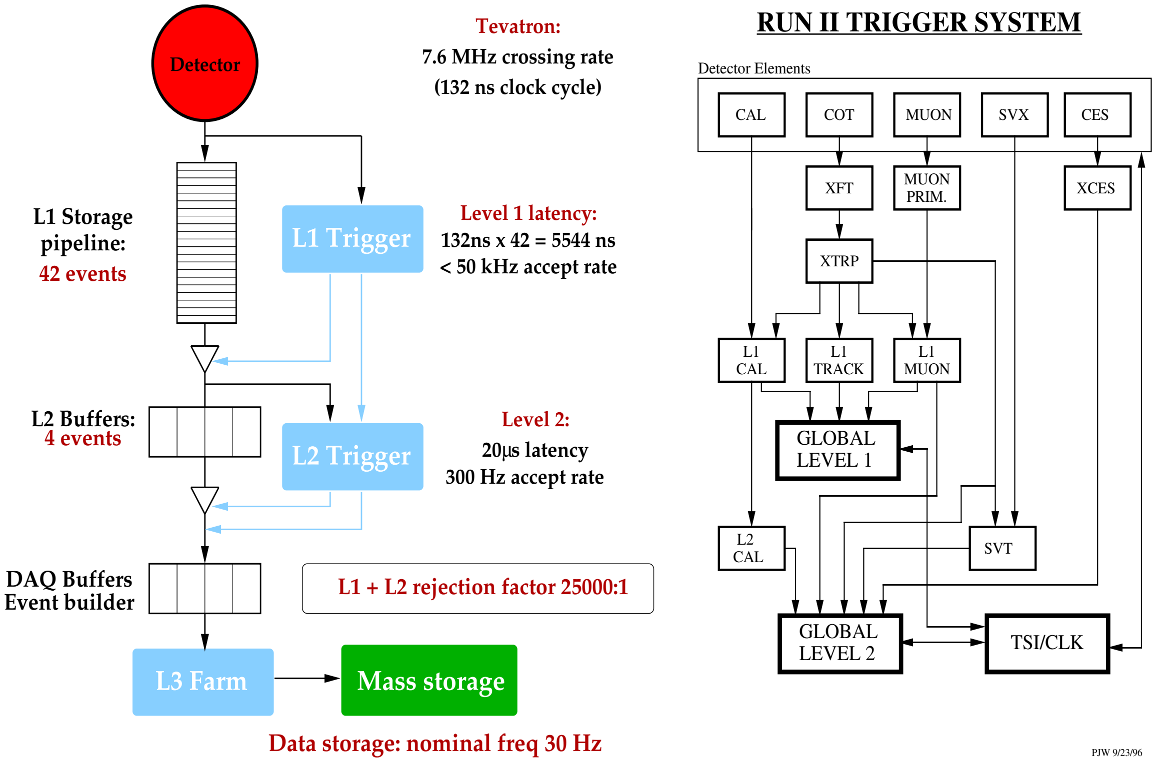


Figure 3.10: The Run II trigger system pipeline (left) and block diagram of detector information usage in trigger (right). The numbers shown here are original Run II estimates which changed in last years. The more appropriate numbers are in text.

The output of the Level-1 trigger limit the rate of the events by factor of ~ 150 to $\approx 18\ kHz$ at the luminosity range of $3 - 7 \cdot 10^{31} cm^{-2} s^{-1}$.

Level 2

At the next stage, Level-2 (L2) hardware trigger use more refined information and additional tracking information from the silicon detector. The data are transfered to one of the four Level-2 buffers, so the total time for processing at the Level-2 is $\sim 20\mu s$. This is asynchronous system, it means the event stays in the buffer until it is rejected or accepted and during that time no other event can be processed. This can cause dead time.

L2 performs limited event reconstruction which can be processed in programmable processors. All the information used in L1 is also available at L2, however with better precision. Jet reconstruction is available at the L2 with cluster finder, and secondary vertex information is available from Silicon Vertex Tracker (SVT) device.

The most significant addition at the Level-2 in Run II comparing to Run I is the information from SVX which enables to trigger on secondary vertices which is for the first time in hadron collider experiment.

The event rate is reduced by factor 60 to $\approx 300\ Hz$. Data are collected to Data Acquisition (DAQ) buffers and then transfered to a Level 3 CPU node.

Level 3

The final level of the trigger – Level-3, with access to the complete event information, uses software algorithms and a Linux computing farm. Results from lower levels are used to drive the algorithms.

Each event is processed by one CPU. Because there are approximately 300 CPUs in the PC farm there is roughly 1 second for event to be reconstructed. As a consequence, the quality of reconstruction is near the quality of "offline" reconstruction.

Level-3 reduces the output rate by factor of 4 to $\approx 75\ Hz$. Such number of events are then written to permanent storage (tapes).

3.3 Offline reconstruction

In this section, we will describe whole machinery of how the data delivered from detector to permanent storage are made useful for physicists to perform their analyses. This is called "offline" procedure because it happen a few days, weeks and months later than data were taken as comparing to "online" procedures which happen directly during data taking.

3.3.1 Data storing

The data events which pass the triggering system (see section above) are divided into "streams" and written on tapes. There are 8 different streams at CDF. In each stream there are usually collected events with the similar event characteristics.Usually, the events are divided into files with approximately 1 Giga bytes (GB) size. The tapes have size of about 200 GB (older ones have just $\sim 40GB$ and the tape robot can handle about 10000 tapes. Up to now, there have been written to tapes roughly 1.6 PB of data (real experimental data and also Monte-Carlo simulated data). Therefore, there was lately need for another robot to be able to store new incoming data. The files where the data are stored are in ROOT [51] format.

Of course, without information about what kind of data and where on which tape are stored, the stored data would not be very useful. It is needed to have all the useful information (name of the file, file size, date it was created, number of events, etc.) about data easily available. Therefore, there exists system which stores every needed information about every file which is stored on tape and according specified informations it can deliver the specified data (datasets, files, events, etc) to the user or just informations about stored data. CDF collaboration started in last years to use SAM (**S**equential data **A**ccess via **M**eta-data) [52] data handling system for this purpose.

3.3.2 Calibrations

Before the data will be processed, there is needed to know some parameters (alignment of detectors, position of beam, etc.) about every kind of detector in order to correctly reconstruct physical quantities. These parameters are called calibrations.

Some preliminary calibrations are already obtained "online", i.e during data taking at the time of processing the events through trigger system. However, the final calibrations are obtained "offline", i.e. only after the data events were passed through trigger system and written to tapes.

Calibrations are extracted from specialized datastreams and performed approximately every 200 pb^{-1}. First, only SVX and COT calibrations are performed. Then, using the information about tracks from SVX and COT (these already use calibrations obtained in previous step), the beam parameters are fitted ("beamlines"). The beamlines are calculated on run by run basis. The knowledge of the beam position in x and y gives a good estimate of the primary vertex position of an event [7]. Moreover, the beams are not centered in the detector and not parallel to the detector z-axis. Therefore, there is non zero slope of the beamline which is also determined. After that, calibrations for other detectors are produced, validated and finally written to database in order to be available for later processing of data.

3.3.3 Reconstruction software

All the CDF software is organized into releases and within each release there exists many packages (order of hundreds) used for each particular task, e.g. there exist Muon packages for reconstruction, validation, etc. of the muons. For completeness, in our analysis, we have used the release 5.3.3 of CDF software.

Each piece of code used by CDF collaboration is stored within CVS management system [53]. This system allows multiplier users to edit the same files, it saves all the versions of the code, it allows to return to previous versions of the code, etc.

---

[7] The precise knowing of the position of the primary vertex is important in many analyses (e.g. all analyses which need to identify b-jets).

Therefore, it is suitable for such a big collaboration where many people contribute to the development of software.

One big technical change happened in reconstruction software during the transition from Run I to Run II. All the programs were rewritten from Fortran to object oriented C++. The framework used is AC++ which have been developed for particle physics experiments. AC++ allows modular structure, so that pieces of the code with different functionality (e.g. reconstruction of electrons) and lots of parameters can be switch on/off or set, directly at the run of the program.

The description of how the particular physics objects (electrons,, muons, jets, etc.) are identified and reconstructed will be presented in next chapter (see Sec. 4.1).

3.3.4 Data processing

For reconstruction of the data event which passed trigger requirements and were stored to tapes ("raw data") is used one reconstruction program (called ProductionExe). This is C++ executable which reconstructs all the physics objects within event.

The data reconstruction is performed on large Linux based PC farm [54, 55]. The number of nodes [8] in the farm used for data reconstruction is order of 500. The Farm Processing System (FPS) was the software that managed, controls and monitored the CDF production farm from 1999 to late 2005. I used to manage this software and used it to perform the data reconstruction on the PC farm.

The processing of the data on the farm consists of a few steps. First, the data are staged from the tapes to "dfarm" consisting from disks areas of all the nodes in the PC farm. Then, the jobs are submitted to the farm through batch manager. The underlying system for submission used to be Fermilab developed FBSNG batch system [56], however it was lately replaced by Condor system [57]. In each job looks for input file which is then copied to local disk. The reconstruction program runs locally on the worker node on one data file (its size is about 1 GB) and reconstruct all the objects for all events. The reconstructed data events are divided into more output streams according what kind of criteria the event pass. It can happen that one event can be stored in more output streams, however, it is not very likely. The output files are written locally. They are of different file size and at the end of job are copied back to dfarm. size. The relevant files like log files, histogram files, etc. are collected to a place where members of the collaboration can easily access them for the need of validation or monitoring purposes. At the end of processing chain, we collect the

---

[8] PC which is part of the farm is called "node". Usually, one node has 2 CPUs, so it runs at least 2 (mostly 3 jobs).

data files to some specially dedicated nodes in farm, where the files from each stream are concatenated together in the files of roughly 1 GB in size. The last part of the reconstruction process is writing the reconstructed data back to the tapes and update the record in database. The farm is able to operate at the level of about 40 MByte/sec.

Lately, there was performed a production farm upgrade, in order to accommodate the increasing data acquisition rate and the migration of the CDF data-handling system to SAM. The PC cluster acting as the data production farm was upgraded with the CDF Analysis Farm (CAF) software [58]. The upgrade of the CDF data processing has incorporated the advances of SAM and CAF and is compatible with recent GRID computing development at Fermilab [59].

Monte-Carlo production

In the high-energy physics experiment, it is very important to have simulated data. They can be compared to the real data and if in agreement, can be used for many studies. The Monte-Carlo (MC) based programs are used for generation of events.

The procedure of generating MC samples is technically almost the same as for data. Of course, the only difference is the generation of MC samples have no input data files as opposite to real data processing. Other than that, the MC data are also processed on PC farms. Moreover, because of essentially no input data, the MC samples can be produced wherever in world. Currently at CDF, there are order of 10 PC farms around the world which are used for MC generation. This procedure is increasingly being included into world wide Grid computing.

3.3.5 Accessing data by users

The final goal of all the computing at the HEP experiment is to provide easy and quick availability of data to physicists. Of course, people can take a look even at “raw data” but over the years at CDF there have been provided better, more efficient way of accessing data. First, as it was described above, there is dedicated farm where all the “raw data” are processed and the events are divided into many outgoing streams. This way, everyone needs to look at only small portion of all data and not at whole 1.6 PetaBytes of data! However, even such processed data are quite big (order of TeraBytes). Moreover, you usually have there only basic objects reconstructed and no cuts applied on objects, so most of the events is not useful for your analysis. Therefore, each physics group at CDF (top quark group, b quark physics group, etc.) have provided programs which will select events according some criteria to select really only interesting events and it will create ntuple files [9] which are already easy to use in the analyses (the size of these data is order of tens of GigaBytes). For example, top quark physics group provides program 'TopFind' which can select the events according many criteria on electrons, muons, jets or even criteria on $Z, W, t\bar{t}$ candidates, etc.

The ntuple files are then stored on tapes and/or dedicated fileservers where people can easily access them and perform their analyses only on really small portion of all data.

---

[9] ntuples are arrays of different variables (momentum, charge, etc.) where you have association of all the variables to particular event.

Chapter 4

$t\bar{t}$ dilepton event selection

We use $t\bar{t}$ dilepton final state in order to measure the top quark mass. The final state particles of the $t\bar{t}$ dilepton event production are charged leptons (e, μ), neutrinos and jets of particles (see sec. 2.2.2).

In first part of chapter, we will describe how we identify all of the above mentioned objects, using the information obtained by detectors.

In second part, we will describe all the criteria we put on these objects to select as much as possible signal $t\bar{t}$ dilepton events while rejecting as much as possible the events which just mimic the signal - background events.

4.1 Analysis Objects Identification

First, we will describe reconstruction of particle tracks in detectors. They will be used for an identification of physical objects (electrons, muons).

Later, we will describe electrons, muons, jets and missing transverse energy $\not{E}_T$.

We will extensively use the names of the detectors in the following sections. For reference, see Sec. 3.2.

4.1.1 Tracks

Using tracking detectors (COT and SVX) allow us to reconstruct trajectories of particles.

There are a few algorithms which reconstruct the tracks depending on which detector the information is used from. In our analysis, we don't necessary require that event contains information from SVX. However, the hits from SVX detector are used whenever available. Therefore, the main algorithms for track reconstruction will use COT information. However, COT is available only for $|\eta| < 1$ and therefore there will be needed other track algorithms for plug region. We will very briefly describe the algorithms available at CDF.

First, we will describe the parameters used for track description at CDF.

Track Parametrization

In homogeneous magnetic field, tracks of charged particles travel on a helix with the axis of the helix parallel to the magnetic field. At CDF, we use the following five parameters to describe the helix of a charged track

z_0 : z position of helix at point of minimum approach to origin

d_0 : signed impact parameter: distance between helix and origin at minimum approach

C : half curvature (same sign as the charge of the particle)

ϕ_0 : direction of track at point of minimum approach

$\cot\theta$: cotangent of the polar angle at minimum approach

The radius of the circle is then defined as $\rho = \frac{1}{|2C|} = \frac{1}{2QC}$, where Q is charge of the particle. The p_T of the track can be then obtained via the following equation known from classical electrodynamics $p_T = B \cdot Q \cdot \rho$.

Track Reconstruction Algorithms

There are a few different track algorithms according detectors informations used, available at CDF:

- Stand-alone algorithm: it use only SVX information for reconstruction.
- Outside-In algorithm: use as the input the COT tracks and extrapolates it to SVX detector. It goes from *outside* to *inside* relatively to interaction point.
- Inside-Out algorithm: it starts from SVX and add the additional COT hits. This way it makes possible tracking even outside $|\eta| < 1$ region covered by COT.
- Phoenix algorithm: it is 2D calorimetry-seeded tracking algorithm.

For our selection of events, we will use Outside-In algorithm for central leptons and Phoenix algorithm for plug electrons.

The Outside-In (OI) algorithm starts with information from COT detector. The COT hardware provides, as its output, the drift time for each detected hit on COT wire. This time measurement is converted into a position measurement. Once the positions of hit candidates are known, the algorithm scans each of the eight COT super-layers looking for line segments. These segments are used in subsequent steps. Segment finding begins by looking for triplets of hits belonging to consecutive layers. A list of candidate segments is formed by selecting cases in which the central hit lies close enough to the midpoint of the external ones, and the slope of the segment with respect to the radial direction is not too high (so that the tracks are sufficient high momentum tracks). The candidate segments are then fitted to a straight line which is then extrapolated to the other layers of the super-layer. Once segments are available, they are assembled into tracks. The strategy for doing this is divided into two parts: at first, axial segment are joint in a 2D track; then, stereo segment and individual hits are attached to each axial track.

When the hits in SVX are available, they are attached to the original COT track and the parameters of the track are refitted.

In Phoenix algorithm, as an input is used an information from the calorimeter. We know that two points and a curvature define a unique helix. Every calorimeter cluster assumed to result from the passage of an electron from primary vertex provides:

- one point from the primary vertex position in three dimensions,
- one point from the shower-max (CES,PES) hit position in three dimensions,
- a measurement of the transverse energy and thus the absolute value of the curvature.

There is enough information to define helix up to ambiguity in the sign of the curvature. Two possible solutions differ in curvature C and the polar angle ϕ_0.

The information from primary vertex and the calorimeter is used to create tracks. These tracks are then used to drive the silicon pattern recognition just as for COT tracks (it require ≥ 3 $r - \phi$ hits and no stereo or z SVX hits). In this way, the informations from calorimeter and the SVX are incorporated to create tracks with the best possible parameters. Another solution would be to start from Stand-alone tracks and match them to EM objects in the plug. However, seeding standalone tracks at very high η where are a few layers of SVX would be very difficult.

4.1.2 Electrons

Electrons[1] are light charged leptons (don't interact through strong interaction).

For all the charged particles, there exists critical energy E_c above which they loose energy mainly by bremsstrahlung (emitting photons) when passing through the material. Below the critical energy dominate the energy losses by ionization.

For electrons, the critical energy is given approximately as $E_c = 800\ \text{MeV}/(Z + 1.2)$ [21], where Z is atomic number of absorber. This is order of tens of MeV for usual materials, e.g. in CDF central electromagnetic calorimeter is used lead (Z=82), for which the critical energy is $E_c \approx 10$ MeV. In the experiment, we are roughly three orders of magnitude above this energy, because we require the electrons to have energy above 20 GeV as we will see later in the chapter. Therefore, the electrons we are interested in loose their energy mainly by bremsstrahlung. The photons radiated by such electrons have still high energy, so they predominantly produce electron-positron pairs.

So, the high energy electron initiates electromagnetic (EM) shower in absorber as bremsstrahlung and e^+e^- pair production produce more and more electrons and photons with lower energy. Energy of the electron eventually fall below critical energy. Ionization dissipate rest of the electron energy. Even when the electron produce the EM shower, it is focused in small angle and usually almost all of the energy is deposited in one calorimeter tower (as opposite to jets).

The procedure of electron reconstruction and all the variables used to identify the electrons will be described later in the section.

Electron reconstruction

The electron reconstruction is as follows. Going through all calorimeter towers, there are found electron seeds (the tower need to pass criteria for E_T, default is $E_T \geq 2$ GeV). Going through all seeds, the EM clusters are formed by summing energies deposited in electromagnetic and hadron (HAD) calorimeter from the seed and the neighboring towers. For CEM, adjacent tower can be only 2 neighbor towers, one on each side, in η direction. For PEM, the 3×3 clusters can be formed at maximum. The ratio of deposited energy in HAD (E_{had}) and EM (E_{em}) calorimeters is calculated and the cluster is checked to pass certain acceptance criteria (by default, EM cluster must have $E_T > 2$ GeV and pass the criteria $E_{had}/E_{em} < 0.125$). After the EM cluster is formed, the tracks are associated to the cluster. The tracks are extrapolated to

---

[1] through this chapter we will not make difference between particle and its anti-particle and will call both the same way according the particle's name.

the plane of CES detector for the wedge containing the associated EM cluster. The extrapolated track is required to be within a prescribed distance of the CES wedge ϕ center and the EM seed tower center in z (the actual values will be presented in next section). Moreover, there are performed electron associations of clusters in CES and/or pre-shower detectors (CPR).

By the way, photons are reconstructed in a similar way. However, photon is not charged particle and therefore it leaves no hits in track detectors. Therefore, the one main difference in reconstruction of photons comparing to electrons is that with photon there is not associated corresponding track.

Electron identification variables

In the following, we describe the variables which are used for identification of the electrons:

- E_T:
  The transverse electromagnetic energy deposited by the electron in the CEM is calculated as the electromagnetic cluster energy multiplied by $\sin\theta$, where θ is the polar angle provided by the best track pointing to the EM cluster.

- P_T:
  The transverse momentum of the track associated with the electron. The track is fitted using COT hits and the constraints coming from the beam position. We don't necessary require that event contains information from SVX. However, the hits from SVX detector are used whenever available.

- E_{had}/E_{em}:
  The ratio of the hadronic calorimeter energy to the electromagnetic calorimeter energy for a cluster. This variable is designed to distinguish electrons (which should have very little hadronic energy) from jets (which should have mostly hadronic energy, but can have some EM energy as well).

- E/P:
  The ratio of the EM cluster transverse energy to the COT track transverse momentum. For our highly relativistic electrons ($E > 20$ GeV), we expect $E \approx P$. However, due to radiation of soft photons, this variable can have large values (the energy of electron is usually not distorted because the photon energy is deposited in the same EM cluster while the momentum of electron is lowered by radiation).

- $\Delta x_{CES} * Q$:
  The distance in the r-ϕ plane between the extrapolated, COT beam constrained track and the best matching CES cluster, times the charge Q of the track.

- Δz_{CES}:
  The distance in the r-z plane between the extrapolated, COT beam constrained track and the best matching CES cluster.

- χ^2_{strip}:
  The χ^2 comparison of the CES shower profile in r-z view with the same profile extracted from test beam electrons.

- *Lshr*:
  The lateral shower profile for electrons. The purpose of this quantity (similarly to the above χ^2_{strip} variable) is to provide some discrimination of electrons from hadronic showers faking electrons in the central electromagnetic calorimeter by comparing the observed sharing of energy deposition between towers to that expected for a "true" electromagnetic shower. The "true" electromagnetic shower profile was obtained by test beam electrons.

  Lshr is just a difference between expected and measured energy in adjacent towers divided by its uncertainty and multiplied by some constant.

- z_0:
  The z intersection of the track associated to the electron with the beam axis.

- COT track quality cuts:
  The electron associated track must have passed through certain number of axial (*TrkAxSeg*) and stereo (*TrkStSeg*) COT super-layers, with certain number of hits in each super-layer (the actual values will be listed in section 4.2.2).

- Isolation (Isol):
  The energy E_T in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$ around the electron cluster excluding the electron cluster divided by the energy in the electron cluster:

  $$Isol = \frac{E_T^{cone} - E_T^{electron}}{E_T^{electron}} \tag{4.1}$$

  This variable is corrected for calorimeter leakage across the ϕ boundary between wedges. The particular value of this variable just says what is the fraction of additional energy (in the cone around the lepton) comparing to energy of lepton (in case of cut 0.1, it means 10% of additional energy in the cone). It has high values for jets (lots of particles close to each other) and small values for electrons.

- Fiduciality:
  This is boolean variable. It checks for the position of CES cluster corresponding to the electron. It is required that $|X_{CES}| < 21$ cm and $9 < |Z_{CES}| < 230$ cm, where X_{CES} and Z_CES are positions of CES cluster in transverse and r-z plane, respectively. Using such cuts, the CEM tower with highest $|\eta|$ is excluded and most of tower next to chimney is included.

All the rest of the variables are used just for identification of the plug electrons:

- Pes2dEta:
  this is the detector η (i.e. η with respect to the nominal origin (0,0,0)) of the best matching 2D PES cluster

- Pem3x3FitTow:
  number of towers used by the 3×3 PEM cluster fit. The fit compares the EM shower profile of a given PEM cluster to shapes derived from plug test beam data.

- χ^2_{Pem3x3}:
  this is the χ^2 value of the 3×3 PEM cluster fit.

- Pes2d5by9U, Pes2d5by9V:
  These are just ratios:

$$Pes2d5by9U(V) = \frac{\text{sum of the energy in the central 5 strips of PES cluster}}{\text{sum of the energy in all 9 strips of PES cluster}} \tag{4.2}$$

  It is calculated for U (Pes2d5by9U) and V (Pes2d5by9V) PES strips of the best matching associated 2D PES cluster. It is essentially an isolation variable for PES cluster.

- $\Delta R(PEM, PES)$:
  The matching distance (in the transverse plane) between the Pem3x3Fit coordinates and the 2D PES coordinates of electron PEM cluster.

- PhxMatch:
  variable which tells whether or not the Phoenix (PHX) track match to the PES cluster.

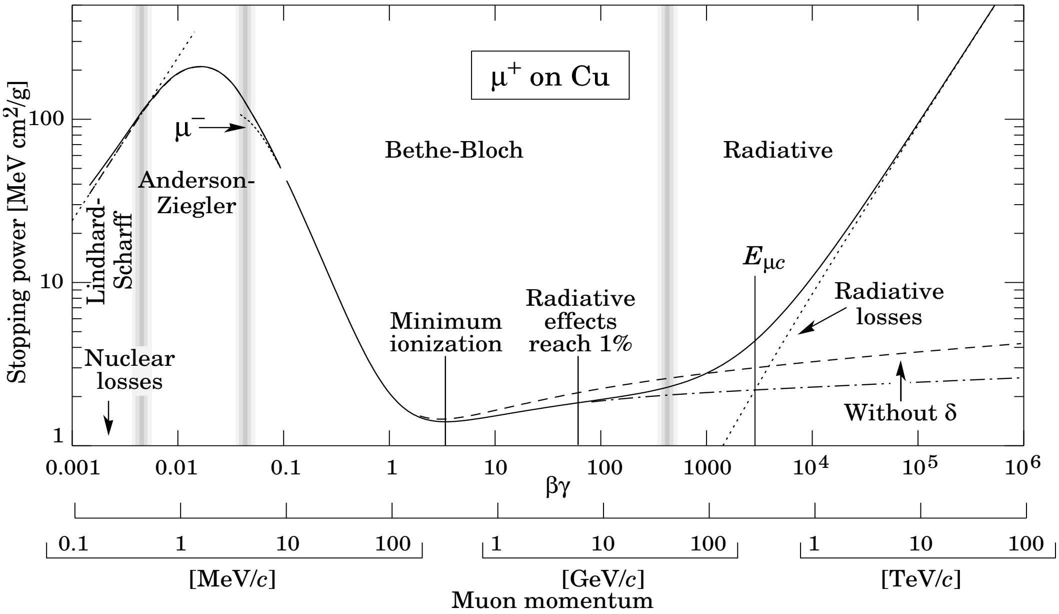


Figure 4.1: Stopping power ($=< -dE/dx >$) for positive muons in cooper as a function of $\beta\gamma = p/Mc$ and momentum p [21].

- N^{SVX}_{hits}:
  number of hits in SVX detector which belong to PHX track

4.1.3 Muons

Muons are just heavier "brothers" of electrons, i.e. they are charged leptons which are $\sim$ 200 times heavier than electrons. Therefore, they are much less decelerated (acceleration $a \sim 1/\text{Mass}$) in electromagnetic field when passing through detector material and consequently they emit much less photons and loose much less energy by bremsstrahlung than electrons (radiation looses $\sim a^2 \sim 1/m^2$). Therefore, the critical energy for the muons occurs at several hundred GeV depending on material which they pass (for iron it is around 350 GeV).

For the practical energies of muons in $t\bar{t}$ events at Tevatron (order of tens GeV) the ionization looses dominates. It can be seen from Fig. 4.1 that in wide range from hundreds of MeV to hundreds of GeV the muon mean energy lose rate is almost constant and close to the minimum. Particles with such energies are said to be minimum ionizing particles (mip's).

Therefore, muons loose only very small portion of its energy in calorimeters and leave them. They must be identified as well as they energy need to be measured in other

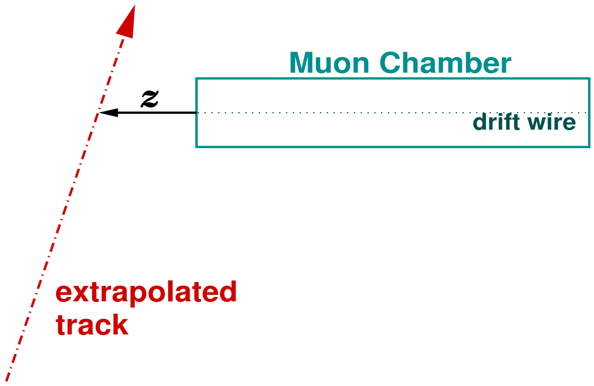


Figure 4.2: An illustration of the definition of the z-fiducial distance.

detectors. For this reason, there are another detectors behind calorimeters specifically placed in order to register muons. As the muons are charged particles, their momentum will be measured in tracking detectors using the curvature of the tracks associated with them.

Muon reconstruction

First, we need to define "fiducial" distance which will be used later in reconstruction.

The $z-$fiducial distance (see Fig. 4.2) is defined to be the distance between extrapolated track position and the edge of the muon chamber, in the direction parallel to the drift wire. The x-fiducial distance is defined an analogous way, except the distance is perpendicular to the drift wire[2]. The convention is that for tracks extrapolated outside the chamber the fiducial distance is greater than 0. For tracks extrapolated inside the chamber the fiducial distance is less than zero.

In reconstruction, the finder examines all pairs of hits taken from layers 0 and 2 that are within a specified fiducial distance of each other in x. These two hits are used as the basis for a line segment. The line parameters are used to estimate the location of hits in layer 1 and 3. The finder then looks for hits in those layers that are within another specified distance (typical distances are order of centimeters, as we will see later) from the prediction in regional x. There needs to be at least one hit in these additional layers to form at least three-hit stub. This entire procedure is repeated, using layers 1 and 3 as the starting layers, and then searching for hits in the layers 0 and 2. All hits in each stub are then fit by linear fit.

---

[2] this means that this "regional" distance in x is not necessary also distance along x axis defined globally for CDF detector,see Sec. 3.2.1

The linking of stubs and tracks (found in COT or SVX) is performed in order to create final muon objects. This is done by extrapolating the tracks to the stubs through the inhomogeneous magnetic field and dense material. Of course, more than one track may make a reasonable match with a stub. Moreover, muon might travel through more than one muon detector subsystem, so a given track may have more than one stub attached to it. Tracks must pass certain criteria in order to be considered. Removed are tracks with low momentum, large impact parameter, large z at the distance of closest approach to the beamline, or few axial COT hits. Every muon is guaranteed to be associated with the track. It may or may not have a stub in any given muon subsystem. Such "stubless muons" do not extrapolate to the fiducial volume of any muon detection system.

Moreover, the tracks associated to the muons are extrapolated to the calorimeters to determine which electromagnetic and hadronic towers have been traversed by the muon. The energies in those towers are summed to determine the electromagnetic and hadronic energy depositions of the muon and must correspond to minimum-ionizing particle.

Muon identification variables

The variables used for identification of muons are:

- P_T:
  The P_T of the best matched track associated to the muon.

- E_{had}, E_{em}:
  The hadronic (E_{had}) and electromagnetic (E_{em}) calorimeter energy associated to the muon.

- COT track quality cuts:
  These quality cuts are the same as for electron's track, see Sec. 4.1.2.

- d_0:
  The impact parameter of the associated track which is corrected for beam position. Different cuts on d_0 are used depending whether the track has SVX hits (N_{hits}^{SVX}) or not.

- z_0:
  The z coordinate of the track at the distance of closest approach to the beamline.

- Δx_{CMU}, Δx_{CMP}, Δx_{CMX}:
  The distance in regional x between extrapolated track and the muon stub in CMU, CMUP, CMX detector, respectively. Note, these cuts are applied only where appropriate, e.g. cut on Δx_{CMX} will be applied only for muon passing through CMX detector, while for muon passing both CMU and CMP detectors, the cuts on Δx_{CMU} and Δx_{CMP} will be applied.

- ρ:
  COT exit radius of the track. This variable is used only for CMX muons. This cut is used for the following reason. The Level 1 CMX18 trigger requires that a track has hits in all 4 axial super-layers of the COT. However, it is possible for high-p_T CMX muon not to traverse all 4 axial COT super-layers. In order the data and simulation agree in ρ, it is recommended to cut on value $\rho > 140$ cm of the track.

- $Isol\_P_T$:
  it is the ratio of sum of electromagnetic and hadronic energies around the muon track in the cone of $R < 0.4$ and the muon P_T.

4.1.4 Jets

Jets are showers of particles flying in roughly the same direction and coming from hadronization of original quarks or gluons. The particles which form jets are mostly hadrons (pions, kaons, etc.) but also leptons can be produced there. Therefore, jets loose the energy mostly by ionization and atomic excitation when passing through the material. Because there are many particles (orders of tens) within the jets which are depositing energy within the same calorimeter towers, there is no practical way of identifying and measuring each particle separately. Therefore, the jets are reconstructed as whole objects within calorimeters.

Since most of these particles are charged, they leave also the hits in tracking detectors.

Jet reconstruction

There exists more algorithms for jet reconstruction within the CDF (cone algorithm, cone algorithm using midpoints, K_T algorithm). The default one is the cone algorithm and we use it in our analysis, so we will describe here mainly that particular algorithm. At the end, we will very briefly describe midpoint and K_T algorithm.

Cone algorithm form jets by associating together particles whose trajectories (i.e. centers of towers) lie within a circle of specific radius R in $\eta \times \phi$ plane. This 2-dimensional space is natural in $p\bar{p}$ collisions where the dynamics are spread out in longitudinal direction.

Starting with a trial geometric center for a cone in $\eta \times \phi$ space, the energy weighted centroid is calculated including contributions from all particles within the cone. This new point in $\eta \times \phi$ is then used as the center for a new trial cone. As this calculation is iterated the cone center "flows" until a "stable" solution is found, i.e. until the centroid of the energy depositions within the cone is aligned with the geometric axis of the cone. For a specified geometric center for the cone (η^C, ϕ^C) the particles i within the cone satisfy

$$i \in C \quad : \quad \sqrt{(\eta^i - \eta^C)^2 + (\phi^i - \phi^C)^2} \leq R \tag{4.3}$$

Naively, we can simply identify these stable cones, and the particles inside, as jets. In originally proposed *Snowmass* scheme [60] [3] (which is actually used at CDF in cone algorithm), the variables of the jet are defined as follows:

$$\eta^{Jet} = \eta^C = \frac{\sum_{i \in C} E_T^i \eta^i}{E_T^C} \tag{4.4}$$

$$\phi^{Jet} = \phi^C = \frac{\sum_{i \in C} E_T^i \phi^i}{E_T^C} \tag{4.5}$$

$$E_T^C = \sum_{i \in C} E_T^i \tag{4.6}$$

where the sums are on towers in the cone "C".

To keep the time for data analysis within reasonable bounds the concept of the seed was introduced. Instead of looking "everywhere" for stable cones, the iteration process starts only at the centers of seed towers that passed a minimum energy cut (by default, 1 GeV). Another modification of cone algorithm with the seeds is 'midpoint' algorithm. This algorithm adds a starting point for clustering at the positions given by $p_i + p_j$,$p_i + p_j + p_k$, etc., where p_i is momentum corresponding to i-th seed. This has advantage that the sensitivity of the algorithm to soft radiation is essentially removed (this is one of main theoretical concerns about cone algorithm).

Moreover, there need to be addressed the question of treating overlapping stable cones. It is quite common for two stable cones to share some subset of their particles. While not all particles in the final state need to be assigned to a jet, particles should not be assigned to more than one jet. Typically, cones whose shared energy is larger

---

[3] there are also other schemes, like $E-scheme$ where the summing over particles is performed for four-momenta components

than a fixed fraction (e.g. $f = 50\%$) of the energy in the lower energy cone are merged. For the cases with shared energy below this cut, the shared particles are assigned to the cone that is closer in $\eta \times \phi$ plane.

There are three sizes of the cone radius R which are used at CDF: 0.4, 0.7, 1.0. The default cone which was also used in this analysis is 0.4. However, one may consider to use 0.7 cone in analysis involving b jets as they are usually wider than jets coming from hadronization of other light quarks.

There is another family of algorithms for reconstruction of jets, called K_T algorithms. Since we will not use this algorithm, we will mention just its idea. The K_T algorithm starts with a list of *preclusters* which are formed from calorimeter cells, particles, or partons. When the precluster is well separated from others, it is defined as a jet. Otherwise, it is merged with other precluster which is closest to it. Fig. 4.3 illustrates how the K_T algorithm successively merges the preclusters in a simplified diagram.

Jet Corrections

The energy deposited in calorimeter which is assigned by the algorithm to the jet (so called "raw jet energy") must be corrected for detector effects as well as physics effects. These corrections scale the "raw jet energy" (calorimeter level) to the particle level and eventually to the parton level. Therefore we are talking about jet energy scale (JES).

The corrections to the "raw jet energy" are due to the following:

- *η dependent calibration* (f_η) - also called "relative jet energy correction". It is applied to raw jet energies measured in the calorimeter to make jet energy uniform along pseudo-rapidity η. This correction is used because the different calorimeters (even different parts of the same calorimeter) have different response to the same energy of incident particles. The response of the central calorimeters at CDF, covering the region $0.2 < |\eta| < 0.6$, are best understood and the selected region is far away from the cracks. Therefore the response of the forward region calorimeters are corrected to be the same as that of the central ones.

- *multiply interactions* ($Mp\bar{p}I$) - In average, 2.3 $p - \bar{p}$ interactions per bunch crossing happen at CDF. In such case, additional energy can appear inside the jet clusters, increasing the energy of the measured jet. This correction subtracts this contribution in average according number of primary vertices in the event.

- *absolute jet energy scale* (f_{jes}) - Corrects the jet energy measured in the calorimeter for any non-linearity and energy loss in the un-instrumented regions of each

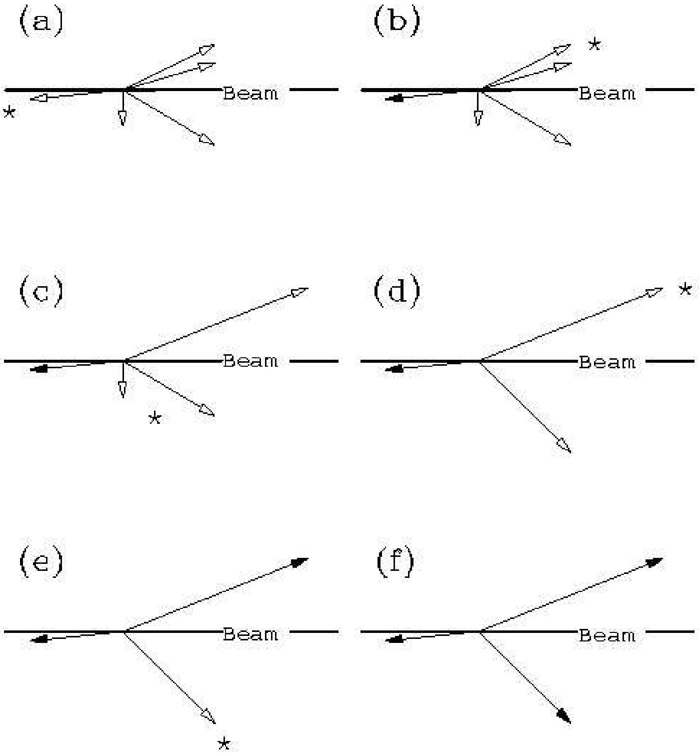


Figure 4.3: A simplified example of the final state of hadron collision. The open arrows represent preclusters in the event, and the solid arrows represent the final jets reconstructed by K_T algorithm. The six diagrams show successive iterations of the algorithm. In each diagram, either a jet is defined (when it is well separated from all other preclusters), or the preclusters are merged. The asterisk labels the relevant precluster(s) at each step [61].

calorimeter. The jet energy is corrected to the sum of the particles p_T within the cone of same size as jet cone and around the jet direction. This conversion factor from calorimeter energy to particle-level energy is called absolute energy correction.

- *underlying event* (UE) - The underlying event energy is defined as the energy associated with the spectator partons in a hard collision event. This energy can fall inside the jet clustering cone and therefore must be subtracted from the particle-level jet energy.
- *out of cone energy* (OOC) - The jet clustering may not include all the energy from the initiating parton. Some of the partons generated during fragmentation may fall outside the cone chosen for clustering algorithm. Out-of-cone corrections

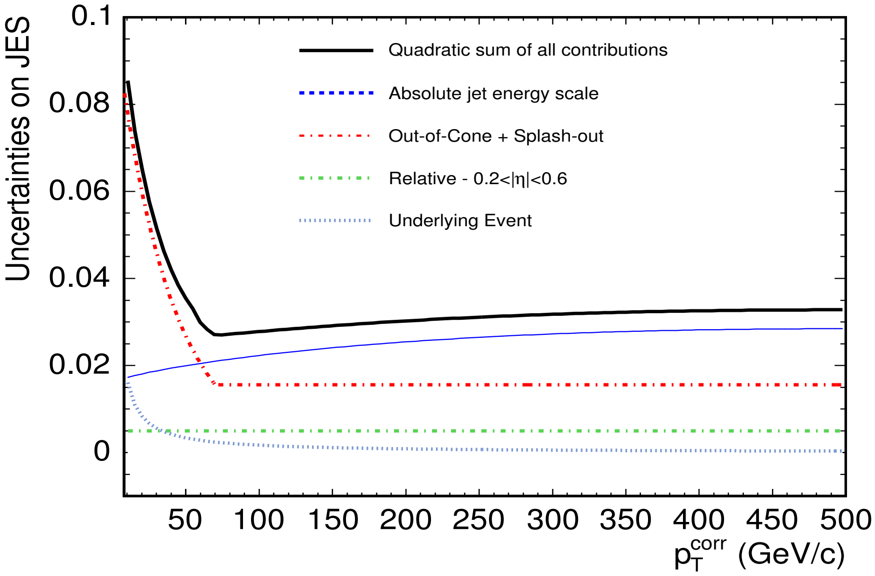


Figure 4.4: Systematic uncertainties on jet energy scale as a function of the p_T in region $0.2 < |\eta| < 0.6$ [62].

were applied in order to correct particle-level jet energy to the parton energy.

These levels of corrections are applied to the raw (measured) jet energy according to:

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

where R is the clustering cone radius, E_T is the raw jet energy measured in the cone, and η is the detector pseudo-rapidity of the jet (pseudo-rapidity calculated with respect to the origin of coordination system, not to the position of vertex). The total systematic uncertainties on jet energy scale as a function of jet p_T are shown in Fig. 4.4. These uncertainties will be later source of the systematic uncertainty on our top quark mass measurement (as we will see later in Chap. 9, it will be the largest source of systematic uncertainty).

It should be mentioned that not all the corrections are applied in all analyses. For example, for measurement of cross-section, the underlying event energy and out of cone energy corrections are not applied because it is enough to reconstruct the energy of the jet and it is not needed to reconstruct the energy of partons.

4.1.5 Missing transverse energy

Neutrinos are neutral leptons which mass seems to be very small. Therefore, when passing through the material, the neutrinos interact only via weak interaction and their probability of interaction with the detector is very small (they can pass whole Earth without interaction). Therefore, we are practically not able to directly measure any kinematic characteristic of neutrino.

However, we can deduce some of its characteristics due to momentum conservation law. The incident protons and anti-protons have zero transverse momentum. Consequently it is assumed, the incident partons have presumably also zero (or negligible) transverse momentum. However, we can not use the same assumption about longitudinal partons momentum, because incident protons (anti-protons) have very large longitudinal momentum (~ 1 TeV) and according PDFs the parton momentum can be anywhere from zero up to total momentum of incident hadron.

Therefore, the final state particles must also have all together zero transverse momentum. Summing up all the reconstructed particles momenta, we obtain imbalance of momentum in transverse plane and we associate the missing transverse momentum with the final state neutrino momentum. However, the $t\bar{t}$ dilepton events which we are interested in, have two neutrinos in final state and so we know just sum of neutrinos momenta, not each momentum individually.

Note, even this variable is called missing transverse energy ($\not{E}_T$)[4], it is not scalar, rather vector variable which is given by its size $\not{E}_T$ and azimuthal angle ϕ (or by its two components $\not{E}_{T_x}$ and $\not{E}_{T_y}$).

$\not{E}_T$ reconstruction

$\not{E}_T$ of the event is defined as the transverse energy needed in order the event to have zero total transverse energy.

First, there is summed all the energy deposited in the calorimeters (electromagnetic or hadronic) where only calorimeter towers with the deposited energy above certain threshold (by default, 0.1 GeV) are considered. This is performed in vector sense and with respect to origin of the coordination system (point $p_0 = (0, 0, 0)$):

$$\vec{\not{E}}_{T_{raw}} = -\sum_i E_{T_i} * \vec{n_i}, \tag{4.8}$$

where E_{T_i} is the magnitude of the transverse energy contained in each calorimeter

---

[4] it should be called transverse momentum. However, because its definition use measurements from calorimeter (see below), it is called energy

tower i and $\vec{n_i}$ is unit vector in the azimuthal plane that points from the beam lie to the i-th calorimeter tower.

Missing transverse energy must have opposite direction (therefore minus sign) as the sum of all transverse energies in order the event to have zero total transverse energy as per definition.

Of course, such definition of $\not{E}_T$ must be corrected for various reasons. First, the interaction does not necessary occurs at point p_0, so the vector sum $\vec{\not{E}}_{T_{raw}}$ must be corrected for the position of the primary vertex to $\vec{\not{E}}_{T_{vertex}}$, where vectors $\vec{n_i}$ will be taken with respect to the primary vertex position and not to origin of coordination system p_0.

Moreover, not all particles leave all the energy in the calorimeters. Muons pass through calorimeters as minimum ionizing particles leaving only small portion of its energy in the calorimeters. Therefore, there must be made the correction on muons. This is done in very simple way, by removing the energy of calorimeter towers corresponding to muon (by adding this energy to the $\vec{\not{E}}_{T_{vertex}}$) and adding the muon momentum P_T measured in tracking detectors (by subtracting this momentum from the $\vec{\not{E}}_{T_{vertex}}$):

$$\vec{\not{E}}_{T_\mu} = -(\sum_i E_{T_i} * \vec{n_i} - \vec{E}_{T_\mu} + \vec{P}_{T_\mu}) = \vec{\not{E}}_{T_{vertex}} + \vec{E}_{T_\mu} - \vec{P}_{T_\mu} \tag{4.9}$$

The final correction of the $\not{E}_T$ comes from jets. We know from Sec. 4.1.4 that jet energies must be corrected in order to go from calorimeter level up to hadron level energy. However, in calculation of $\not{E}_T$ we just used the "raw energy" of the jets which was deposited in the calorimeter. In order to have in the $\not{E}_T$ included more appropriate jet energies, we need to add to $\not{E}_T$, energies of all the towers belonging to jets and after that subtract the corrected jet energies:

$$\vec{\not{E}}_{T_{jets}} = \vec{\not{E}}_{T_\mu} - \vec{E}^{raw}_{T_{jets}} + \vec{E}^{corr}_{T_{jets}} \tag{4.10}$$

4.2 Selection criteria

In order to understand the philosophy of $t\bar{t}$ dilepton events selection criteria, we will first describe the signature of $t\bar{t}$ signal and background events, i.e. what are the final products of these interactions.

The signature of $t\bar{t}$ dilepton decay channel (see section 2.2.2) is two high-energetic leptons (e or μ) and high missing energy $\not{E}_T$ (due to two unobserved neutrinos) from the W decays and two high-energetic jets coming from hadronization of the b quarks.

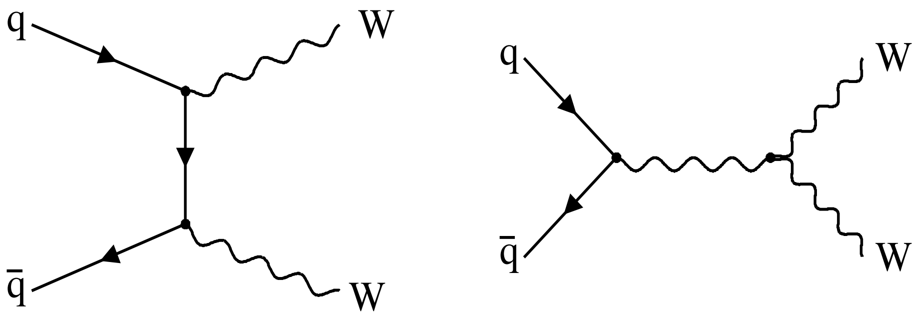


Figure 4.5: Feynman diagrams for tree level WW production in the Standard Model. Analogous diagrams exist also for WZ production.

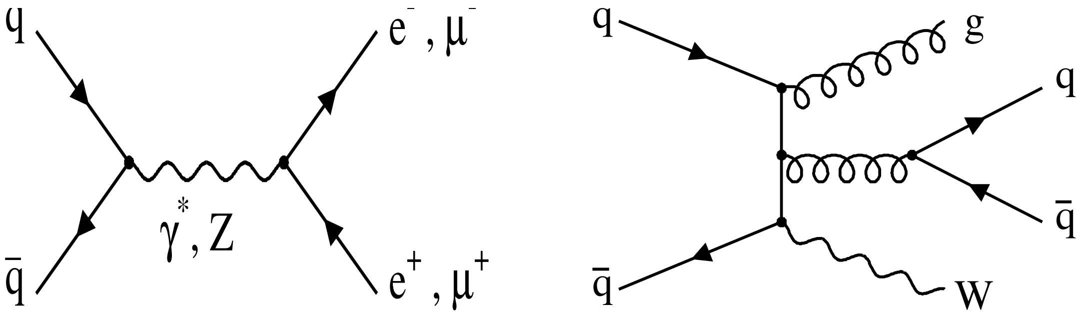


Figure 4.6: Lowest order diagram for Drell-Yan production of lepton pairs at hadron colliders (left) and example of W boson production in association with 3 jets (right) which is main source of “FAKE” events.

These particles are highly energetic because they all come from the decay of high mass top quark. There is possibility of extra jets due to hard gluon radiation from incoming parton or from a final state quark or gluon, so called initial and final state radiation (ISR and FSR), see Fig. 2.4.

The major backgrounds for dilepton $t\bar{t}$ events are diboson production, Drell-Yan events and “Fake” events (see Fig. 4.5 and Fig. 4.6) with the signatures described below:

- Diboson production : this includes WW ($q\bar{q} \rightarrow WW$) events in which both W decays leptonically (there must be also present at least 2 jets from ISR,in order to have the same event signature as $t\bar{t}$ dilepton events) and WZ events ($q\bar{q} \rightarrow WZ$), in which the Z decays leptonically and the W hadronically (in such WZ events, the mis-measurement of jets or leptons must occur to have $\not{E}_T$, because it does not have the neutrinos in final state). We don’t consider the ZZ events as their contribution is negligible.

- Drell-Yan production ($q\bar{q} \rightarrow Z/\gamma^* \rightarrow \ell^+\ell^-, l \in e, \mu, \tau$): this is the production of Z boson or virtual photon decaying into charged leptons. There must be at least 2 jets from ISR. Moreover, this process does not have real physics $\not\!\!E_T$, it can have $\not\!\!E_T$ just due to mis-measurement of jets and/or leptons energies.

- "Fakes" production ($W \rightarrow \ell\nu + jets$): events where one jet is misleadingly reconstructed as a lepton, i.e. jet "fakes" the signature of a lepton. The main source of such events is production of W boson with associated production of at least 3 jets ($W+ \geq 3$ jets).

4.2.1 Trigger requirements

There are three primary data sets used in this analysis. Two of them are based on central electron (CEM18) or muon triggers (CMUP18, CMX18), respectively. The third dataset is based on the trigger requiring plug electron and $\not\!\!E_T$ (METPEM).

The Central Electron Trigger (CEM18)

This trigger looks for electron candidate in central EM detector by requiring the following on different trigger levels:

- Level 1: requires a central EM tower with $E_T^{EM} \geq 8$ GeV. and $E_{HAD}/E_{EM} \leq 0.125$. The CEM tower must have a matched XFT track with $P_T \geq 8$ GeV.

- Level 2: requires a central EM cluster with $E_T^{EM} \geq 16$ GeV and $E_{HAD}/E_{EM} \leq 0.125$. The EM cluster must have a matched XFT track with $P_T \geq 8$ GeV.

- Level 3: requires a central EM cluster with $E_T^{EM} \geq 18$ GeV and $E_{HAD}/E_{EM} \leq 0.125$. The seed tower of the EM cluster must be matched to a fully reconstructed three dimensional (3D) track with $P_T \geq 9$ GeV.

The Central muon trigger (CMUP18)

This central muon trigger utilizes information from both central muon detectors (CMU and CMP):

- Level 1: requires an XFT track with $P_T \geq 4$ matched to a muon track segment with $P_T > 6$ GeV from CMU and CMP

- Level 2: no further requirements

- Level 3: requires muon track with $P_T \geq 18$ GeV and matched to the CMU stub within $|\Delta x| < 10$cm and within $|\Delta x| < 20$cm of the CMP stub.

The CMX muon trigger (CMX18)

This is another central muon trigger, which:

- Level 1: requires an XFT track with $P_T \geq 8$ matched to a muon track segment with $P_T > 6$ GeV from CMX
- Level 2: no further requirements
- Level 3: COT track with $P_T \geq 18$ GeV matched to the CMX stub to within 10 cm in $r - \phi$ plane ($|\Delta x| < 10$ cm).

The Forward Electron Trigger (METPEM)

The tracking coverage is not sufficient to allow for triggers based on calorimetry and tracking to be used in forward region. To compensate for this, the electron trigger in the forward region looks for $\not{E}_T$ in addition to an EM calorimeter object:

- Level 1: forward EM tower with $E_T^{EM} \geq 10$ GeV and $E_{HAD}/E_{EM} \leq 0.125$ (this applies only for $E_T^{EM} < 14$ GeV). The $\not{E}_T$ must be more than 15 GeV where no correction on vertex position is made.
- Level 2: forward EM cluster with $E_T^{EM} \geq 20$ GeV and $E_{HAD}/E_{EM} \leq 0.125$. It is required isolation of the E_T cluster to be $< 10\%$.
- Level 3: forward EM cluster with $E_T^{EM} \geq 20$ GeV and $E_{HAD}/E_{EM} \leq 0.075$ and $\not{E}_T$ must be more than 15 GeV ($z_{vertex} = 0$ assumed). It is also required isolation of the E_T cluster to be $< 10\%$.

4.2.2 $t\bar{t}$ Dilepton Selection

For our top mass analysis, we employ set of selection criteria developed for the top cross section measurement in the dilepton decay channel [63], referred as "dilepton" (DIL) analysis, which demand a "tight" lepton in combination with a "loose" lepton of opposite charge. This is the method similar to that used in the CDF Run I measurement [64], and requires both candidate leptons to be specifically identified as either electron or muon as opposite to other, so called "LTRK" method which requires only one well identified lepton and high-P_T track.

The first step in the selection is to classify any lepton in the event as tight or loose, following the cuts listed in Tab. 4.1, 4.2, 4.3, 4.4. For description of all variables listed in tables, see Sec. 4.1.2, 4.1.3.

variable	CEM cut value
E_T	> 20 GeV
P_T	> 10 GeV
E/P	< 2.0, unless $P_T > 50$ GeV
$\|z_0\|$	< 60 cm
E_{had}/E_{em}	$< 0.055 + 0.00045 \times E/GeV$
Isol	< 0.1
Lshr	< 0.2
$\Delta x_{CES} * Q$	-3.0 cm $< \Delta x_{CES} * Q <$ 1.5 cm
$\|\Delta z_{CES}\|$	< 3.0 cm
χ^2_{strip}	< 10
Fiduciality	1 (yes)
COT track quality cuts	≥ 3 TrkAxSeg AND ≥ 2 TrkStSeg with ≥ 5 hits/Seg

Table 4.1: Central electron (CEM) identification (ID) cuts.

variable	PHX cut value
E_T	> 20 GeV
E_{had}/E_{em}	< 0.05
Isol	< 0.1
Pes2dEta	$1.2 < \|\eta_{PES}\| < 2.0$
Pem3x3FitTow	$\neq 0$
χ^2_{Pem3x3}	< 10.0
Pes2d5by9U	< 0.65
Pes2d5by9V	< 0.65
PhxMatch	yes
N^{SVX}_{hits}	≥ 3
$\|z_0\|$	< 60.0 cm
$\Delta R(PES - PEM)$	< 3.0 cm

Table 4.2: Plug electron (PHX) ID cuts.

In summary, “tight” leptons must have well-measured tracks, shower profile which is typical for them, and must be isolated electrons or muons which can trigger the event. Tight electrons are either central (CEM[5]) or plug (PHX). Tight muons are pointing to the region covered by both central muon chambers and with a stub in both (CMUP), or to the region covered by the muon extension chamber and with a stub in it (CMX).

---

[5] the leptons are usually labeled according name of detector they hit. The exception is PHX electron, which is detected in plug calorimeter (PEM), but it has named according its matched Phoenix (PHX) track.

variable	cut value
P_T	> 20 GeV
E_{em}	$< 2.0 + Max(0, (P - 100) * 0.0115)$ GeV
E_{had}	$< 6.0 + Max(0, (P - 100) * 0.028)$ GeV
$\|z_0\|$	< 60 cm
$\|d_0\|$	< 0.02 cm (if $N_{hits}^{SVX} > 0$) OR < 0.2 cm (if $N_{hits}^{SVX} = 0$)
COT track quality cuts	≥ 3 TrkAxSeg AND ≥ 2 TrkStSeg with ≥ 5 hits/Seg
$Isol\_P_T$	< 0.1
Δx_{CMU}	< 3 cm
Δx_{CMP}	< 5 cm
Δx_{CMX}	< 6 cm
ρ	> 140 cm (only for CMX muons)

Table 4.3: Central muons ID cuts. "Max" function will choose the larger of the two values.

variable	cut value
P_T	> 20 GeV
E_{em}	$< 2.0 + Max(0, (P - 100) * 0.0115)$ GeV
E_{had}	$< 6.0 + Max(0, (P - 100) * 0.028)$ GeV
$E_{em} + E_{had}$	> 0.1 GeV
$\|z_0\|$	< 60 cm
$\|d_0\|$	< 0.02 cm (if $N_{hits}^{SVX} > 0$) OR < 0.2 cm (if $N_{hits}^{SVX} = 0$)
COT track quality cuts	≥ 3 TrkAxSeg AND ≥ 2 TrkStSeg with ≥ 5 hits/Seg
$Isol\_P_T$	< 0.1

Table 4.4: CMIO muons ID cuts. "Max" function will choose the larger of the two values. Note, there are no Δx stub matching cuts since we are only allowing non-fiducial CMIO muons.

"Loose" leptons are typically electrons or muons which pass the same identification cuts as the tight leptons, but fail the isolation requirement (Isol ≥ 0.1). Loose electrons can only be central (NICEM). Loose muons are not only the non isolated counterparts of the tight muons (NICMUP and NICMX), but also some special muon categories that cannot trigger the event. These are:

- muons with track pointing to the region covered only by one of the central muon chambers and with a stub only in that detector (CMU and CMP),
- CMX muons with COT exit radius ρ smaller than 140 cm (RHOCMX),

- stubless tracks with energy deposition typical for minimum ionizing particle pointing to regions not covered by any muon chamber(CMIO).

While for the stubbed muon categories we allow for the non isolated muons (NICMU, NICMP and NIRHOCMX), CMIO muon must always be isolated.

The DIL selection accepts events containing exactly two leptons. Tab. 4.5 list all of the dilepton categories considered in the DIL selection. For each category, it requires that at least one lepton be a tight trigger lepton, i.e. CEM or PHX or CMUP or CMX. The other lepton in the event can be tight or loose. However, we don't allow PHX-PHX category, because it contributes very little to signal ($t\bar{t}$ events are mostly central) and due to difficulties in estimating Drell-Yan background for events that come through trigger requiring $\not\!\!E_T$ (METPEM).

After establishing a dilepton selection, we proceed to apply some event topology cuts aiming at rejecting backgrounds from other Standard Model dilepton states, mainly Drell-Yan and diboson, while keeping as much as possible of the top dilepton signal. Jet corrections are applied before calculating any of the event topology variable. All corrections described in Sec. 4.1.4 are applied except for underlying event correction and out of cone correction. The event topology cuts are applied in the following order:

1. *cosmic and conversion removal*:

   This cut reject the events which were marked as the events with the particles coming from cosmic or the events where the electrons are coming from the conversion of the photon to e^+e^- pair within detector.

   The cosmic removal is used to remove events where the muons are identified as an cosmic muons.

   The conversion removal is applied only to central electrons. The track associated with the electron is checked to all other tracks whether it is close in all three spatial dimensions to another track and that it has opposite charge. In such case, the electron is flagged as the conversion electron.

2. $Z - veto\ cut$:

   Aimed to reject events with Z boson decaying into dileptons (dielectrons or dimuons). Therefore, this cut is not applied to events with different types of leptons (e and μ).

   Rejected are events when:

$$(76 \text{ GeV} < M_{\ell\ell} < 106 \text{ GeV}) \textbf{ AND } (JetSig \geq 8.0\sqrt{GeV} \textbf{ OR } dPhiJet \leq 10^\circ)$$

category	$\ell_1\ell_2$	trigger
ee	CEM – (NI)CEM	CEM18
	PHX – (NI)CEM	CEM18
$\mu\mu$	CMUP – (NI)CMUP	CMUP18
	CMUP – (NI)CMU	CMUP18
	CMUP – (NI)CMP	CMUP18
	CMUP – CMIO	CMUP18
	CMUP – (NI)RHOCMX	CMUP18
	CMUP – (NI)CMX **OR** CMX – (NI)CMUP	CMUP18 **OR** CMX18
	CMX – (NI)CMX	CMX18
	CMX – (NI)CMU	CMX18
	CMX – (NI)CMP	CMX18
	CMX – CMIO	CMX18
	CMX – (NI)RHOCMX	CMX18
$e\mu$	CEM – (NI)CMUP **OR** CMUP – (NI)CEM	CEM18 **OR** CMUP18
	CEM – (NI)CMX **OR** CMX – (NI)CEM	CEM18**OR** CMX18
	CEM – (NI)CMU	CEM18
	CEM – (NI)CMP	CEM18
	CEM – CMIO	CEM18
	CEM – (NI)RHOCMX	CEM18
	PHX – (NI)CMUP	CMUP18
	PHX – (NI)CMX	CMX18
	PHX – (NI)CMU	METPEM
	PHX – (NI)CMP	METPEM
	PHX – CMIO	METPEM
	PHX – (NI)RHOCMX	METPEM

Table 4.5: Dilepton categories considered in DIL $t\bar{t}$ dilepton event selection. The symbol (NI) in front of the second lepton in the event means that isolation is not required as part of lepton identification.

It means , the cut is applied only to events with dilepton mass $M_{\ell\ell}$ lying near the Z boson mass ($\sim 91GeV$). However, it is expected that roughly for 24% of $t\bar{t}$ dilepton events, $M_{\ell\ell}$ will fail into the Z mass region. Therefore, there are applied additional cuts in order to keep such $t\bar{t}$ events while remove background events as much as possible (according the studies on MC, using such cuts, roughly 90 % of the $t\bar{t}$ events in Z boson mass window can be kept, while rejecting about 80 % of Z events). Events must pass the cut on jet significance[6] $JetSig > 8\sqrt{GeV}$ and moreover, it is also required that the jets be at an angle $dPhiJet$ greater than 10° from the $\not{E}_T$ direction.

---

[6] defined below

As we already know, in Z events ($Z \rightarrow ee, Z \rightarrow \mu\mu$ have no neutrinos in final state), false $\not\!\!E_T$ may arise only through mismeasurements of the jets (it can be also leptons but they are usually measured much better than jets). In such case, the $\not\!\!E_T$ vector is expected to point along a direction close to the jet direction while in $t\bar{t}$ events $\not\!\!E_T$ is due to unobserved neutrinos and is usually far away from jet direction. Therefore, there is the cut on angle $dPhiJet$ between jets and $\not\!\!E_T$ which is expected to be high for $t\bar{t}$ events and small for Z events with mismeasured jets.

Jet significance variable takes the above facts also into account. It is defined as:

$$JetSig = \frac{\not\!\!E_T}{\sqrt{\sum_{cos\Delta\phi(\not\!\!E_T, \vec{E}_{T_j})>0} \vec{E}_{T_j} \cdot \left(\frac{\vec{\not\!\!E}_T}{\not\!\!E_T}\right)}}, \tag{4.11}$$

where the denominator sum is over all the jets in the same hemisphere as the missing transverse energy direction $\hat{\not\!\!E}_T = \vec{\not\!\!E}_T / \not\!\!E_T$. It is expected to be high for $t\bar{t}$ events (denominator is small due to large angles between $\not\!\!E_T$ and jets). For Z events with mismeasured jets, it takes into account the fact that the very high energy jet may have naturally larger fluctuation (in this case, larger $\not\!\!E_T$) and therefore the jet significance is expected to be small for Z events with mismeasured jets.

3. *MET cut*:
   Aimed to remove events which don't have neutrinos in final state and consequently only low measured $\not\!\!E_T$ (mainly Drell-Yan events).

   $\not\!\!E_T$ is required to be greater than 25 GeV, where the $\not\!\!E_T$ is already corrected for the primary vertex position, for the presence of any tight or loose muon, and for difference between raw and corrected jet energies.

4. *L cut*:
   another, more advanced cut on $\not\!\!E_T$.

   Rejected are events which have $\not\!\!E_T \leq 50$ GeV and minimum angle between $\not\!\!E_T$ and any jet or lepton in the event is less than or equal 20°.

   This cut takes into account the previously mentioned fact that in events with mismeasured jets the fake $\not\!\!E_T$ is pointing close to the direction of the jet. Moreover, in $Z \rightarrow \tau\tau$ events the $\not\!\!E_T$ vector often points close to the direction of one of the leptons.

5. *NJET cut*:
   this cut is aimed to reject events with not enough jets in final state. Mainly

Drell-Yan and dibosons will be rejected because there must be ISR/FSR in these events in order to have jets in final state.

At least 2 jets are required with corrected $E_T > 15$ GeV and $|\eta_{Det}| < 2.5$ (these are called "tight" jets). η_{Det} means detector η, i.e. η with respect to the nominal origin point (0,0,0), not to the primary vertex position of the event.

6. *HT cut*:
   Because of such heavy top quark, there is higher activity in transverse plane in $t\bar{t}$ events comparing to background events. Therefore, the variable which sums (scalarly) transverse energies of all particles in the event has higher values for $t\bar{t}$ events than in background events. Such variable we call H_T and in our case it is the sum of $\not\!\!E_T$ plus leptons transverse energies plus the E_T of all tight jets in the event.

   H_T is required to be greater than 200 GeV.

7. *OS cut*:
   this cut requires that the charged leptons have the charges with the opposite sign. In $t\bar{t}$ events the charged leptons have indeed the opposite sign of charges because they are coming from opposite charged top quarks (top quark vs. anti-top quark). However, some of the backgrounds (FAKE events) don't necessary have opposite sign leptons, so such events are suppressed by this cut.

After event selection is performed, we apply the remaining corrections on jets, i.e. underlying event correction and out of cone correction. Such corrected jets we use then in mass reconstruction.

For Monte Carlo simulated $t\bar{t}$ events we impose three more pre-requisites for the event to be considered as top dilepton signal:

i. run number $\geq$ 150145: this avoid to bias the acceptance of dilepton categories containing a CMX muon as this cut is hardcoded in the tight CMX muon selection and it affects the muon corrected $\not\!\!E_T$. The effect of this cut in the data is modeled by the lower luminosity used for dilepton categories with a CMX muon.

ii. longitudinal position of event primary vertex, z_{VTX}, has to be inside ±60 cm of the nominal CDF origin.

iii. require a leptonic decay of both W's from $t\bar{t}$ at level of generation (HEPG level)

event category	number of events
Expected $t\bar{t}$	15.7 ± 1.3
Drell-Yan	5.5 ± 1.2
Fakes	3.5 ± 1.4
Diboson	1.6 ± 0.3
Total background	10.5 ± 1.9
Total predicted	26.2 ± 2.3
Observed	33

Table 4.6: Expected $t\bar{t}$ signal and background, and observed number of events for DIL selection method corresponding to luminosity 340 pb^{-1}. A $t\bar{t}$ cross section of 6.1 pb^{-1} is assumed, corresponding to a top mass of 178 GeV (for 175 GeV top mass, the expectation is 17.2 ± 1.4 events).

In the Tab. 4.6 we summarize the results of the selection method applied to the inclusive lepton data set collected from the beginning of Run II (march 2002) up to august 2004. The expected number of $t\bar{t}$ signal events are calculated using PYTHIA [65] Monte Carlo simulation, assuming a production cross section of 6.1 pb^{-1}, corresponding to a top mass of 178 GeV. The Drell-Yan, fakes, and diboson background acceptances are estimated using a combination of Drell-Yan and W+jets data, and PYTHIA and ALPGEN+HERWIG [66, 67] Monte Carlo simulation.

Chapter 5

Top quark mass measurement methods

In this chapter, we will classify, describe and compare the methods for top quark mass measurement.

First, we will try to make description of the present methods from general point of view. Then, we will present particular methods and their principles. However, the details of these methods can be found in separate articles written by the corresponding authors [68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 22]. Later, we will discuss the advantages and disadvantages of particular methods.

In second part of the chapter, we will talk about possible methods and what precision in the measurement they can lead to at the future colliders - Large Hadron Collider (LHC) and International Linear Collider (ILC).

5.1 Top quark mass measurement methods at present

In the high energy physics experiment, we measure particular properties (momentum, energy, sign of charge, decay length, etc.) of the final products of the interaction. Using such measurables, we try to reconstruct the mass of the rapidly decaying top quark which can not be detected directly with the apparatus. However, there are a few reasons which complicates the reconstruction of the top quark mass. First of all, top mass is physics variable which have intrinsic uncertainty [1] due to Heisenberg uncertainty principle [2], therefore it can not be measured exactly even in principle. Consequently, we need as many events with $t\bar{t}$ produced as possible in order to measure top mass

---

[1] the width of the top quark mass distribution $\Gamma_{top} \approx 1.5$ GeV.

[2] in fact, this is true for all the 4-momenta of all the particles

with better statistical uncertainty. Another complication is that some of the products of top quark decay we can not detect at all (neutrinos), due to their extremely small probability of interaction (neutrinos interact only via weak interaction). Another top quark products (b quarks) can not be detected directly, rather only the products (jets of particles) of their hadronization in the process of interaction evolution. Finally, there is uncertainty coming from not perfect detectors and our inability to measure perfectly. It should be noted, that from these complications to our measurement, we can possibly improve only in the last point.

There exists two main approaches to the measurement of top quark mass as of now (autumn 2006). Moreover, one method ("ideogram method") exists at CDF experiment which tries to combine advantages of the both approaches. In both of these two approaches, it is assumed that $t\bar{t}$ production goes via SM decay chain, see Fig. 2.10.

In the first approach, there is always chosen some variable which is sensitive to the top quark mass. Usually, such variable is reconstructed top quark mass from the products of the interaction. In ideal case, such reconstructed mass would be the same as the true input top quark mass. In reality, this is not true. However, as we will see later, many other variables can be used (and actually they are) as top quark mass sensitive variable, e.g. scalar sum of the transverse energy of the products of the interaction - H_T, decay length of b-hadrons in transverse plane, invariant mass of the lepton and b-jet coming from the decay of the same top quark - $M_{\ell b}$, jet transverse energy - E_T, etc. All these variables depend, more or less, on the mass of the input top quark.

After choosing top mass sensitive variable, we reconstruct such variable for the simulated events with known generated top quark mass and for the real experimental data events. The final top quark mass estimate is obtained by comparing the reconstructed sensitive variable between simulated and data events. Usually, the comparison is performed by creating the probability distribution of such sensitive variable for simulated samples (each with different generated top quark mass), so called **templates** [3]. Finally, the top quark mass for which is the distribution most similar with the distribution coming from the real data events is considered to be final top quark mass estimate. This method is therefore called **template method**.

In the second approach, there is calculated the probability that the event with certain input top quark mass will lead to measured variables as we see them in the detector. This probability is calculated according theoretical description of the $t\bar{t}$ production using $t\bar{t}$ production matrix element. Therefore this method is called **matrix**

---

[3] some of the easiest methods produce just one number (e.g. average) and not the distribution of the variable for the sample of the events. Such average is then compared to the top mass dependence of such variable obtained from MC simulated events

element(ME) method. For each event, there is then calculated probability distribution as a function of top quark mass. The final top mass estimate is calculated by taking the mass value for which the product of individual event probabilities is maximal (see Fig. 5.4).

The method which tries to combine both template and matrix element approaches to extract the maximum amount of mass information out of a $t\bar{t}$ candidate event is called **ideogram** method. This method is applied to events in all hadronic channel and we will describe it in corresponding subsection.

This was only general description of both approaches to the top quark mass measurement. More detailed description of actual methods which are used in the experiment will be described later in the chapter. Moreover, even the methods with the same approach differ in many particular ways, so we will even describe more methods within one approach.

We will describe here the methods for top mass measurement which are currently used at CDF. Only other experiment which have been able to observe the top quark and to measure its mass is D0 experiment. However, the methods used there are in principle the same as the ones used at CDF, so we will not mention them.

We will describe the methods according to which of the final-state channel of $t\bar{t}$ production (see Sec. 2.2.2) they use for the measurement, because each channel has its own specialties.

5.1.1 The Dilepton channel

The dilepton channel (see Sec. 2.2.2) is a special because there is not enough measurables to be able to kinematically reconstruct top mass. It is due to two undetected neutrinos in the final state. We are missing six constraints from two not-measured momenta of neutrinos assuming zero neutrino mass. However, we have particular information about neutrinos from the imbalance of the energy of interaction in transverse plane - missing E_T, i.e. its two components - E_{T_x},E_{T_y}. Moreover, we assume value for the $W^{\pm}$ mass and equality between top quark and anti-top quark mass. So, we finally end up with -1 constraints. Therefore, each method which tries kinematically solve the system must make additional assumptions about some kinematic variables (at least one).

In dilepton channel there exists three template methods and two matrix-element methods which are in use as of now (autumn 2006) at CDF. We describe the principles of these methods in the following. One of these methods is our method which we will describe deeply in the next chapter.

The three template methods are in principle different just in the kinematic variables which they make assumption about. All other differences are just in the way of implementation. Our method is one of these template methods and it makes assumption about longitudinal momentum of $t\bar{t}$ system - $p_z^{t\bar{t}}$ (KIN method). Other two template methods make assumption about neutrinos pseudorapidities ("neutrino η weighting algorithm"– NWA) and about neutrinos ϕ angles ("neutrino ϕ weighting method"– PHI), respectively. In our method, we make assumption about only one variable so the system becomes constrained and we can solve the system. The other two method make assumption about 2 variables, the system becomes overconstrained and there is need to select one solution from more possible ones or somehow to weight the solutions.

We will describe the procedure how the template methods obtain top quark mass estimate for particular event in the following subsections. Once the template methods produce a single top quark mass estimate for each event in their corresponding data samples, which are mixture of signal and background samples, the rest of the procedure is very similar for all the template methods. To arrive at a final top quark mass measurement, data events are compared with probability density functions (p.d.f.'s) for signal and background within a likelihood minimization. The p.d.f.'s are developed from template mass distributions created by applying the methods to simulated $t\bar{t}$ signal and background samples, which are then parametrized. For NWA and PHI methods, this parametrization uses a combination of Gaussian and gamma distribution terms. Similarly our KIN method parametrization contains a Gaussian term in conjunction with an approximate Landau distribution.

The final step for each dilepton template analysis is the determination of a representative top quark mass from the data sample by maximizing a likelihood. The likelihood function finds the probability that our data are described by an admixture of background events and dilepton $t\bar{t}$ decays with a certain top quark mass. As input we use the top quark mass values returned by the particular measurement technique applied to the data sample, and the parametrized probability density functions of the signal and background templates derived from simulation. The total likelihood takes the form

$$\mathcal{L}(m_t) \equiv \mathcal{L}_{shape}(m_t) \times \mathcal{L}_{n_{bg}} \tag{5.1}$$

,where

$$\mathcal{L}_{shape}(m_t) = \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \prod_{i=1}^{N} \frac{n_s \times f_s({m_{t_i}}^{rec}, m_t) + n_b \times f_b({m_{t_i}}^{rec})}{n_s+n_b} \tag{5.2}$$

and

$$-\ln \mathcal{L}_{n_{bg}} = \frac{(n_b - n_b^{exp})^2}{2\sigma_{n_b}^2} \tag{5.3}$$

The likelihood returns a true top quark mass hypothesis (m_t), and estimated number of signal (n_s) and background (n_b) events. In the above equation, f_s, f_b are the functions which parametrize the signal and background templates, respectively. The product goes through all N data events which the mass was reconstructed for. n_b^{exp} and σ_{n_b} is the expected number of background events and its uncertainty. The true top quark mass hypothesis (m_t) which minimizes $-\ln \mathcal{L}$ is retained. The statistical uncertainty on m_t is given by the difference between the minimization mass result and the mass at $-\ln(\mathcal{L}/\mathcal{L}_{max}) + 0.5$.

Neutrino η weighting algorithm

This method is one of the template methods. It tries to solve for the unknown neutrino and anti-neutrino momenta, independently of the measured missing energy, by making additional assumptions about the $t\bar{t}$ decay. The neutrino (anti-neutrino) solutions are then compared with the measured $\not{E}_T$ through a weight function in order to create a probability distribution for the event as a function of top quark mass.

The NWA weight function is constructed as follows. There are assumed values for the top quark and W boson masses, the pseudorapidities of the neutrino and anti-neutrino, and the lepton-jet pairings associated with the top (anti-top) decays. Then, there is applied energy-momentum conservation to the top quark decay and two possible solutions for the 4-momentum (ν) of the neutrino are obtained. This procedure is repeated on the anti-top decay, resulting in up to four possible pairs of neutrino -anti-neutrino solutions ($\nu, \bar{\nu}$). Each of the four solutions is assigned a probability (weight, w_i) that it describes the observed missing transverse energy components $\not{E}_{T_x}$ and $\not{E}_{T_y}$ within uncertainties σ_x and σ_y, respectively

$$w_i = exp\left(-\frac{(\not{E}_{T_x} - p_x^{\nu} - p_x^{\bar{\nu}})^2}{2\sigma_x^2}\right) \times exp\left(-\frac{(\not{E}_{T_y} - p_y^{\nu} - p_y^{\bar{\nu}})^2}{2\sigma_y^2}\right) \tag{5.4}$$

The values for σ_x, σ_y ($\sigma_x = \sigma_y = 15$ GeV) are obtained from $t\bar{t}$ Monte Carlo simulation.

Given the assumed top quark mass and assumed neutrino η values, any of the four solution pairs ($\nu, \bar{\nu}$) have *a priori* equal probability. Therefore, it is summed all four weights

$$w(m_t, \eta_\nu, \eta_{\bar{\nu}}, l - jet) = \sum_{i=1}^{4} w_i. \tag{5.5}$$

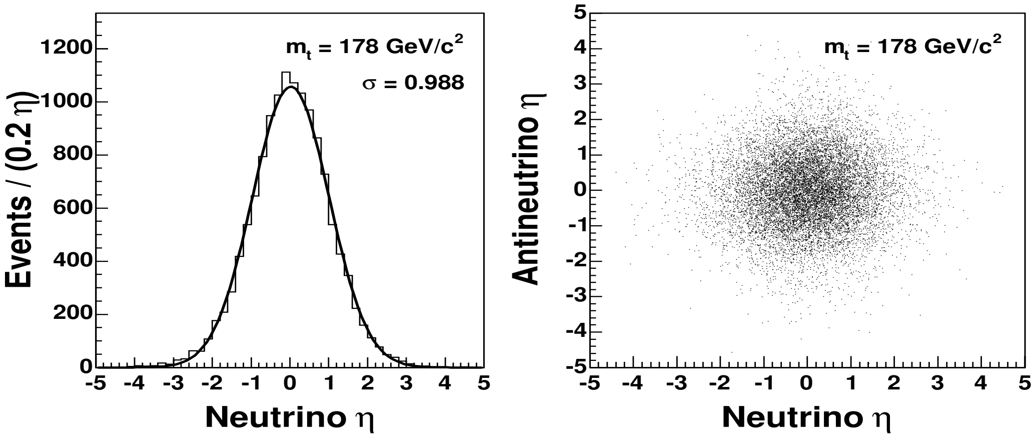


Figure 5.1: Neutrino η distribution with Gaussian fit (left) and neutrino vs anti-neutrino η (right) from a HERWIG $t\bar{t}$ sample with $m_t = 178$ GeV.

Not knowing which are the true neutrino $\eta's$ in our events, we repeat the above steps for many possible $(\eta_\nu, \eta_{\bar{\nu}})$ pairs. As seen in Fig. 5.1, Monte Carlo $t\bar{t}$ simulation indicates the neutrino $\eta's$ are uncorrelated, and follow Gaussian distribution centered at zero with a width near one. The neutrino η is scanned from -3 to +3 in steps of 0.1 and each $(\eta_\nu, \eta_{\bar{\nu}})$ pair is assigned a probability of occurrence $P(\eta_\nu, \eta_{\bar{\nu}})$ derived from a Gaussian of width 0.988. Each trial $(\eta_\nu, \eta_{\bar{\nu}})$ pair contributes to the event according to its weight (Eq.(5.5)) and the probability of occurrence, $P(\eta_\nu, \eta_{\bar{\nu}})$

$$w(m_t, l - jet) = \sum_{\eta_\nu, \eta_{\bar{\nu}}} P(\eta_\nu, \eta_{\bar{\nu}}) \cdot w(m_t, \eta_\nu, \eta_{\bar{\nu}}, l - jet) \tag{5.6}$$

Since we don't distinguish b jets from $\bar{b}$ jets, both possible lepton-jet pairings are summed. Then, the final weight becomes a function only of the top quark mass, after integrating over all other unknowns

$$W(m_t) = \sum_{\ell^+ - jet_1}^{\ell^+ - jet_2} w(m_t, l - jet) \tag{5.7}$$

We scan m_t from 80 to 380 GeV in steps of 1 GeV. Fig. 5.2 shows the resulting normalized weight distribution from Eq. 5.7 after applying the NWA method to a HERWIG Monte Carlo $t\bar{t}$ event, with a simulated top quark mass of 170 GeV. We choose one indicative top quark mass for each event, selecting the most probable value (MPV) of the weight distribution as that which best explains the event as a $t\bar{t}$ dilepton decay.

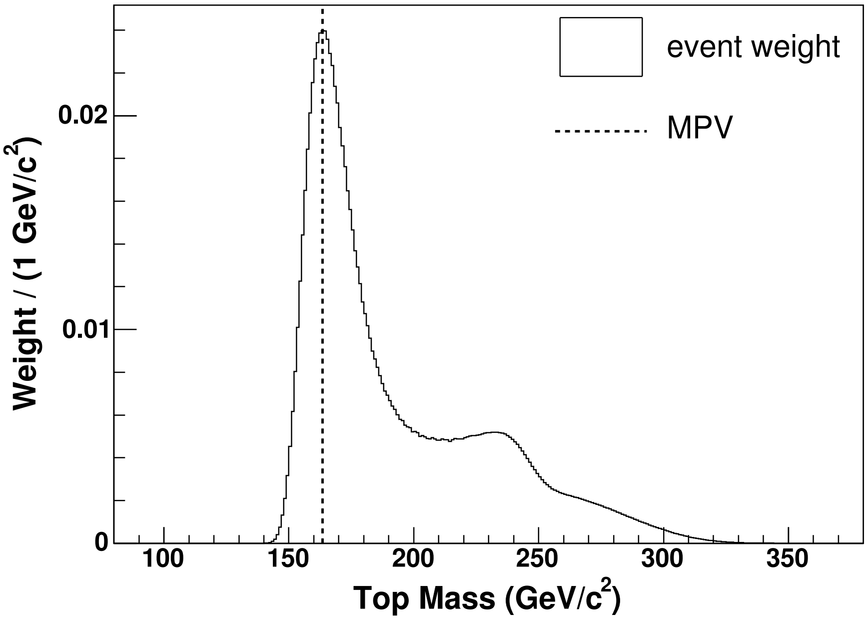


Figure 5.2: NWA weight distribution as a function of top quark mass hypothesis for a HERWIG Monte Carlo $t\bar{t}$ event with $m_t = 170$ GeV. The vertical line denotes the most probable value (MPV) of m_t chosen by the method.

Neutrino ϕ weighting method

This method is also one of the template methods. Introducing additional assumptions about the azimuthal angle ϕ of the final state neutrinos, this method reconstructs dilepton decays through the minimization of a chi-square functional (χ^2) to arrive at a single top quark mass for each event.

The χ^2 functional to be minimized takes the form

$$\begin{aligned}\chi^2 \;=\; & \sum_{\ell=1}^{2}\frac{(p_T^\ell-\widetilde{p_T^\ell})^2}{(\sigma_{p_T}^\ell)^2}+\sum_{j=1}^{2}\frac{(p_T^j-\widetilde{p_T^j})^2}{(\sigma_{p_T}^j)^2}+\sum_{i=1}^{N}\frac{(UE^i-\widetilde{UE^i})^2}{(\sigma_{UE}^i)^2}+ \\ & \frac{(m_{\ell_1\nu_1}-m_W)^2}{\Gamma_W^2}+\frac{(m_{\ell_2\nu_2}-m_W)^2}{\Gamma_W^2}+\frac{(m_{j_1\ell_1\nu_1}-\widetilde{m_t})^2}{\Gamma_t^2}+\frac{(m_{j_2\ell_2\nu_2}-\widetilde{m_t})^2}{\Gamma_t^2}\end{aligned} \tag{5.8}$$

The first two terms sum over primary lepton transverse momenta and transverse momenta of leading two jets, respectively. The quantity UE in the third term denotes the unclustered energy in the calorimeter, summed over all towers, which is not associated with a lepton or leading jet calorimeter cluster, but includes any additional jets. The last four terms restrict the invariant mass of lepton and neutrino ($m_{\ell\nu}$) to the mass of W boson (m_W) and invariant mass of jet, lepton and neutrino to the top quark mass. Variables with tilde refer to the output of the minimization procedure. The quantity $\widetilde{m_t}$ is the fit parameter returned as the reconstructed top quark mass for the combination being considered.

To resolve the neutrino momentum used in the W and top decay constraints of Eq.(5.8), two additional assumptions are introduced. Assuming values for both neutrino azimuthal angles ($\phi_{\nu 1}, \phi_{\nu 2}$), the transverse momenta of the neutrinos are linked through the measured $\not{E}_T$ by

$$\begin{aligned} p_T^{\nu 1}\cos(\phi_{\nu 1}) + p_T^{\nu 2}\cos(\phi_{\nu 2}) &= \not{E}_{T_x} \\ p_T^{\nu 1}\sin(\phi_{\nu 1}) + p_T^{\nu 2}\sin(\phi_{\nu 2}) &= \not{E}_{T_y} \end{aligned} \tag{5.9}$$

leading to the solutions for $p_x^{\nu 1}, p_y^{\nu 1}, p_x^{\nu 2}, p_y^{\nu 2}$.

Performing the χ^2 minimization of Eq.(5.8) on all allowed values of neutrino ϕ creates a set of solutions in the ($\phi_{\nu 1}, \phi_{\nu 2}$) plane. A grid of 12×12 points in the ($\phi_{\nu 1}, \phi_{\nu 2}$) plane is chosen. In practice, only points in the quadrant ($0 < \phi_{\nu 1} < \pi, 0 < \phi_{\nu 2} < \pi$) need to be sampled since identical neutrino components occur for the four points ($\phi_{\nu 1}, \phi_{\nu 2}$), ($\phi_{\nu 1}+\pi, \phi_{\nu 2}$), ($\phi_{\nu 1}, \phi_{\nu 2}+\pi$), ($\phi_{\nu 1}+\pi, \phi_{\nu 2}+\pi$) .At each point, 8 solutions exist due to the two-fold ambiguity in longitudinal momentum for each neutrino, and the two possible lepton-jet combinations. Thus, for each event, $12 \times 12 \times 8 = 1152$ minimizations of Eq.(5.8) are performed, each returning an output χ^2 and reconstructed top quark mass m^{rec}. The minimal value for χ^2 among eight possible solutions at each point is retained, reducing each event to an array of 144 χ^2_{ij} and m^{rec}_{ij} values, where $i, j = 1, .., 12$ referee to the ($\phi_{\nu 1}, \phi_{\nu 2}$) grid points. Each point is weighted by its returned χ^2 value according to

$$w_{ij} = \frac{exp(-\chi^2_{ij}/2)}{\sum_{i=1}^{12}\sum_{j=1}^{12} exp(-\chi^2_{ij}/2)} \tag{5.10}$$

to create a probability density distribution normalized to unity.

To arrive at a single top quark mass value per event, the reconstructed mass values m^{rec}_{ij} of the array are averaged, using the weights derived from Eq.(5.10). The sensitivity to top quark mass is enhanced by averaging only points with a weight at least 30 % that of the most probable value in the probability density distribution. Fig. 5.3 shows the results of the PHI method applied to a HERWIG Monte Carlo $t\bar{t}$ event with $m_t = 170$ GeV.

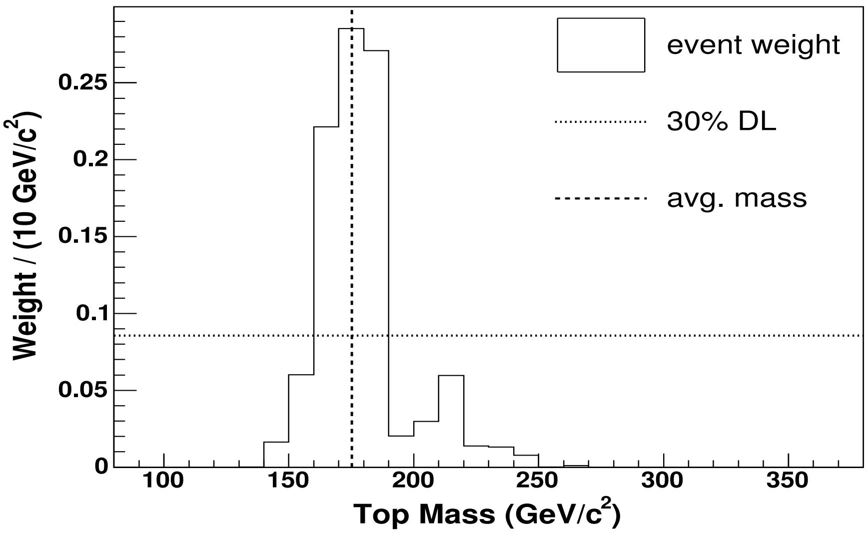


Figure 5.3: Binned weight distribution from the PHI method for a HERWIG Monte Carlo $t\bar{t}$ event with $m_t = 170$ GeV, showing the resulting average mass for bins above the 30 % discriminantion level (DL).

Matrix element method

This method, obviously according its title, use ME approach to the measurement of the top quark mass. The probability density for $t\bar{t}$ decays is expressed as $P_s(\mathbf{x}|M_t)$, where M_t is the top quark mass and $\mathbf{x}$ contains the measured lepton and jet momenta. This probability is calculated using theoretical description of the $t\bar{t}$ production process expressed with the respect to $\mathbf{x}$, $P_s(\mathbf{x}|M_t) = \frac{1}{\sigma(M_t)}\frac{d\sigma(M_t)}{d\mathbf{x}}$, where $\frac{d\sigma}{d\mathbf{x}}$ is the differential cross-section and σ is the total cross-section. To evaluate the probability density, we integrate the leading-order matrix element of $t\bar{t}$ production process over quantities which are not directly measured by the detector, i.e. neutrino and quark energies. We assume that lepton momenta are perfectly measured, that quark angles are perfectly measured by the corresponding jets angles, and that the two most energetic jets correspond to the b b quarks from top quark decay. Quark energies, while not directly measured, are estimated from the observed energies of the corresponding jets. We define the transfer function $W(p, j)$ to be the probability of measuring jet energy j given quark energy p. We approximate $W(p, j)$ as a sum of two Gaussians fitted to the predicted distribution of quark-jet energy difference from $t\bar{t}$ events generated with HERWIG and the CDF II detector simulation. The expression for the probability density at a given mass for a

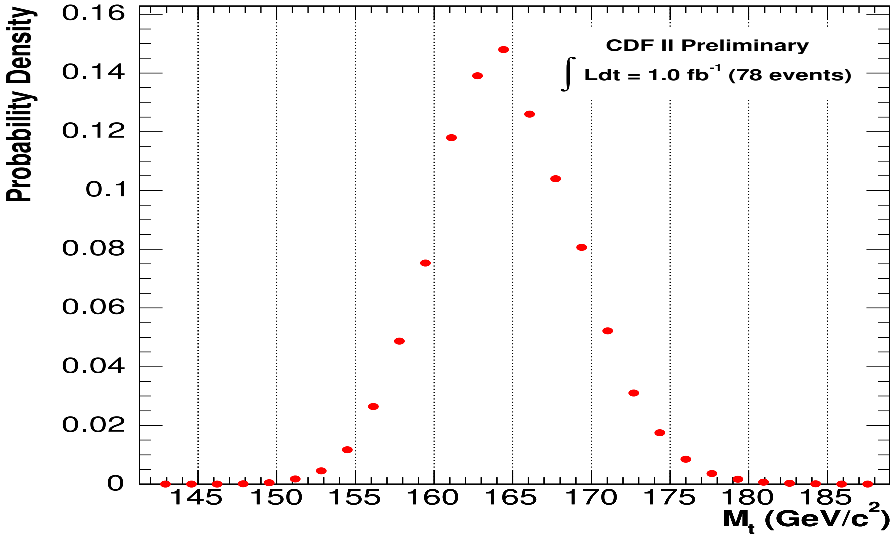


Figure 5.4: Final posterior probability density as a function of top quark mass for the 78 dilepton candidate events in data.

specific event can be written as

$$P_s(\mathbf{x}|M_t) = \frac{1}{\sigma(M_t)} \int d\Phi |\mathcal{M}_{t\bar{t}}(q_i, p_i, M_t)|^2 \prod_{jets} W(p_i, j_i) f_{PDF}(q_1) f_{PDF}(q_2), \qquad (5.11)$$

where the integral is over the momenta of the initial and final state particles, q_1 and q_2 are the incoming momenta, p_i are the outgoing momenta, f_{PDF} are the parton distribution functions and $|\mathcal{M}_{t\bar{t}}(q_i, p_i, M_t)|$ is the $t\bar{t}$ production and decay matrix element. The term $1/\sigma(M_t)$ in front of the integral ensures that the normalization condition for the probability density is satisfied. The probability as in Eq. 5.11 is calculated also for main background processes using background matrix element evaluated numerically. The generalized per-event probability density is formed $P(\mathbf{x}|M_t) = P_s(\mathbf{x}|M_t)p_s(M_t) + P_{bg1}(\mathbf{x})p_{bg_1} + P_{bg2}(\mathbf{x})p_{bg_2} + ...$, where p_s and p_{bg_i} are determined from the expected number of events for signal and background processes. The probability density in top mass for a given data sample is the product of individual event likelihoods, see Fig. 5.4. The mass measurement (M_t) is the mean of the posterior probability. Of course, due to various reasons, the top quark mass estimate performed by the above described method can be shifted form the true value. Therefore there is performed test of the method using pseudo-experiments. According results in pseudo-experiments there is derived correction to top mass estimate and also to estimate on

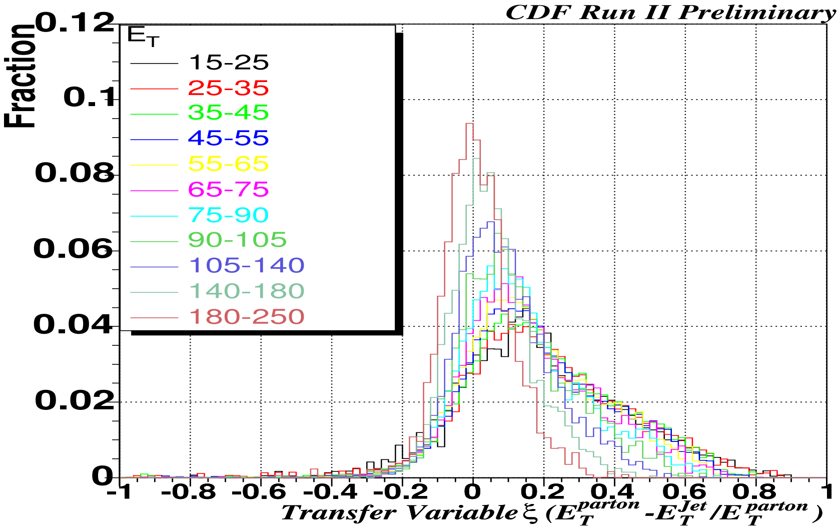


Figure 5.5: Transfer functions for jet E_T. Different histograms correspond to different ranges of partons E_T.

top mass uncertainty.

Dynamical likelihood method

This method is also one of the ME methods, and in principal, it is very similar to the above described “ME method”. There are some small differences in assumptions between these two methods.

In Dynamical likelihood method (DLM), the definition of likelihood for a each event is based on the differential cross section per unit phase space volume of the final partons, $\mathrm{d}\sigma_{t\bar{t}}/d\Phi$ using the leading order matrix element as a function of M_{top}.

Detector resolution effects are accounted for using $t\bar{t}$ events generated by the HERWIG MC simulation and full detector simulation to derive a transfer function (TF). The transfer function are defined the same way as in ME method, i.e. it relates the quarks and observed jets energies, see Fig. 5.5.

For a given event, a Monte Carlo integration is performed over the possible $t\bar{t}$ final state kinematics in the following way. we first generate a random value for a virtual mass squared of the W bosons and top quarks, $s_{W^{\pm}}, s_t, s_{\bar{t}}$, according Breit-Wigner distribution. We identify the momentum of the electron or muon daughter with the measured value. We then generate random values for the momenta of final state quarks according to the TF probabilities. To reconstruct neutrino momenta, some resonance mass relations are required and they can be transfered to the fourth-order equation which can be analytically solved. Because there are 2 missing particles (neutrinos) in dilepton channel, and while measurement of $\not\!\!E_T$ gives only 2 constraints, so there are 4 constraints still needed to solve neutrino momenta analytically. Therefore, there are required 4 mass relations of $t, \bar{t}, W^+, W^-$

$$\begin{gathered}(b + l^+ + \nu)^2 = s_t, (\bar{b} + l^- + \bar{\nu})^2 = s_{\bar{t}} \\ (l^+ + \nu)^2 = s_{\bar{l}\nu}, (l^- + \bar{\nu})^2 = s_{l\bar{\nu}}\end{gathered} \tag{5.12}$$

However, some solutions are coming from multiple possible neutrino momenta and are not appropriate. Therefore, it is applied the neutrino kinematics weight to enhance only good solutions. The weighting function simply employed the variables $U_X = \nu_x - \bar{\nu}_x, U_y = \nu_y - \bar{\nu}_y, \nu_z, \bar{\nu}_z$. The distributions of U, ν variables are fitted with Gauss function convoluted with exponential function because the weights have slightly top mass dependence.

Thus, for a given set of $\mathbf{x}$ and $s_{W^\pm}, s_t, s_{\bar{t}}$, we fully determine the event kinematics, and the event likelihood as a function of top mass M_{top} is given by

$$L(M_{top}) = N \sum_{I_j} \sum_{I_\nu} \frac{\mathrm{d}\sigma_{t\bar{t}}}{\mathrm{d}\Phi}(M_{top}; x, s_{W^\pm}, s_t, s_{\bar{t}}) \tag{5.13}$$

where the normalization factor N is independent of M_{top} for a given event, and the indices I_J and I_ν run over the parton-jet assignments and the two neutrino solutions, respectively. The event likelihood is obtained by numerically integrating over $\mathbf{x}$ given by the TF and s_W given by the Breit-Wigner distribution. We account for the presence of the background events by evaluating the shift they make in the measured top quark mass.

5.1.2 Lepton-jets channel

The lepton+jets channel (see Sec. 2.2.2) has been up to now the most preferred channel for the measurement of top quark mass because the most precise measurements are coming from this channel. In final state, there is only one neutrino, so we can identify a missing transverse energy $\not\!\!E_T$ with the energy of neutrino. The only ambiguity is coming from longitudinal component of neutrino momentum (there are two possibilities). Therefore, we are able to kinematically solve the system.

As of now, there exists three template methods and two matrix element methods in this channel.

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

This template method is in its implementation similar to Neutrino ϕ template method in dilepton channel (see previous section). Moreover, it takes the advantage of the possibility to calibrate jet energy scale in lepton+jets channel. By doing so, it reduce the uncertainty due to jet energy scale which is the main source of systematic uncertainty at present.

We generate a set of Monte Carlo samples at a range of true top quark masses (M_{top}) and a range of jet energy scale calibrations shifts (Δ_{JES}). We form good estimators of both of these parameters: reconstructed top mass m_t^{reco} and qq' dijet mass m_{jj}. The reconstructed top mass m_t^{reco} is extracted from each event using a χ^2 fitter. Measurement of M_{top} and the *in situ* jet energy scale calibration Δ_{JES} is performed by comparing the m_t^{reco} and m_{jj} distributions obtained from the data to these templates using an unbinned likelihood fit.

For each event, an invariant mass of the top quark is reconstructed from the top decay products using a χ^2 kinematic fit. The χ^2 definition is in principal the same as the one defined in Eq.(5.8) for Neutrino ϕ weighting method, just in lepton+jets channel there is one less charged lepton and two more jets. The χ^2 is minimized for all jet-to-parton assignments consistent with the b-tagging information. It is used m_t from the fit that yields the lowest χ^2, provided that $\chi^2 < 9$.

To perform *in situ* jet energy scale calibration we construct the dijet mass m_{jj}. This is simply invariant mass of the measured jets pairs which are not b-tagged. Note that in an event where one or fewer b tags are present it is possible to form more than one m_{jj}.

The rest of the procedure is almost identical to other template methods except of additional templates for invariant dijet mass. The m_t and m_{jj} distributions (templates) are parametrized as a function of two parameters, M_{top} and Δ_{JES} to obtain two probability density functions (the one-dimensional visualizations of the templates are in Fig. 5.6).

The reconstructed mass distributions from data are compared to the signal and background templates using unbinned likelihood fit. The likelihood fits for the expectation values of the number of signal and background events, and for the top quark mass and jet energy scale. The likelihood is given by:

$$\mathcal{L}(M_{top}, \Delta_{JES}, n_s, n_b) \equiv \mathcal{L}_{shape}^{m_t^{reco}} \times \mathcal{L}_{shape}^{m_{jj}} \times \mathcal{L}_{nev} \times \mathcal{L}_{bg} \tag{5.14}$$

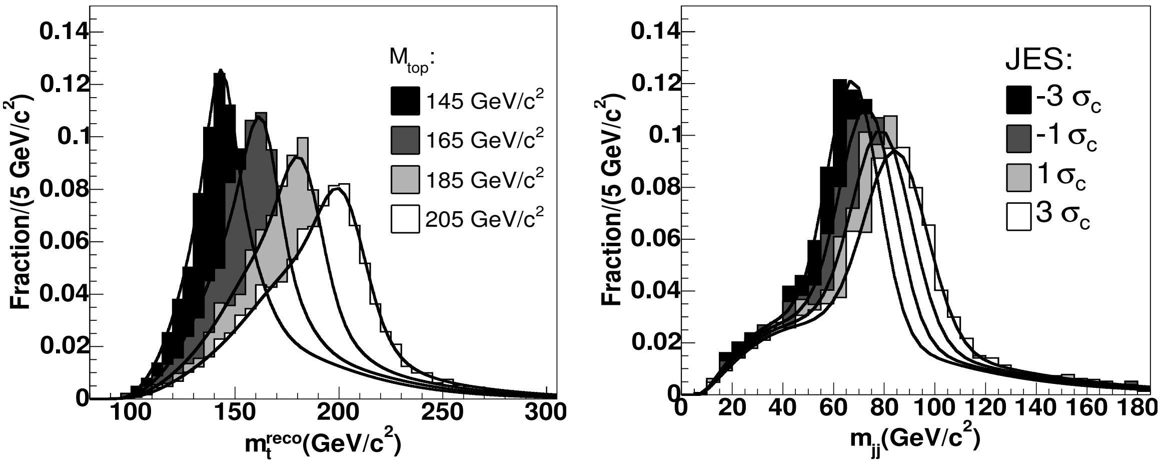


Figure 5.6: Top mass templates and parameterizations for a range of true top masses at $\Delta_{JES} = 0$ (left). m_{jj} templates for several values of Δ_{JES} and m_{top} fixed at 180 GeV.

,where

$$\begin{aligned}
\mathcal{L}_{shape}^{m_t^{reco}} &= \prod_{k=1}^{N^{\chi^2}} \frac{\epsilon_s n_s P_s(m_{t,k}^{reco}; M_{top}, \Delta_{JES}) + \epsilon_b n_b P_b(m_{t,k}^{reco})}{\epsilon_s n_s + \epsilon_b n_b} \\
\mathcal{L}_{shape}^{m_{jj}} &= \prod_{k=1}^{N \cdot C_i} \frac{n_s P_s(m_{jj,k}; M_{top}, \Delta_{JES}) + n_b P_b(m_{jj,k})}{n_s + n_b} \\
\mathcal{L}_{nev} &= \sum_{N_s + N_b = N} Pois(N_s; n_s) \times Pois(N_b; n_b) \times \\
& \left[\sum_{N_s^{\chi^2} + N_b^{\chi^2} = N^{\chi^2}}^{N_{s,b}^{\chi^2} \leq N_{s,b}} B(N_s^{\chi^2}; N_s, \epsilon_s) B(N_b^{\chi^2}; N_b, \epsilon_b) \right] \\
-\ln \mathcal{L}_{n_{bg}} &= \frac{(n_b - n_b^{exp})^2}{2\sigma_{n_b}^2}
\end{aligned} \quad (5.15)$$

The free parameters in the fit are $M_{top}, \Delta_{JES}, n_b, n_s$ - top mass, jet energy scale shift, expected number of background events and expected number of signal events.

The values ϵ_s and ϵ_b represents the efficiency of the χ^2 cut for signal and background events, respectively. N and N^{χ^2} are the numbers of events observed in the data before and after the χ^2 cut.

The $k-th$ term of $\mathcal{L}_{shape}^{m_t^{reco}}$ gives the probability of observing the $k-th$ data event with reconstructed mass $m_{t,k}^{reco}$ for given M_{top} and Δ_{JES} values. This term is most sensitive to fitted M_{top} value.

Similarly, $\mathcal{L}_{shape}^{m_{jj}}$ is sensitive to Δ_{JES} value. Note that all the m_{jj} contributions from different dijet combinations C_i are taken in the product.

The third term in likelihood,$\mathcal{L}_{nev}$, captures the information arising from the number of signal and background events in the top quark mass and dijet mass samples, which are correlated. The number of observed events N and number of events passing χ^2 cut N^{χ^2} are known, but we don't know how many signal N_s and background N_b events exist within the sample. Also unknown are $N_s^{\chi^2}$ and $N_b^{\chi^2}$ - numbers of signal and background events passing χ^2 cut. We sum over all possibilities of assigning $N_s, N_b, N_s^{\chi^2}$ and $N_b^{\chi^2}$ and calculate the probability of such assignment given n_s and n_b values.

The $\mathcal{L}_{bg}$ term is a Gaussian constraint on the expected number of background events.

In order to increase sensitivity of the measurement, the measurements for samples with different number of b-tagged jets are combined together. The total likelihood is therefore product of likelihood for all subsamples with a Gaussian constraint on Δ_{JES} with mean of 0 and with of 1. The constraint is imposed since we trust our *a priori* jet energy scale calibration to within $\pm 1\sigma$.

Multivariate template method

This is, as title already says, one of the template methods. In its principals, it is very similar to template method described above as it use energy-momentum conservation law to reconstruct top quark mass in a given event. However, it tries to make several improvements. First, the method attempts to reduce jet systematics by calibrating the jet energy scale using W mass constraint event-by-event [4]. Moreover, it reduces statistical uncertainty by estimating the probability to pick the correct jet permutation on event-by-event basis and reweight the events according to this probability. Finally, it improves the signal/background separation by utilizing other kinematic variables in addition to the reconstructed top mass, i.e. it extracts the top mass from multidimensional templates (2D) obtained from Monte Carlo simulated samples.

The more detailed description of the method is as follows. In the procedure of top mass reconstruction, we consider only those parton assignments to the four leading jets which are consistent with the b-tagging information For each parton assignment, we attempt to improve the resolution by integrating over the W mass Breit-Wigner distribution. For each of the two W bosons in the event, we use a set of seven m_W^2 values equidistant in the cumulative Breit-Wigner probability. A jet energy scale factor, JES,

---

[4] at the time when this method produced the results, this was a new idea. However, since then, also other methods implemented this calibration of jet energy in order to reduce systematic uncertainty

is included in the W mass kinematic fit . The momenta of all jets in the event for a particular permutation are multiplied by the jet energy scale factor obtained from the mass fit of the hadronically-decaying W. The mass of top quark for a given event and given m_W is then determined. The top mass value for a given event is then weighted combination of top masses corresponding to different points in the W mass integration grid.

The presence of four jets in the final state allows many permutations for the assignment of a jet to parton from $t\bar{t}$ process. In this analysis, there are used events $\geq 1b$-tagged events, so we can have up to 6 jet to parton assignments possible. We split our signal MC events into three subsamples:

- events in which the four leading jets correspond to the four daughter quarks and the jet permutation with the smallest χ^2 results in the correct jet to parton assignment.
- events in which the four leading jets correspond to the four daughter quarks, but the jet permutation with the smallest χ^2 does not results in the correct jet to parton assignment.
- events in which one or more leading jets do not come from $t\bar{t} \to l\nu bq\bar{q}\bar{b}$. This subsample includes all $t\bar{t}$ MC events which do not belong to the first two subsamples.

The reconstructed mass distribution for these subsamples is illustrated in Fig. 5.7. These plots show that knowledge of the subsample to which each event belongs is crucial to the resolution of the reconstructed mass distribution. Perfect $t\bar{t}$ subsample separation can not be achieved in practice. Instead, to each data event we assign the probabilities to belong to different subsamples. We employ a model in which we consider differences between χ^2 values of the best permutation and all other permutations, as well as kinematic properties of $t\bar{t}$ events. We consider two kinematic variables: $\cos(\phi)$, where ϕ is the angle between the lepton and b (the b associated with the leptonic top) in the rest frame of the W which decays into $\ell\nu$ and the product $\cos(\theta_1)\cos(\theta_2)$ where θ_1 is the angle between lepton and the beam axis in the leptonic top quark rest frame and θ_2 is the angle between the direction of the light quark and beam axis in hadronic top quark rest frame. We use that light quark which in the rest frame of the parent W has smaller angle with the b quark originating from the hadronically decaying top.

For the multivariate template method we seek a set of “good” variables which both increase the sensitivity of the likelihood to the top quark mass and improve the discrim-

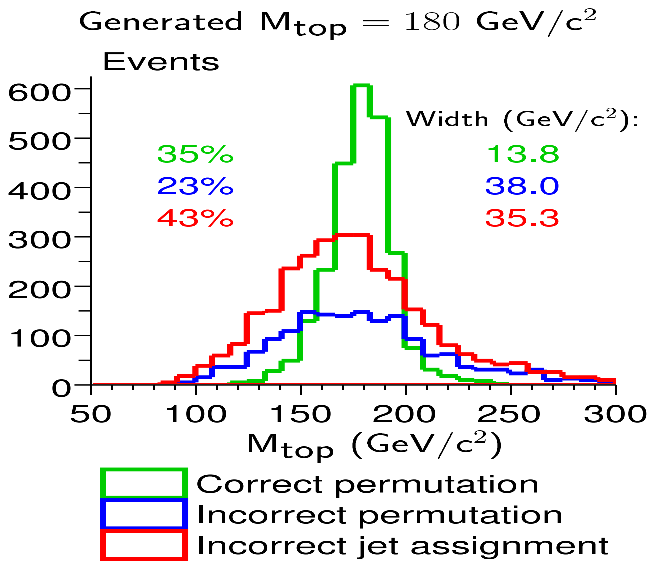


Figure 5.7: Reconstructed top mass for different permutations of jets.

ination between signal and background. The set of variables should be also optimal with respect to the systematic uncertainties of the measurement. To select between many possible kinematic and angular/shape variables we compared the probability density distributions for particular variables between signal and background (W + 4 jets) sample. Using a few different measures of distance (e.g. Kolmogorov-Smirnov distance) lead to the following conclusion: kinematic energy quantities discriminate signal from background better than angle or shape variables. It was found that the sum of the transverse energies of the four leading jets is the best performer.

For the construction of the templates we have chosen a nonparametric multivariate density reconstruction method called *Kernel Density Estimation* (KDE). The density at any given point in the multivariate space receives contribution from Monte Carlo events in such a way that events are weighted more heavily when they appear closer to the desired point. The reason for choosing KDE was as follows. Using histograms for building the templates becomes impractical in more than one dimension because the number of events needed to populate the histogram rises very quickly as more variables are added. It is also desirable to try a variety of variable and their combinations in order to optimize the expected top mass resolution. Therefore, it appear that one of the nonparametric multivariate density reconstruction method should be used. Although a wide variety of such techniques is available (neural nets, etc.) we have chosen to use KDE method due to its maturity, conceptual simplicity and straightforward software implementation. As mentioned above, we choose sum of the transverse energies of the four leading jets as the variable for creation of the templates besides the top mass. The templates for different top mass values and for different samples are in Fig. 5.8

The final top mass estimate is, as usual, done by likelihood maximization. The likelihood is defined the following way

$$\mathcal{L}(m_t) = \prod_{i=1}^{N} (f_b P_b(m_i, x_i) + (1 - f_b) P_s(m_i, x_i, m_t)) \tag{5.16}$$

where N is the number of observed events, m_i is the top mass in the i-th event (determined for the jet permutation with the lowest χ^2), x_i symbolizes all template variables besides measured mass, P_s and P_b are the signal and background densities in the (m_t, x) space, f_b is the background fraction, which is allowed to float freely. The signal density is composed as follows

$$P_s(m, x, m_t) = (1 - p_{CJ}) S_{2,m_t}(m, x) + p_{CJ}(p_{GP} S_{0,m_t}(m, x) + (1 - p_{GP}) S_{1,m_t}(m, x)) \tag{5.17}$$

where S_{0,m_t}, S_{1,m_t} and S_{2,m_t} are the tree signal templates for different subsamples. P_b is just combination of background templates according expected number of background events of particular types.

Template method using decay length technique

I consider this method once again the template method, because, as it will be described below, it use simulated events to produce certain template (even it is simplified one comparing to other template methods) which data events are compared to.

This method use the transverse decay length of b-hadrons from top decay. This technique relays solely on tracking and thus avoids the jet energy scale uncertainty that is largest systematic uncertainty common to all other methods (with increasing statistics systematic errors will dominate the statistical error).

The method exploits the fact that top quarks at the Tevatron are produced nearly at rest (similarly it will be true for the LHC due to dominant production by gluon fusion). In top quark rest frame, the W and b daughters have equal and opposite momenta of magnitude:

$$p = \frac{m_t}{2} \sqrt{(1 - (M_W^2 + m_b^2)/m_t^2)^2 - 4(M_W m_b / m_t^2)^2} \tag{5.18}$$

Thus, $p \sim 0.4 m_t$ for large m_t and the boost of the b quark is substantial. As a consequence, the b quark momentum is potentially sensitive gauge of the top quark

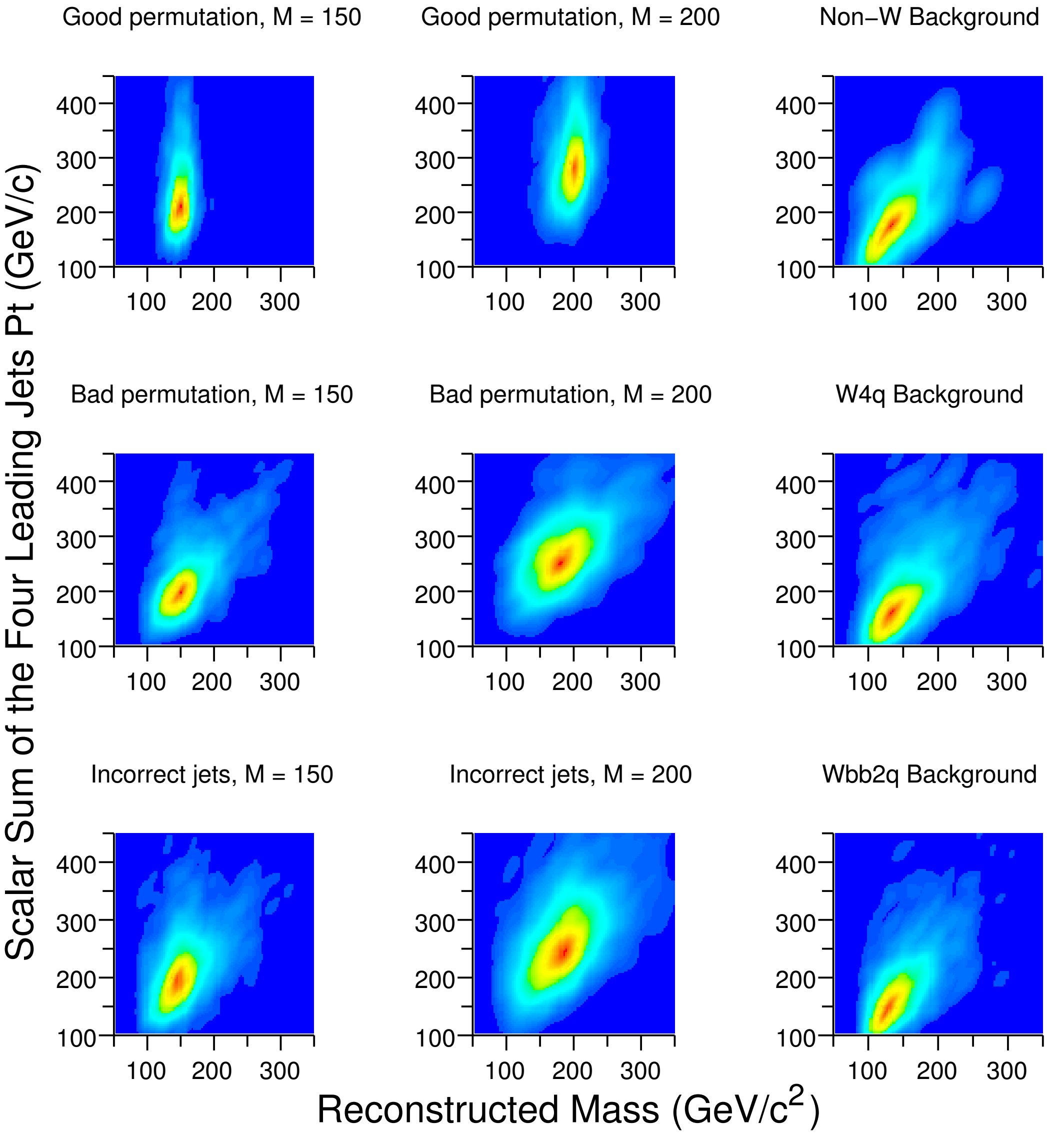


Figure 5.8: Illustration of 2-dimensional templates based on the invariant top mass and the scalar sum of the transverse energies of the first four leading jets. The templates for events with the correct jet combination, for combination that use the correct jets but not with the correct assignment and for the events for which one or more of the 4 jets are not decay products of the two top candidates are presented.

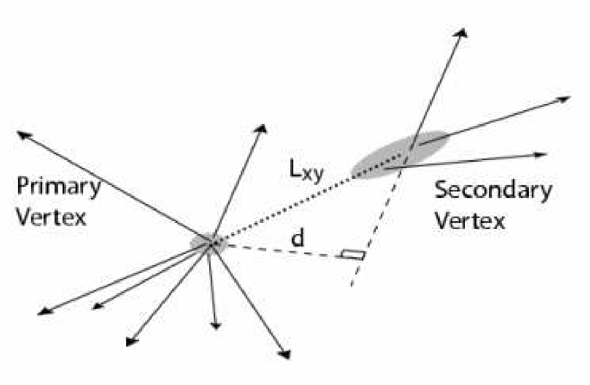


Figure 5.9: The visualization of transverse decay length.

mass. One could use the correlation of the b jet energy to the top quark mass directly. The drawback of such approach, however, is that it depends upon the b jet energy measurement which suffers from a jet energy scale uncertainty like other methods.

As an alternative, the average lifetime of the b-hadrons resultant from top decays can be used because the momentum of particle is correlated with the boost of the particle which is consequently correlated with the lifetime of particle in laboratory frame (or equivalently to the distance traveled by the particle until it decays).

In this analysis, rather than measuring the lifetime, it simply measures the experimentally more accessible average transverse decay length of the b-hadrons, see Fig. 5.9. For a b hadron of momentum p, mass m_b, and proper life time τ_0 one obtains:

$$< L >= c\tau_0\beta\gamma = \tau_0\frac{p}{m_b} \tag{5.19}$$

The key point is that $< L >$ can be found directly without measuring energy E. It is simply the average measured distance from the primary interaction vertex to the b hadron decay vertex. It thus depends upon charged particle track reconstruction as opposed to jet energy reconstruction.

In practice, it is used average of the transverse decay length , $L_{xy} = L|\sin\theta|$ where θ is the angle of the b hadron flight path with respect to the beam axis. This is necessitated by the fact that the net longitudinal momentum of the $t\bar{t}$ pair is not known in hadron collisions. The partons themselves have broad momentum distributions within the proton or anti-proton. The sum over the transverse momenta of all objects in the event must however be zero.

The mass measurement procedure itself is similar to other template methods. The distribution of L_{xy} variable is obtained for simulated events for wide range of input top quark mass. The little difference comparing to other template methods comes now. The

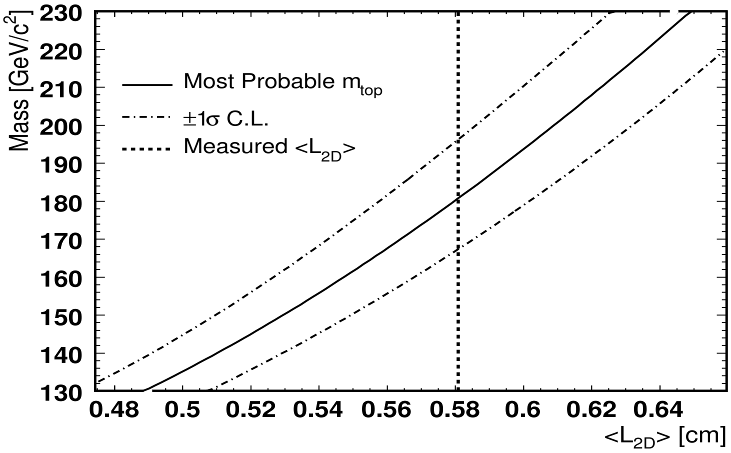


Figure 5.10: Most probable top mass and 1 σ confidence intervals as a function of mean transverse decay length. The mean decay length measured in 695 pb-1 of CDF data is overlaid as the dashed line, from which the measured top mass and its uncertainty can be read off.

L_{xy} distributions are not fitted by some smooth functions and the L_{xy} distribution from data events is not compared to these templates. Rather the measurement use a little bit simplified version of template method. The signal and background L_{xy} distributions are treated as probability density functions from which ensembles of pseudo-experiments are formed. In forming each ensemble, the number of events are converted to a number of b-tags . This procedure is repeated many times (1000) for each top quark mass point. The mean L_{xy} that results from each pseudo-experiment is histogrammed from which the mean and $\pm 1\sigma$ variance are extracted as a function of mass. These points are then fit to smooth function (in this case, it is third degree polynomial). This is just one template used in this method (comparing to many templates corresponding to each mass point in other template methods). The fit to the mean establishes the most probable value for a true top mass given a measured mean L_{xy} and is the function that is used to make the top mass measurement from the measured mean L_{xy} extracted from data, see Fig. 5.10.

Matrix Element analysis technique with *in situ* $W \to jj$ calibration

This method belongs to ME methods and is quite similar to ME method used in dilepton channel. It calculates the event probability $P(\vec{x})$ for a given set of measured parameters, $\vec{x}$, by integrating over parton level differential cross-section, $d\sigma(\vec{y})$, using a *transverse function*, $W(\vec{x}, \vec{y})$, to translate the measured parameters to final state parton level with their corresponding resolution

$$P(\vec{x}) = \frac{1}{\sigma_{t\bar{t}}} \int d\sigma(\vec{y}) W(\vec{x}, \vec{y}), \tag{5.20}$$

where $\sigma_{t\bar{t}}$ normalizes the signal probability. In order to obtain the differential cross-section $d\sigma$ for $p\bar{p} \rightarrow t\bar{t}$, we convolute the integral with the *parton distribution function* (PDF) for initial partons, $f(\tilde{q}_i)$, and include an integral over the transverse components of the momentum of $t\bar{t}$ system

$$P(\vec{x}) = \frac{1}{\sigma_{t\bar{t}}} \int d\sigma(\vec{y}) f(\tilde{q}_1) f(\tilde{q}_2) W(\vec{x}, \vec{y}) d\tilde{q}_1 d\tilde{q}_2 d\tilde{p}^x_{t\bar{t}} d\tilde{p}^y_{t\bar{t}}, \tag{5.21}$$

The PDFs account for the unknown longitudinal momentum of colliding partons and are defined in the usual manner as the probability a parton carries longitudinal momentum fraction $\tilde{q}$ and $\tilde{q} + d\tilde{q}$. The integral over transverse momentum of the $t\bar{t}$ system allows for initial state radiation and any p_T in the $q\bar{q}$ system [5], where $d\tilde{p}^i_{t\bar{t}} = dp^i_{t\bar{t}} / \int dp^i_{t\bar{t}}$.

There are six final state particles in $t\bar{t}$ decay, thus Eq.(5.21) can be expressed as

$$P(\vec{x}) = \frac{1}{\sigma_{t\bar{t}}} \frac{(2\pi)^4}{16} \int |\mathcal{M}|^2 \frac{f(\tilde{q}_1)}{|q_1|} \frac{f(\tilde{q}_2)}{|q_2|} W(\vec{x}, \vec{y}) d\Phi_6 d\tilde{q}_1 d\tilde{q}_2 d\tilde{p}^x_{t\bar{t}} d\tilde{p}^y_{t\bar{t}}, \tag{5.22}$$

where we neglected masses and transverse momenta of the initial state particles and the 6-body phase space, $d\Phi_6$ is given by

$$d\Phi_6 = \delta^4(P - \sum_{i=1}^{6} p_i) \prod_{i=1}^{4q,\ell,\nu} \frac{d^3 p_i}{(2\pi)^3 2E_i} \tag{5.23}$$

and the product is over the outgoing 4 quarks (4q), a lepton (ℓ) and a neutrino (ν).

It is assumed that detector measure well lepton momenta and also angles of the jets, so the transfer function simplifies just to transfer function for jet energies $W_{jet}(\vec{x}, \vec{y})$. The transfer function is just normalized probability of observing a jet with E^{jet} when a outgoing parton with E^{parton} have been produced and have been obtained using simulated $t\bar{t}$ events.

Using above described procedure, the probability distribution for a given event is obtained. The final top mass estimate is performed by multiplication the probabilities for all events (obtaining likelihood) and the top mass which the likelihood takes maximal value for is considered as final top mass estimate, see Fig. 5.11.

---

[5] this is small difference to ME method in dilepton channel, where such integration have been left out. However, in last round of results using $1fb^{-1}$ it was implemented also in ME dilepton method.

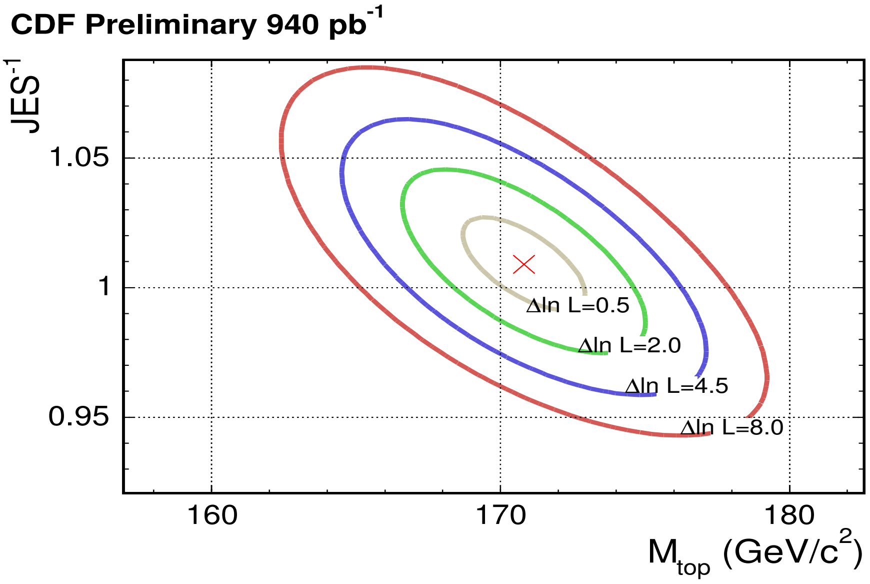


Figure 5.11: Likelihood contours at different distances from the minimum (corresponding to consecutive 1 σ distances in the Gaussian approximation) in the plane of M_{top} vs. JES factor.

There are, of course, some complications to above introduced picture. First, we perform *in situ* calibration of jet energy scale (JES) similarly to what template method in lepton + jet channel does. Therefore, the individual probability for a given event is dependent not only on input top mass but also on JES factor, $P = P(\vec{x}|M_{top}, JES)$. The second complication comes from fact that there are still some background processes which mimic the signature of signal $t\bar{t}$ events. Therefore, the final event probability can be written as

$$P(\vec{x}; M_{top}, JES) = C_s P_{t\bar{t}}(\vec{x}; M_{top}, JES) + (1 - C_s) P_{W+jets}(\vec{x}, JES), \tag{5.24}$$

where $0 \leq C_s \leq 1$ is the fraction of events consistent with our leading-order $t\bar{t}$ hypothesis and as a background process is assumed production of $W + jets$ events. The $P(\vec{x}; M_{top}, JES)$ gives a probability that, for a given event with input top quark mass M_{top} and jet energy scale factor JES [6], the measured variables will be equal to $\vec{x}$.

---

[6] the JES factor is defined as the factor by which is divided the jet energy obtained with default JES, i.e. increasing this factor will decrease the energy of the jet

DLM method

The DLM method used in this channel is almost identical not only in general principles, but also in the implementation to the DLM method used in dilepton channel, see Sec. 5.1.1.

There are only small differences due to the different kinematics. For a given event, a Monte Carlo integration is performed over the possible $t\bar{t}$ final state kinematics in the following way. We first generate a random value for a virtual mass squared of the W boson s_W in leptonic channel, according Breit-Wigner distribution. We identify the momentum of the electron or muon daughter with the measured value and the neutrino transverse momentum with the measured missing transverse energy. We then generate random values for the momenta of final state quarks according to the TF probabilities. We determine the z component of the neutrino momentum, with a twofold ambiguity, using s_W as a constraint.

Thus, for a given set of $\mathbf{x}$ and s_W, we fully determine the event kinematics.

The rest is totally same as in the DLM method used in dilepton channel. For each event, there is calculated the probability distribution as a function of the top mass (see Fig. 5.12) and the final top mass estimate is obtained as a maximum of the product of likelihoods for all data events. For the details see page 79.

5.1.3 All-hadronic channel

This channel is characterized by a 6 jet final state topology with a little missing energy in the event coming from unprecisely measured jet energies. The reconstruction of the top quark mass in this channel is challenging for two main reasons: the large ambiguities to reconstruct two top quarks out of 6 jets and a large amount of QCD multi-jet background of which the cross sections are badly known.

The number of possible jet combinations can be calculated following way. The total number of combinations of 6 jets is $6! = 720$. However, the combination which pairs jet_1 with jet_2 to form one W boson is equivalent (in a sense, that it obtains the parent W and top quark with the same 4-momentum) to combination of jet_2 with jet_1. Therefore the number of combinations should be decreased by a factor of two. The same is true for jets forming the second W boson [7] and we get another factor of two smaller number of combinations. Finally, it also does not matter or better said,

---

[7] usually we don't distinguish between W^+ and W^- and call both particles just "W boson". It is because in experiment, we usually don't know the charges of jets, and so we can not distinguish between W^+ and W^- when reconstructing them from pair of jets. However, it is possible to reconstruct charge of the jets, which is used in analysis measuring top quark charge, for example. Therefore, using such information would possible help to reduce even more the number of combinations

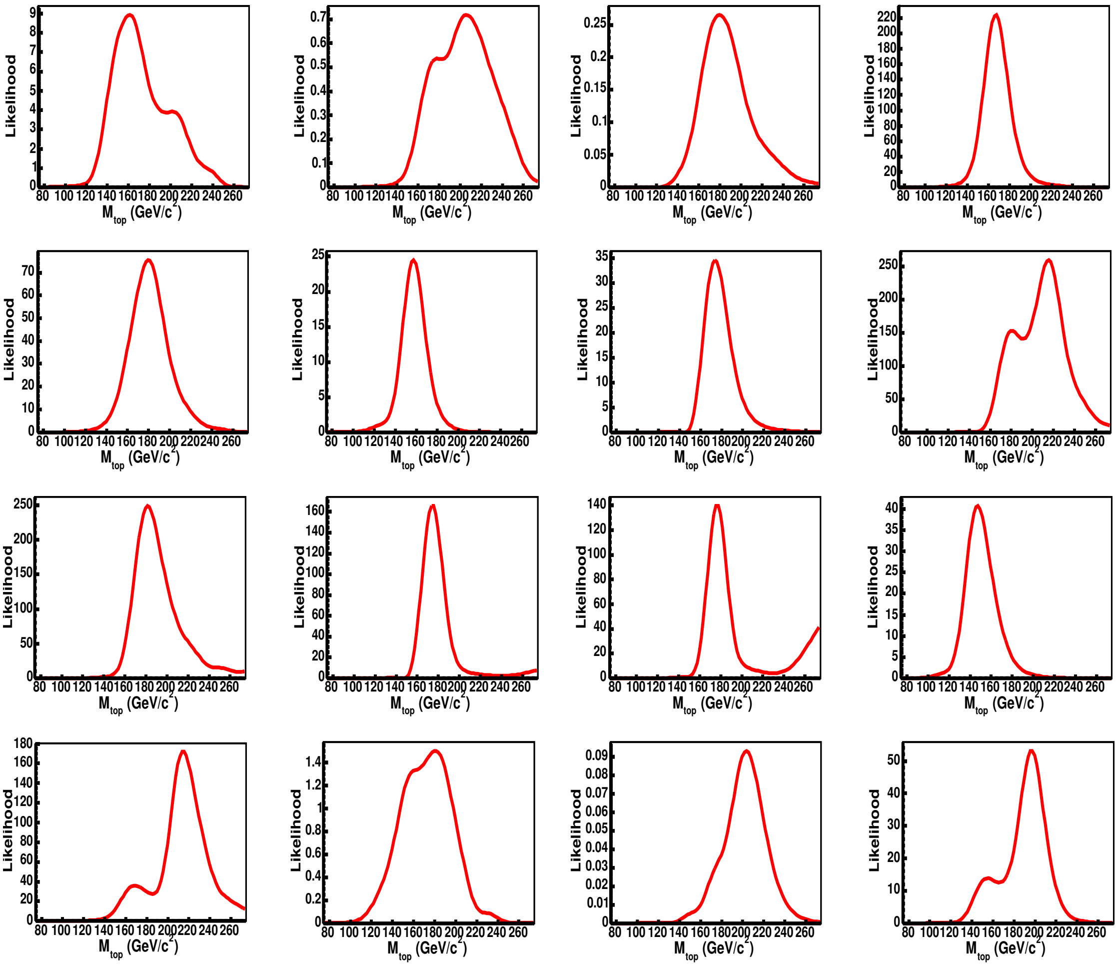


Figure 5.12: Event likelihood distributions as a function of top quark mass for double b-tagged events in lepton+jets data sample.

we are not able to distinguish, whether the first tree jets form top quark or anti-top due to the same reason as for W bosons and we suppose the mass of particle and its antiparticle is the same. Therefore there is final factor of two which lowers the number of combinations. So, in summary we have $720/(2*2*2) = 90$ combinations of how to form top and anti-top quarks from 6 final state jets.

When requiring that b-jets are identified with b-tagging method, the amount of ambiguities goes down comparing to 90 combinations when no requirement on b-tagging is applied. In case we require both b-jets are tagged in the event, the number of combinations of combining jets is lowered to only 6 possibilities. This can be calculated similar way as in previous case of no b-tagging requirement. The total number of combinations is now $2*4! = 48$ due to possible 4 non b-jets combination in forming two $W's$ and moreover two possible combinations of assigning particular b-jet to pair of non b-jets. However, identically as in case of no b-tag requirement, such number of combinations should be decreased by factor (2*2*2=8) due to reasons described above. Therefore there are left only 6 possible combinations. However, usually it is required that there is b-tagged only one jet and not necessarily both b-jets, because the b-tagging lowers the total number of events as the efficiency is not 100 % (e.g. SVX b-tagging algorithm has about 30 % efficiency). In case that there is one jet b-tagged, we have altogether 30 different combinations. This is calculated similarly to previous case with only one exception, that there is now known only one jet which is b-tagged and therefore there are 5! possibilities of how to assign remaining 5 jets. However, due to exactly same reasons as above, the total number of different combinations with respect to reconstructed top mass is 2*5!/(2*2*2) = 30 combinations.

There are two measurements at CDF in all-hadronic channel. One use template method, the other one use "ideogram method".

Template method

The procedure used here is the similar to other template methods in a sense that it produce top mass dependent quantity, which will be used to compare real data to simulated samples in order to extract the mass. However, there is a difference comparing to other template methods. It is in that the top mass dependent quantity (reconstructed top mass) is not the event-by-event measurement. The procedure is to append a mass to each b-tag that is found in the event surviving the selection.

In order to reduce the combinatorics, only the 6 highest E_T jets are used (even when by selection, the events can have up to 8 jets). Moreover, permutations which assign the tagged jet (at least one tagged jet is required by the selection) to the 4 jets forming the W's are rejected, thus there are only 30 possible combinations left.

Now, it is defined the χ^2 as a function of top mass. χ^2 is in principal the same as in template method in dilepton (Neutrino ϕ weighting method) and lepton+jet channel (see 5.8), just there are some terms related to leptons changed because there is no charged lepton and neutrino in all-hadronic final state. The χ^2 contains two terms which constrain the light quark jets to form the two W masses. Then a third jet momentum is added in order to form two objects which masses are closely spaced in the unknown mass. An additional term is added to account for the uncertainties in jet momentum measurements. The definition is thus as follows:

$$\chi^2(\widetilde{m_t},\widetilde{p_T^j}) = \frac{(m_{jj_1}-m_W)^2}{\Gamma_W^2} + \frac{(m_{jj_2}-m_W)^2}{\Gamma_W^2} + \frac{(m_{jjj_1}-\widetilde{m_t})^2}{\Gamma_t^2} + \frac{(m_{jjj_2}-\widetilde{m_t})^2}{\Gamma_t^2} + \sum_{j=1}^{N} \frac{(p_T^j-\widetilde{p_T^j})^2}{(\sigma_{p_T}^j)^2} \tag{5.25}$$

,where m_{jj} is invariant mass of di-jet,m_{jjj} is invariant mass of tri-jet, m_W, Γ_W are the PDG mass and width of W boson, Γ_t is the PDG width of top quark and p_T^j is the measured p_T of the j-th jet. Variables with tilde refer to the output of the minimization procedure. The quantity $\widetilde{m_t}$ is the fit parameter returned as the reconstructed top quark mass for the combination being considered.

Out of 30 combinations, the combination with the lowest χ^2 is selected and we get "tag" mass as a top mass output from minimization procedure. If there is more than one b-tag in the event, we run the kinematic fitter in iterative procedure using as input b-tagged jet all other tagged jets. This way, for each b-tagged jet is obtained one mass, and so we can end up with more top masses for one event.

The rest of the procedure is the same as for other template methods. The templates for different input top mass samples are created. They are approximated by some combination of smooth functions (in this case, it is combination of two Gaussian functions and one Gamma function for signal templates and two Gamma functions and one Gaussian for background template). Final top mass estimate is performed by likelihood minimization. The likelihood has the exact same form as in dilepton case (Eq.(5.1),(5.2),(5.3)) except that the total number of events is constrained the same way as a background, i.e. by Gaussian function instead of Poisson constrain in dilepton case.

Ideogram method

The basic idea of this method is to abandon analysis based on fitting the line shape to the global invariant mass spectrum, and to use instead event-by-event likelihoods describing the full ambiguity of the mass information in each event as correctly as possible.

The first ingredient of the method is the kinematic fit. The form of χ^2 is the same as in template hadronic method, Eq.(5.25), except there is no explicit condition on top mass being the same in both cases. Therefore, as the output of χ^2 minimization procedure there are obtained two fitted top quark masses m_1, m_2. All 90 independent jet permutations are evaluated and the χ^2 value is used as a combinatorial weight in the evens likelihood described below. The two-dimensional mass distributions for background Monte Carlo events $BG(m_1, m_2)$ and signal events that cannot be matched with the original partons coming from $t\bar{t}$ decay ($S_{NM}(m_1, m_2)$) are used as additional terms in the event likelihood. These distributions are shown in Fig. 5.14.

The second ingredient of the method is jet b-tagging. In the all-hadronic top quark analysis, it is crucial to identify the two jets, which are coming from the b-quarks. The tagging algorithm used in this analysis is called Jet Probability (JP) tagging algorithm. The algorithm makes use of the information of the tracks which are associated to the jet to determine the probability for this ensemble of the tracks to be consistent with coming out of the primary vertex of interaction. The probability distribution of set of tracks originating from the primary vertex is by construction uniformly distributed from 0 to 1. For a jet coming from heavy flavor hadronization, the distribution peaks at 0, due to tracks from long lived particles.

The likelihood ratio for a given jet i, with JP value, JP_i to be a b-quark jet can be expressed as

$$\mathcal{L}_i^b = \frac{P(b|JP_i)}{P(b|JP_i) + P(q|JP_i)}, \tag{5.26}$$

where $P(b|JP_i)$ is the probability of jet i being a b-quark jet and $P(q|JP_i)$ the probability of jet being a light quark jet, for a measured value of JP_i. Similarly, the likelihood ratio for jet j with JP value JP_j to e a light quark jet can be expressed as

$$\mathcal{L}_i^q = \frac{P(q|JP_i)}{P(b|JP_i) + P(q|JP_i)}, \tag{5.27}$$

Assuming a signal topology of six jets of which two are b-jets, each jet permutation, k, to reconstruct two top masses can be given a weight

$$w_{JP}^k = \prod_{i=1}^{2} \mathcal{L}_i^b \prod_{j=3}^{6} \mathcal{L}_i^q \tag{5.28}$$

The distributions of jet probability for light and heavy quarks together with JP weights for correct and wrong jet permutations in all-hadronic channel is in Fig. 5.13

The final event likelihood is constructed the following way. Each combination is weighted by its goodness-of-fit χ^2 value, obtained from the kinematic fit. The weight for each permutation is calculated as

$$w_{fit}^i = exp(-\frac{1}{2}\chi_i^2) \tag{5.29}$$

The total weight for each possible combination is then simply product

$$w_i = w_{fit}^i \cdot w_{JP}^i \tag{5.30}$$

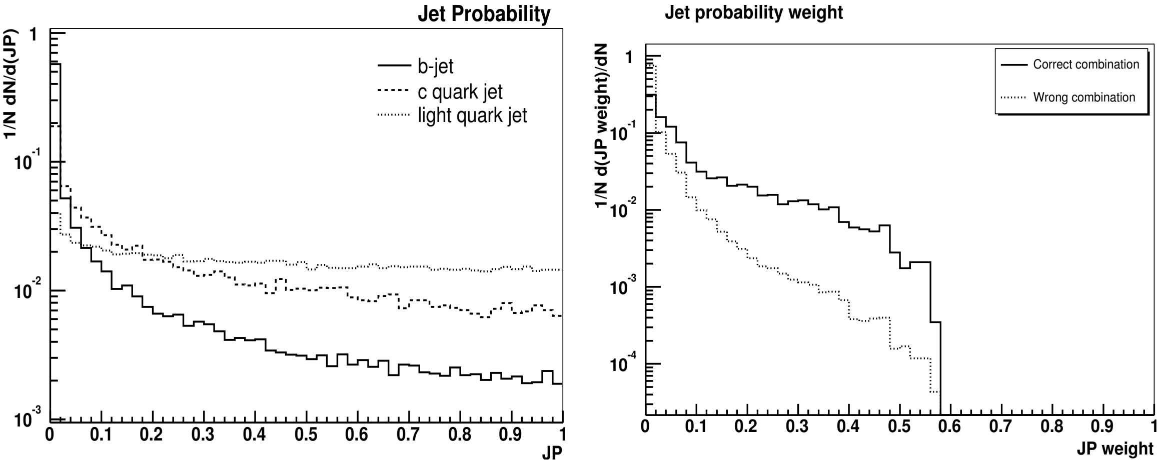


Figure 5.13: The jet probability for b-jets, c-quark jets, and for light quark jets (left). The probability (weight) for correct and incorrect jet permutations in all-hadronic channel (right).

Probability for the event to be a signal P_{evt} is used in the total event likelihood, which has then form

$$\begin{aligned}\mathcal{L}_{evt}(M_t, \mathcal{P}_{evt}) &= \mathcal{P}_{evt} \sum_{i=1}^{90} w_i \left[S_M(M_t, m_1^i, m_2^i) P_M + (1 - P_M) S_{NM}(m_1^i, m_2^i) \right] \\ &+ (1 - \mathcal{P}_{evt}) \sum_{i=1}^{90} w_i BG(m_1^i, m_2^i), \end{aligned} \quad (5.31)$$

where P_M represents the efficiency of matching 6 quark to 6 jets, S_M the signal likelihood for matched events and S_{NM} the signal likelihood for non-matched events. The matching efficiency P_M is calculated from signal MC to be 0.45, averaged over wide top mass range.

The likelihood for matched signal events is given by

$$\begin{aligned} S_M(M_t, m_1^i, m_2^i) &= \int \int_{M_t \pm 10GeV} G(m_1^i, m', \sigma_1^i) G(m_2^i, m'', \sigma_2^i) \\ & \qquad BW(m', M_t) BW(m'', M_t) dm' dm''. \end{aligned} \quad (5.32)$$

This represents the probability that the top masses m_1, m_2 will be measured, in case the true top quark mass was M_{top}. The term is just a convolution of Gaussian resolution functions $G(m_i, m', \sigma_i)$, describing the experimental resolution, with a Breit-Wigner $BW(m', M_t)$, representing the expected distribution of the average of the two invariant masses in the event, for a top mass M_t. The explicit form of these function is

$$G(m_i, m', \sigma_i) = \frac{1}{\sqrt{2\pi\sigma^2}} exp \left(- \frac{(m' - m_i)^2}{2\sigma_i^2} \right) \quad (5.33)$$

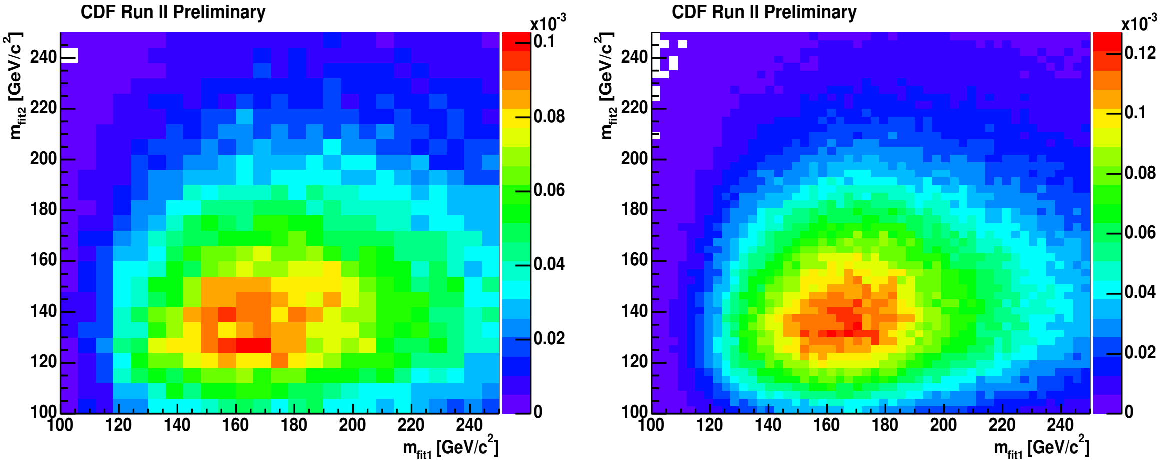


Figure 5.14: The 2D mass template for background (left) and signal non-matched events (right).

$$BW(m', M_t) = \frac{1}{\pi} \frac{M_t \Gamma}{(m'^2 - M_t^2)^2 + M_t^2 \Gamma^2} \tag{5.34}$$

The background term consists of the weighted sum, where $BG(m_i^1, m_i^2)$ is the 2D mass likelihood for background events, obtained from MC simulation.

The combined likelihood for a sample of N_{ev} events is just product of the likelihood for the individual events:

$$\mathcal{L}_{tot}(M_t, \mathcal{P}_{evt}) = \prod_{i=1}^{N_{ev}} \mathcal{L}^i_{evt}(M_t, \mathcal{P}_{evt}) \tag{5.35}$$

Maximizing this likelihood we obtain simultaneously estimate of top mass and probability that the event is signal event (and consequently also cross-section).

Ht template method in missing E_T + jets channel

This method is once again template method. It tries to exploit the correlation between m_{top} and scalar sum of transverse energies of all particles in the event, so called H_T. It's specialty is that it cannot be categorized into one of the standard channel (dilepton,l+j,all-hadronic), because it does not select events according number of charged leptons, rather according missing E_T. Therefore, passing this selection can events from dilepton channel ($t\bar{t} \to \tau\nu\tau\nu b\bar{b}$) and lepton+jets channel($t\bar{t} \to \tau\nu jjb\bar{b}$).

It have been demonstrated the possibility of isolating good purity $t\bar{t}$ sample by inclusively selecting events characterized by high and significant $\not{E}_T$. This results in a sensitivity to leptonic W decays independent on the lepton flavor, and in particular, in

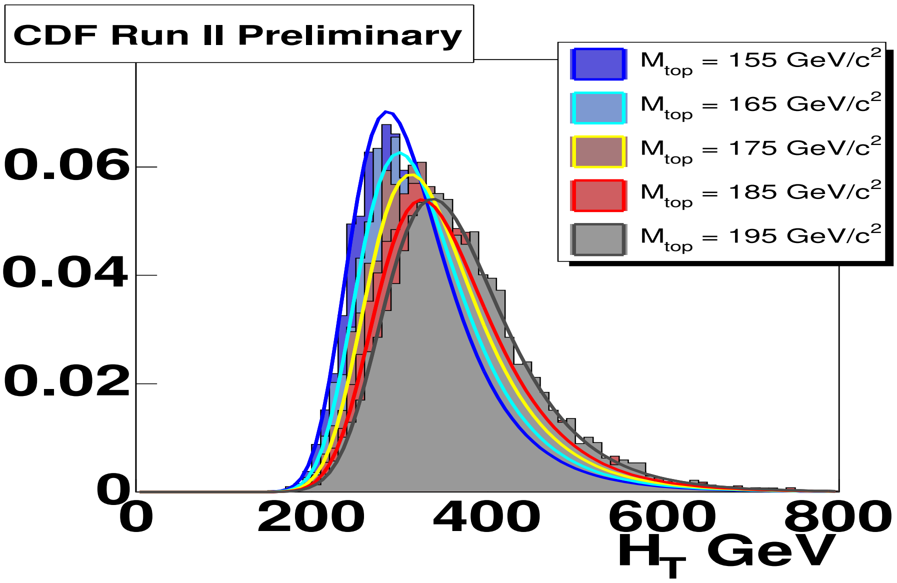


Figure 5.15: The expected signal H_T distributions as a function of the top quark mass.

a large acceptance with respect to $W \to \tau\nu$ decays (other analyzes are throwing away such cases).

Therefore, this channel provides complementary and independent results with respect to lepton-based measurements and is therefore of interest because such measurement has potential to improve the combined measurement significantly. Moreover, this actual measurement at CDF removed the events containing high-p_T electrons or muons , in order to increase the relative contribution of $W \to \tau\nu$ decays and to provide statistically independent sample with respect to lepton+jets analyzes. It's results are therefore statistically independent and orthogonal to other measurements.

Other specialty of this method (similar to the all-hadronic template method) is that it use b-tags as elementary pieces to build the templates as opposite to standard analyzes which use events as basic elements. It is because their background prediction is jet-based. This feature yields the necessity to build templates with one entry per each tagged jet, that is two identical values of H_T are filled in the case of double tagged event. The net effect is that the uncertainties on the top mass are a little bit underestimated when checking the method by pseudo-experiments and therefore they needed to be corrected in final mass estimate from data events.

The rest of the procedure is pretty much the same as for other template methods. H_T templates for different input top quark masses are constructed using simulated events, see Fig. 5.15. These signal templates are approximated by smooth function (in this case it was used Pearson type IV function) by simultaneous fit to all the templates. The background template is also approximated by smooth function (in this case it is combination of two Gaussians and one gamma function).

Final top mass estimate is performed by comparing the data events to MC simulated templates using the un-binned likelihood fit with exactly same formula as is used in dilepton template methods, see Eq.(5.1), Eq.(5.2), Eq.(5.3).

5.2 Discussion about the measurement methods

As we saw in previous section, there are many methods of top mass measurement at CDF in many different channels. The actual CDF Run II results are shown in Fig. 5.16. Basically, there are two different approaches to the top quark mass measurement: template methods and matrix-element based methods. In ME methods, there is used more initial information about the $t\bar{t}$ event than in template methods, because they use information about production mechanism of $t\bar{t}$ (matrix element $\mathcal{M}_{t\bar{t}}(M_t)$). Therefore, it is expected that the ME methods should provide better statistical uncertainty in top mass than template methods.

Moreover, not all the template methods give the similar uncertainties on top quark mass. Usually, the most sensitive variable to true (input) top mass is reconstructed top mass. Naively, this is expected, since in reconstructed top mass are included all the kinematic information about the event. The simple template methods which use just some particular information about the $t\bar{t}$ event (e.g. L_{xy} distance, p_T of leptons, etc.) are up to now statistically limited. However, they have also their advantages. Because they don't use all the information from event, they are also not sensitive to systematic effects related to such informations (e.g. L_{xy}, p_T of leptons are not sensitive to jet energy scale uncertainty, which is the largest systematic uncertainty at present for other methods). The systematic uncertainty on top mass already started to dominate over the statistical error at Tevatron. This will be the case even more at LHC. Therefore, in future, these simple template methods can pay off, since they will be immune to more systematic effects than standard template methods.

In template methods, a special case is the dilepton channel, where is needed to be introduced at least one additional variable in order to kinematically solve the system. There are more possibilities how to choose such a variable. Up to now, three different kinematic variables were used: $p_z^{t\bar{t}}$, η_ν,ϕ_ν. The methods which use these variables give very consistent results and of about the same sensitivity.

Having as much as possible different top mass measurements has a few advantages. First, the more measurements can be combined into one which will have smaller un-

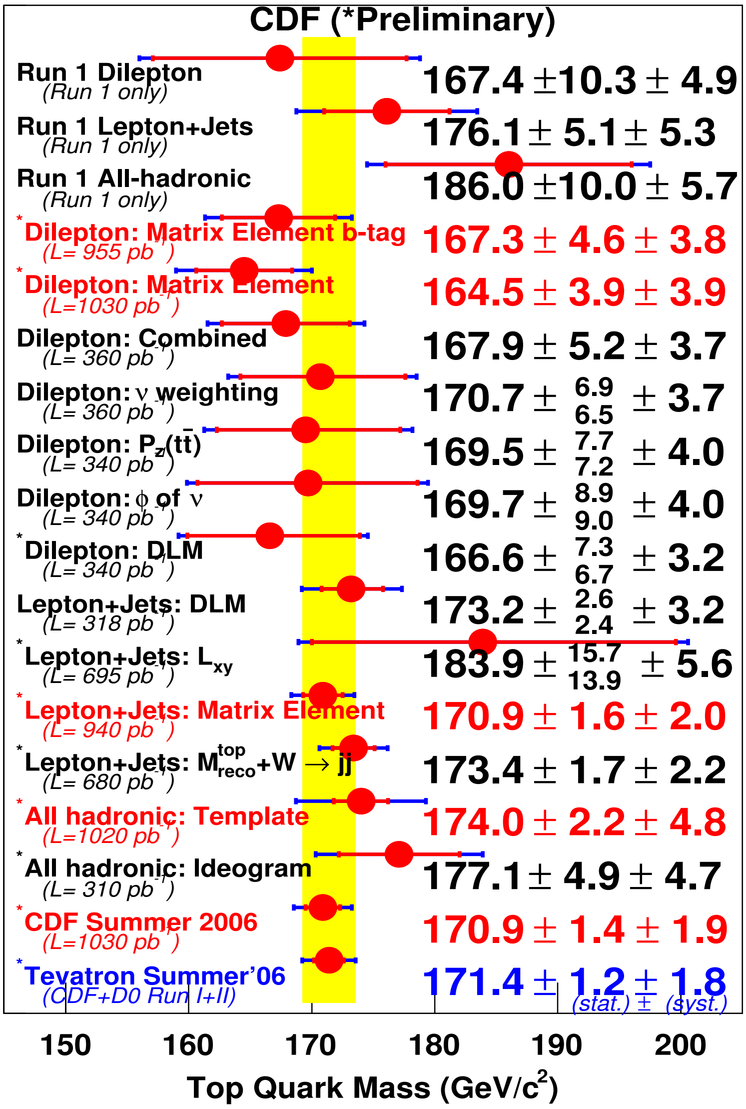


Figure 5.16: All of the CDF Run II top mass measurements. Also Run I results and current world average are shown.

certainty than the individual measurements. Second, in the case, the measurements across different decay channels would significantly differ, this would be hint for physics beyond the Standard Model. Third, even having a few similar methods in the same decay channel is useful, as it can serve as a cross-check.

5.3 Future top quark mass measurement methods

There are two different colliders expected to be run in near future: Large Hadronic Collider (LHC) where protons will collide with protons and International Linear Collider (ILC) which should be electron-positron collider.

We will describe the methods which have potential to produce the most precise measurements of the top quark mass at these colliders. The possible measurement methods are the same for both colliders except for one method possible only at electron-positron collider. This method should produce the best measurement of the top quark mass.

5.3.1 LHC

The Large Hadronic Collider (LHC) is already built and the data taking is expected to start in middle of year 2007. As already mentioned above, it will be proton-proton collider with the central-of-mass energy $\sqrt{s} = 14$ TeV (comparing to Tevatron proton-antiproton collider with $\sqrt{s} = 1.96$ TeV).

LHC will be "top factory". The top quark will be produced mainly in pairs through process $gg \to t\bar{t}$ (90 %) and $q\bar{q} \to t\bar{t}$ (10 %). The corresponding cross-section, at the next-to-leading order, is equal to 833 pb [80]. Therefore, it is expected roughly 8 million $t\bar{t}$ pairs to be produced with 100 days at low luminosity ($10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$) which corresponds to 10 fb^{-1}.

There are expected many methods for top mass reconstruction to be used. I will mention just a few of them which seems will produce the best results. Lepton+jets channel seems to be most promising channel.

In the first method, the top mass will be estimated from the **reconstruction of the invariant mass of three-jet system**: the two light jets from the W and one of the two b-jets. The determination of the combination of three-jets proceeds in two steps: the choice of the two light jets, and the choice of the b-jet associated to the reconstructed hadronic W. The choice of two light jets is performed according minimization of χ^2. Moreover, there is calibrated the energy of the jets in-situ. The expression of χ^2, given by equation 5.36, is the sum of three terms: the first one corresponds to the constrain of the jet pair invariant mass m_{jj} to the PDG W mass (m_W), the others correspond to the jet energy correction factors, $\alpha_i (i = 1, 2)$, to be determined by this minimization:

$$\chi^2 = \frac{(m_{jj}(\alpha_1, \alpha_2) - m_w)^2}{\Gamma_W^2} + \frac{(E_{j1}(1-\alpha_1))^2}{\sigma_1^2} + \frac{(E_{j2}(1-\alpha_2))^2}{\sigma_2^2} \qquad (5.36)$$

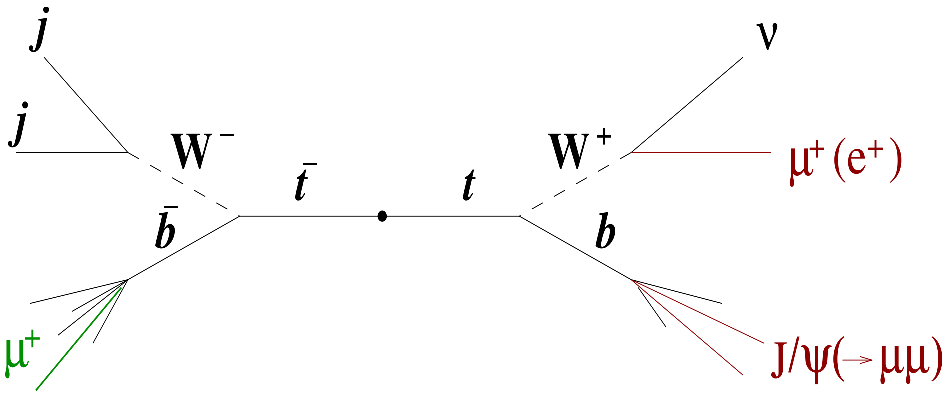


Figure 5.17: Diagram of the $t\bar{t}$ decay to semi-leptonic final state with J/Ψ.

The χ^2 is minimized , event by event, for each light pair jet. The light pair $j_1 j_2$ corresponding to the minimal χ^2 is kept as the hadronic W candidate.

The corresponding b-jet among two possible candidates is the one leading to highest P_T for the top quark.

In the second method, top mass will be measured using **kinematic fit**. Here will be reconstructed entire $t\bar{t}$ state, in order to reduce the systematic error due to FSR. The hadronic part ($W \rightarrow jj$) is reconstructed in a similar way to the previous method. The leptonic side ($W \rightarrow \ell\nu$) can not be directly reconstructed due to the presence of the undetected neutrino, but can be estimated in three steps:

- $p_T(\nu) = E_T^{miss}$
- $p_z(\nu)$ is obtained by constraining the invariant mass of the lepton-neutrino system to the PDG W mass value. This kinematic equation leads to two $p_z(\nu)$ solutions
- the remaining b-jet is associated to the reconstructed W

The top mass determination is performed through a kinematic fit, relying on a χ^2 based on mass constraints ($m_{jj} = m_W^{PDG} = m_{\ell\nu}$; $m_{jjb} = m_{\ell\nu b}$) and kinematic constraints (energy and direction of leptons and jets can vary within their resolutions). The minimization of this χ^2 is performed event by event, for the two solutions: the one giving the lower χ^2 is kept. The final top mass estimate is determined as the linear approximation of $m_{top}(\chi^2)$ for $\chi^2 = 0$.

The last method I will mention is the measurement in lepton+jets channel where a J/Ψ arises from the b-quark associated to the leptonic decaying W (Fig. 5.17).

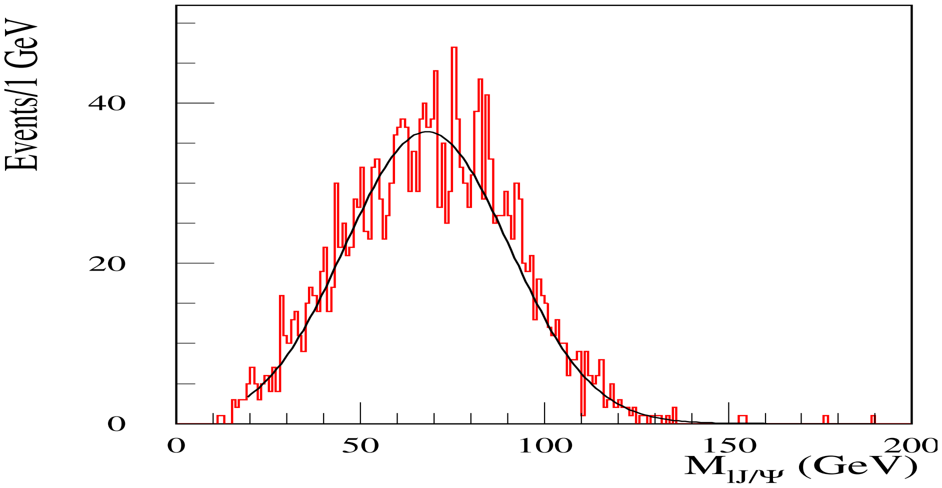


Figure 5.18: Lepton – J/Ψ invariant mass distribution.

The large mass of J/Ψ induces strong correlation with the top mass, as will be shown below. To solve ambiguity between the two b-jets , a charge tagging of the b decaying to J/Ψ is applied, requiring a muon of the same electric charge as the isolated lepton in the other b-jet. The overall branching ratio ($5.3 \cdot 10^{-5}$) is low, therefore, this analysis will be performed at high luminosity (10^{34}cm^{-2}s^{-1}). Four-lepton events (≥ 3 muons) are selected with the invariant mass of two muons being consistent with the J/Ψ mass. This cuts reduce background to a negligible contribution.

An example of the $\ell - J/\Psi$ mass distribution and the correlation of the peak of the distribution to the top quark mass is shown in Fig. 5.18 and Fig. 5.19, respectively. The correlation is linear. Since the method just calculates the invariant mass of leptons, it will be not sensitive to jet energy scale, the main source of systematic uncertainty in direct top mass measurement. Another advantage of this method is that identification of J/ψ provides a much cleaner signal than in other methods using leptons.

The comparison between different methods of top mass reconstruction is in Tab. 5.1. It can be seen there that with more data accumulated, the statistical error will be negligible and dominant will be systematic error. Therefore, it is expected that method using J/Ψ in the final state should eventually give the best resolution on top quark mass from all the methods at LHC ($\Delta m_{top} \sim 0.8$ GeV).

5.3.2 ILC

The top quark mass can be measured in e^+e^- collisions by reconstructing top quark decay products in much the same way as at the LHC or Tevatron. In addition, there

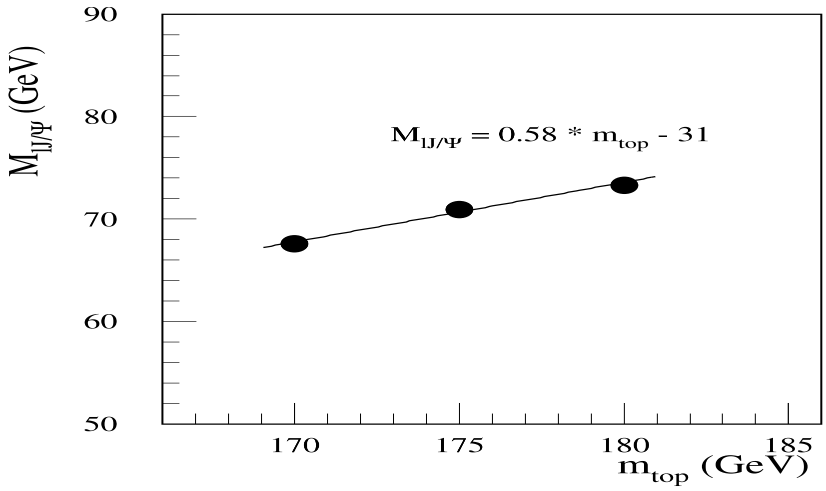


Figure 5.19: Correlation between lepton – J/Ψ invariant mass and the top quark mass.

	3-jets mass	Kinematic fit	J/Ψ
num. of events (10 fb^{-1})	64000	26000	30
statis. uncert.	0.05	0.1	1.0
syst. uncert.	1.3	0.9	< 0.8

Table 5.1: The number of events, expected systematic and statistic uncertaninties on top quark mass measurements, in the lepton + jets channel, for the tree methods described above. The statistic and systematic errors for method with J/Ψ in final state corresponds to one year of running at high luminosity ($10^{34}\text{cm}^{-2}\text{s}^{-1}$).

exists the unique possibility of determining the mass in pair production near threshold.

The ILC is expected be electron-positron collider. Regarding to top quark mass measurement, it will have one big advantage comparing to LHC proton-proton collider. At hadron colliders, we actually don't know the initial energy of interaction, because we don't know the energy of interacting partons coming from protons. Partons can have any energy from zero up to energy of proton. However, at $e^+ - e^-$ collider, the electrons can interact directly, so we know exactly (better said, with high precision) what is the incident energy of interaction. Therefore, at $e^+ - e^-$ collider, we can perform scan near the threshold energy for top quark pair production. This scanning can be used for measurement of the top quark mass.

The idea of the method which use the energy scan near the threshold is as follows. For small quark masses, the long time which the particles remain close to each other, allows the strong interactions to build up rich structures of bound states and resonances.

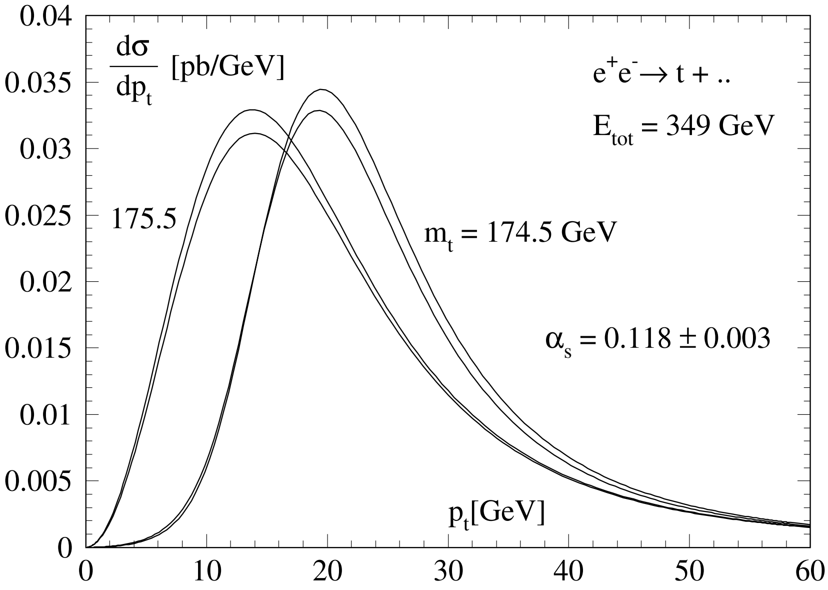


Figure 5.20: The momentum spectrum of the top quarks near the threshold for a fixed total center-of-mass energy. The momentum depends strongly on the top mass, yet less on the QCD coupling.

For the large mass, the picture is different. The topponium resonances can no longer be formed. Despite their transitory existence, the remnants of the toponium resonances nevertheless induce a fast rise of the cross section near the threshold, see Fig. 5.21. The steep rise provides the best basis for high-precision measurements of the top quark mass. Detailed simulations at $m_t \approx 175$ GeV predict sensitivity $\delta m_t \approx 0.15$ GeV to the top mass when the measurements of the excitation curve (see Fig. 5.21) and the t momentum spectrum (see Fig. 5.20) are combined.

However, it was shown in [81] that even by direct kinematic reconstruction of dilepton events, the similar precision of about 200 MeV can be obtained. Such method would be insensitive to uncertainties in luminosity measurement (as opposite to above mentioned 'threshold scan' method).

Here, it should be noted that although one speaks about "the" top quark mass, one should keep in mind that the concept of quark mass is convention-dependent. The definition of top quark pole mass is often implicit, because the measuring of "the top quark mass" from the products of top quark is considered to be measurement of top quark pole mass. However, it can be useful to choose another convention (for more detailed discussion of top mass definitions, conventions, etc. see Sec. 2.2.3).

The "energy scan method" described here is considered to be the most accurate method with potential uncertainty of $\delta m_t \approx 0.15$ GeV which can be achieved for the

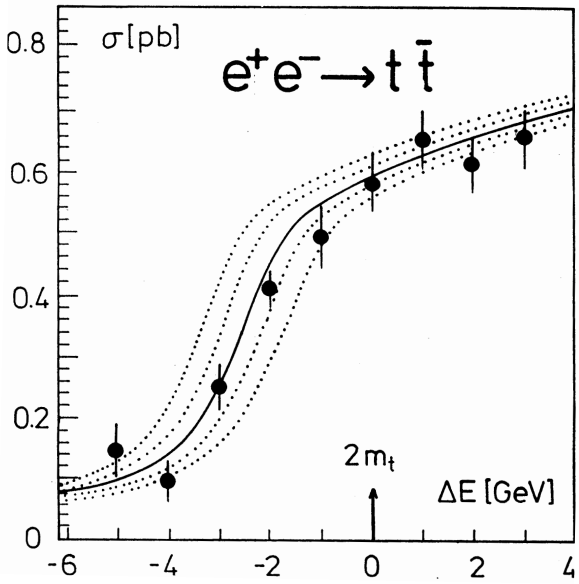


Figure 5.21: Total cross-section of $t\bar{t}$ production. The errors of the data points correspond to an integrated luminosity of 50 fb^{-1}. The dotted curves indicate shifts of the top mass by 200 and 400 MeV.

top quark $\overline{MS}$ mass. This is a factor two improvement compared to the accuracy that could be achieved for the top quark pole mass. However, under LHC conditions the experimental systematic uncertainty discussed above in the previous section is the limiting factor in practice. A potential exception is the measurement with J/Ψ in final state, since there the systematic uncertainty is to a large extent theoretical.

Chapter 6

Template Kinematic Method

In this chapter we will deeply describe the reconstruction method which we will use for top quark mass measurement.

Our method for measuring the top quark mass[1] is considered to be kinematic method. It's because we use kinematic variables (4-momenta of final state particles + one kinematic variable which we make assumption about) in order to kinematically[2] reconstruct top quark mass. There exists also non-kinematic methods for top mass measurement. For detailed discussion, see Sec. 5.1, 5.2.

Due to various reasons which will be described later, the reconstructed mass is biased from original true mass. However, it is clearly very sensitive variable to the original top mass, so we can still use such variable for estimation of top quark mass. We compare the reconstructed top quark mass between simulated and data events. We do this by producing the mass distributions, templates, for simulated events and then comparing data to these templates. Hence, the other part of methods name. For general description of template methods, see Sec. 5.1.

We will in detail describe the principles and the practical use of the method in the following sections of this chapter.

6.1 Principles of the method

As it was already mentioned in previous chapter, we use $t\bar{t}$ dilepton decay channel (see Sec. 2.2.2) for our top mass measurement.

Even the dilepton channel has the smallest branching ratio comparing to other channels, it is the cleanest channel, i.e. it has highest signal to background ratio, see

---

[1] from now on, we will often refer to the mass of the top quark as “top mass”.

[2] i.e. using law of 4-momentum conservation

Tab. 2.1. It is important to measure the top mass in all channels. First, such measurements provide independent (at least, partially) estimates of top quark mass which can be combined in order to improve the overall uncertainty. Moreover, measurements across the different channels serve also as a cross-checks. If there would be significant difference between measured mass in different channels, this would mean that in some channel we use also some events, from up to now unknown sources, and this could point us to some new physics beyond the Standard Model.

Dilepton channel is special comparing to other channels, in a sense that it does not have enough measured kinematic characteristics. Consequently, we are not able to kinematically solve the system, we say the system is under-constrained. This is due to fact that in final state there are two neutrinos which we are not able to detect. However, there is partial information about escaped neutrinos which can be obtained from the imbalance of the energy flow in the transverse plane. That is missing transverse energy $\not{E}_T$, i.e. its two components: $\not{E}_{T_x}, \not{E}_{T_y}$ (for definitions, see section 3.2.1). Other than that, we detect charged leptons and jets coming from b-quarks (there could be more jets coming from ISR and/or FSR). Therefore, we have only 18 constraints from measurement – four-momenta of two charged leptons and two jets plus two components of $\not{E}_T$ [3]. Moreover, we can make an assumption about the mass of the neutrinos ($m_{\nu_1} = m_{\nu_2} = 0$) [4], mass of W bosons (PDG value $m_{W^\pm} = 80.4$ GeV) [21] and about equality between top and anti-top masses ($m_t = m_{\bar{t}}$, according Standard Model).

In total, we end up with 23 constraints. However, we need at least 24 to kinematically fully describe system with six final state particles (four-momenta for 2 leptons, 2 jets, 2 neutrinos).

The simplest possible way how to overcome the above problem, is to introduce one variable, which we will make assumption about. However, some other methods, see Sec. 5.1.1, make assumptions about 2 variables, e.g. pseudo-rapidities or azimuthal angles of both neutrinos. This way, they obtain over-constrained system and from the more solutions is chosen the "best" one according certain criteria or the solutions are weighted.

In our method, we introduce a constrain on **longitudinal momentum of $t\bar{t}$ system** $p_z^{t\bar{t}}$. Basically, this variable is just the replacement for the unmeasured longitudinal component of neutrinos momenta. However, it includes in itself not only neutrinos momenta but also momenta of all other particles in event (leptons, jets)

---

[3] for b-jets, we assume $m_b = 5$ GeV

[4] even there exists evidence for non-zero neutrino masses [21], in our case, we can neglect them because typical values of E_ν in $t\bar{t}$ events are a few tens of GeV which is much more than possible neutrino masses $m_\nu^\ell, \ell \in \{e, \mu, \tau\}$.

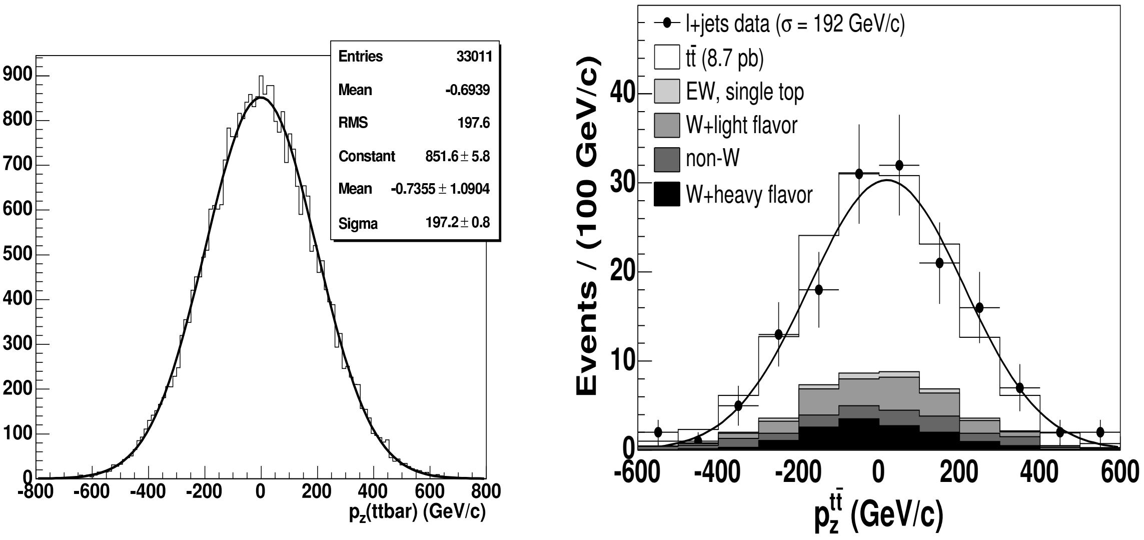


Figure 6.1: Longitudinal momentum of $t\bar{t}$ system $p_z^{t\bar{t}}$ for $t\bar{t}$ simulated events, with generated top mass 175 GeV, passing dilepton selection criteria (left). On right plot, it is shown $p_z^{t\bar{t}}$ for Monte Carlo simulated events and the data from lepton+jets channel.

$p_z^{t\bar{t}} = p_z^t + p_z^{\bar{t}} = \sum_\ell p_z^\ell + \sum_{jets} p_z^{jet} + \sum_{\nu,\bar{\nu}} p_z^\nu$. Ideally, the quantity $p^{t\bar{t}_z}$ should be determined theoretically, e.g. by simulation and should be virtually independent on top quark mass. In first order, the $p_z^{t\bar{t}}$ is given by parton distribution functions (PDFs) of proton (and anti-proton) and it is indeed independent on top quark mass.

Studies from Monte Carlo simulation over wide range of top masses from 150 – 200 GeV show that $p_z^{t\bar{t}}$ has Gaussian behavior, with a mean of zero and width around 195 GeV, see Fig. 6.1 and Tab. 6.1. This width increases by roughly 5 % across the top quark mass range studied. The validity of our Monte Carlo simulation can be tested with data from lepton+jets $t\bar{t}$ events, where $p_z^{t\bar{t}}$ can be reconstructed explicitly. Fig. 6.1 compares $p_z^{t\bar{t}}$ from the lepton+jets CDF data sample with $t\bar{t}$ and background Monte Carlo samples, showing good agreement between data and simulation (Kolmogorov-Smirnov probability is 0.85). The lepton+jets event selection, using secondary vertex b-quark identification, and subsequent backgrounds are similar to those of the lepton+jets cross section measurement [82].

So, finally we can write the following set of 24 equations which can be used to

generated top mass / GeV	width of $p_z^{t\bar{t}}$ distribution / GeV
150	192.5 ± 2.5
160	189.0 ± 2.5
170	194.4 ± 1.5
175	197.2 ± 0.8
180	198.8 ± 1.5
190	200.5 ± 2.5
200	200.7 ± 2.5

Table 6.1: Uncertainties in $p_z^{t\bar{t}}$ for events passing dilepton event selection. Uncertainties are taken from the Gaussian fits to the distributions.

kinematically constrain the $t\bar{t}$ dilepton system:

$$\begin{aligned}
p_{W^+} &= p_{l^+} + p_\nu \\
p_{W^-} &= p_{l^-} + p_{\bar{\nu}} \\
p_t &= p_b + p_{W^+} \\
p_{\bar{t}} &= p_{\bar{b}} + p_{W^-} \\
p_x^\nu + p_x^{\bar{\nu}} &= \not{E}_{T_x} \\
p_y^\nu + p_y^{\bar{\nu}} &= \not{E}_{T_y} \\
m_{W^\pm} &= 80.4 \text{ GeV} \\
m_{\nu,\bar{\nu}} &= 0 \text{ GeV} \\
m_t &= m_{\bar{t}} \\
p_z^t + p_z^{\bar{t}} &= p_z^{t\bar{t}},
\end{aligned} \tag{6.1}$$

where p_i means four-momentum of particle i, while p_k^i means k-th component of momentum of particle i. On right side of these equations, there are the measured variables or the values and variables we make assumption about[5].

The set of equations (6.1) can be narrowed to the following 3 equations by squaring first two equations and using the fact that $p^2 = m^2$, and by inserting first two equations into 3rd and 4th equation, squaring these equations, using $p^2 = m^2$ and using the assumption about equality of top and anti-top masses:

$$\begin{aligned}
(E_{l^+} + E_\nu)^2 - (\vec{p}_{l^+} + \vec{p}_\nu)^2 - m_W^2 &= 0 \\
(E_{l^-} + E_{\bar{\nu}})^2 - (\vec{p}_{l^-} + \vec{p}_{\bar{\nu}})^2 - m_W^2 &= 0 \\
(E_{l^+} + E_\nu + E_b)^2 - (\vec{p}_{l^+} + \vec{p}_\nu + \vec{p}_b)^2 & \\
-(E_{l^-} + E_{\bar{\nu}} + E_{\bar{b}})^2 + (\vec{p}_{l^-} + \vec{p}_{\bar{\nu}} + \vec{p}_{\bar{b}})^2 &= 0
\end{aligned} \tag{6.2}$$

The unknowns are just 3 components of one neutrino momentum vector $\vec{p}_\nu$. E_ν is given by $\vec{p}_\nu$ as $E_\nu = |\vec{p}_\nu| = \sqrt{p_x^2 + p_y^2 + p_z^2}$. The $\vec{p}_{\bar{\nu}}$ can be expressed by $\vec{p}_\nu$ and known measured quantities: $p_x^{\bar{\nu}} = \not{E}_{T_x} - p_x^\nu$, $p_y^{\bar{\nu}} = \not{E}_{T_y} - p_y^\nu$, $p_z^{\bar{\nu}} = p_z^{t\bar{t}} - \sum_{leptons+jets} p_z - p_z^\nu$.

---

[5] except $p_{W^\pm}$, which will be expressed by lepton and neutrino four-momenta.

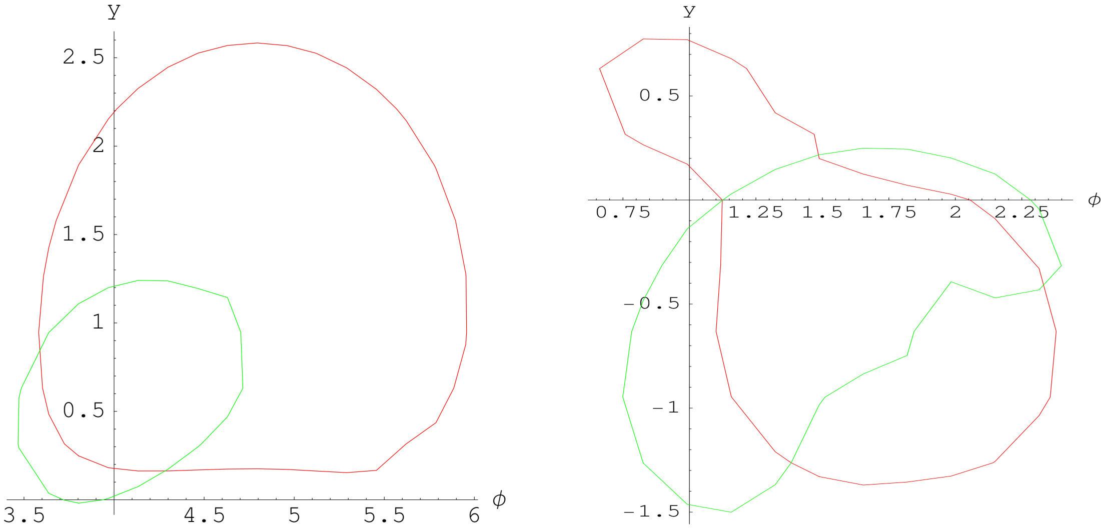


Figure 6.2: Possible neutrino solutions in plane of azimuthal angle ϕ and rapidity y (we assume $m_\nu = 0$, so rapidity y is the same as pseudo-rapidity η). Two different events are shown, one with 2 neutrino kinematic solutions (left) and one with 4 solutions (right).

It can be easily seen, that first two equations in (6.2) express the equality of squared mass of $W^\pm$ and squared invariant mass of its decay products and the last equation says that squares of invariant masses of top and anti-top decay products are equal.

Solving set of equations (6.2), the neutrino vector momentum $\vec{p}_\nu$ is obtained, the full kinematic chain can be reconstructed and the mass of top and anti-top quark can be calculated.

The set of equations (6.2) has up to four solutions, see Fig. 6.2. However, usually only up to two solutions are found. We checked with events on parton level, that 4 solutions exist only in about 3 % of all events. In the Fig. 6.2, there are plotted different neutrino solutions. The neutrino momentum (p_x, p_y, p_z) can be expressed as a function of p_T, azimuthal angle ϕ and pseudo-rapidity η. Using first equation from the set of equations (6.2) to express p_T as a function of ϕ and η, one ends up with two equations for two unknowns ϕ and η. The two curves in Fig. 6.2 correspond to each particular equation. The neutrino solutions corresponds to points where both curves cross each other.

To solve a set of nonlinear equations, we applied Newton's numerical method [83]. In general case, for the equations (in our case, we have 3 equations for 3 unknowns p_x^ν,

p_y^ν, p_z^ν)

$$\left.\begin{array}{lcl} F_1(x_1,x_2,...,x_n) & = & 0 \\ F_2(x_1,x_2,...,x_n) & = & 0 \\ ... & & \\ F_n(x_1,x_2,...,x_n) & = & 0 \end{array}\right\} \mathbf{F}(\mathbf{x}) = 0 \tag{6.3}$$

the following iterative procedure is applied (in matrix notation):

$$\begin{array}{lcl} \mathbf{x} & = & \mathbf{x}_0 \\ \mathbf{A}.\Delta\mathbf{x} & = & -\mathbf{F}(\mathbf{x}) \\ \mathbf{x} & = & \mathbf{x}_0 + \Delta\mathbf{x} \end{array} \tag{6.4}$$

where

$$A_{ij} = \frac{\delta F_i(\mathbf{x})}{\delta x_j}\Bigg|_{\mathbf{x}=\mathbf{x_0}}$$

Starting from some initial vector $\mathbf{x} = \mathbf{x_0}$, the correction $\Delta\mathbf{x}$ is calculated. The corrected $\mathbf{x}$ is used for the calculation of another correction in next iteration until a satisfactory solution is found. As the starting vector $\mathbf{x_0}$, we use vector with components $(0.1, 0.1, 0.1)$ GeV.

We consider the solution to be satisfactory when at least one of these three conditions is met:

- $max\ |F_i(\mathbf{x})| \leq \epsilon_1$ (by default, we use $\epsilon_1 = 1$). *max* is the function which takes the maximal value from the given set of values. Note, such precision in solution corresponds to very, very good precision in top mass ($\sim$ 3 MeV), so we don't expect any impact of such value of ϵ_1 on measured top mass.

- $max|\Delta x_i| \leq \epsilon_2$ (by default, $\epsilon_2 = 0.1$) and $max\ |F_i(\mathbf{x})| \leq 10\epsilon_1$

- if procedure does not converge according the two above criteria after ϵ_4 steps ($\epsilon_4 = 300$), the closest solution found in iteration procedure should have $max\ |F_i(\mathbf{x})| \leq \epsilon_5$ (by default, $\epsilon_5 = 500$). This precision corresponds to difference of m_W from PDG value (80.4 GeV) by about 3 GeV or to the difference between top and anti-top masses $\sim$ 2 GeV.

  We performed simple simulation, where the mass of the particle was randomly generated according Breit-Wigner distribution with corresponding mean and width (M_W = 80.4 GeV, Γ_W = 2.1 GeV, M_{top} = 170 GeV, Γ_{top} = 1.5 GeV), see Fig. 6.3. The above mentioned mass differences (3 GeV and 2 GeV, respectively) are fulfilled in $\sim$ 65% of cases for top – anti-top mass difference and in about 82 % cases for generated W mass shift from PDG value.

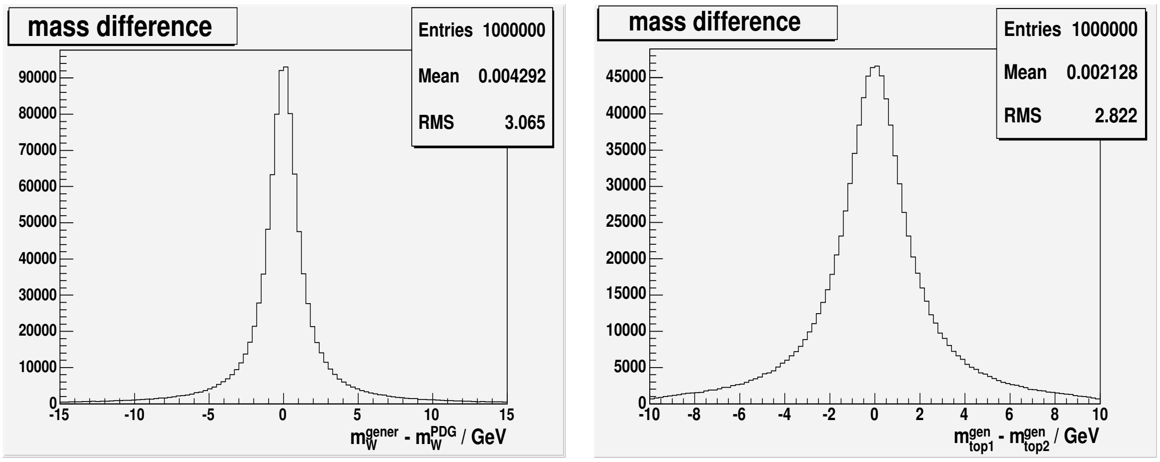


Figure 6.3: Difference between randomly generated W mass according Breit-Wigner distribution and PDG value of W mass (left). Difference between two Breit-Wigner generated top masses (right).

Usually it takes a few iteration steps to find a satisfactory solution. However, if the solution was not found after $\epsilon_6 = 10$ steps in iteration process, there is performed reinitialization of starting vector $\mathbf{x_0}$ (each component of the vector is randomly generated within range $(-100, 100)$) and new iteration process starts. However, the steps from previous initializations are counted into variable which is checked whether the maximum number of steps in iteration (ϵ_4) was not achieved.

We investigated the influence of above mentioned numerical parameters on top quark mass. However, by changing the parameters within some reasonable range (by factor 2, 4), we did not see any improvement with respect to top mass resolution.

After we find solution for p_ν, we can reconstruct m_t and $m_{\bar{t}}$. As the top mass estimate for given event we then use the average $m_{top} = (m_t + m_{\bar{t}})/2$.

As was stated above, when one solution is found, there should be also another one. The other solution $\vec{p_\nu}\,'$

$$\vec{p_\nu}\,' = \vec{p_\nu} + \vec{X} \tag{6.5}$$

can be found by applying Newton's method to the following set of equations:

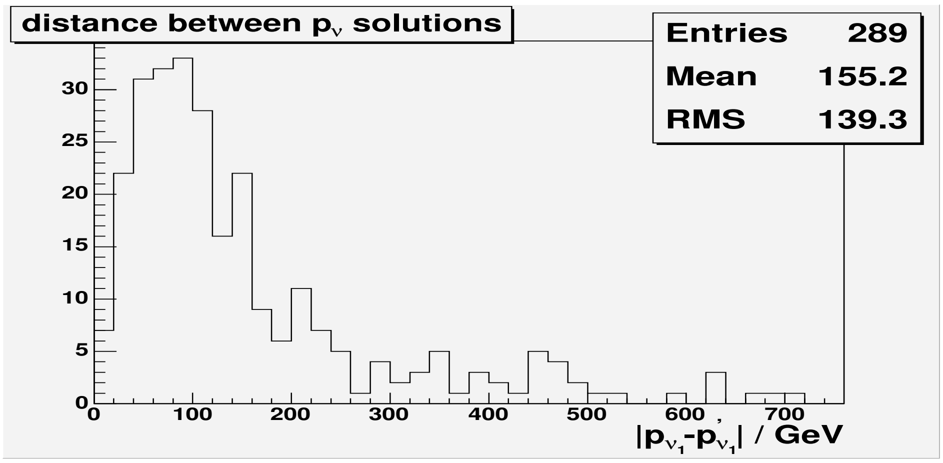


Figure 6.4: The distance $|\vec{p}_\nu - \vec{p}_\nu{}'|$ between neutrinos momenta for different kinematic solutions for a given event at generator level.

$$
\begin{aligned}
f_1(X_1,X_2,X_3) &\equiv \sqrt{m_W^2 + (\vec{p}_{W^+} + \vec{X})^2} - E_{l^+} - \sqrt{(\vec{p}_\nu + \vec{X})^2} \\
f_2(X_1,X_2,X_3) &\equiv \sqrt{m_W^2 + (\vec{p}_{W^-} - \vec{X})^2} - E_{l^-} - \sqrt{(\vec{p}_{\bar{\nu}} - \vec{X})^2} \\
f_3(X_1,X_2,X_3) &\equiv \sqrt{(\sqrt{m_W^2 + (\vec{p}_{W^+} + \vec{X})^2} + E_b)^2 - (\vec{p}_t + \vec{X})^2} \\
&- \sqrt{(\sqrt{m_W^2 + (\vec{p}_{W^-} - \vec{X})^2} + E_{\bar{b}})^2 - (\vec{p}_{\bar{t}} - \vec{X})^2}
\end{aligned}
\qquad (6.6)
$$

Applying Newton numerical method on this set of equations, vector $\vec{X}$ is found and consequently also the other solution $\vec{p}_\nu{}'$ for neutrino momentum. There is parameter which controls that second solution does not coincide with the first one. The minimal distance Δp_ν between two solutions for neutrino momentum ($\Delta p_\nu \equiv |\vec{p}_\nu{}' - \vec{p}_{\nu_1}| = |\vec{X}|$) is required to be greater than 1 GeV. We performed a check, to see what is the typical distance between neutrinos momenta for different kinematic solutions. We used Monte Carlo sample at generator level with input top mass 178 GeV [6]. The distribution is shown in Fig. 6.4. We can see that our cut (1 GeV) is well below the most probable value ($\sim$ 80 GeV) of distance between the solutions.

By the above introduced procedure, we can not get the precise solution even in the principal, i.e. even if we would measure the kinematic properties of variables precisely.

---

[6] We usually use for the MC studies sample with generated top mass of 178 GeV. It is because, at the time prior to this measurement, the world top mass average was 178 GeV (combined Run I average), see Sec. 2.2.3. Where is more appropriate, we use other masses like 175 or 170 GeV.

The reason for that is not in the method itself, rather in the physical variables involved. The masses of top quark and W boson are not precise constants, but they are statistical variables with some uncertainties. They are given by Breit-Wigner resolutions with the width of $\Gamma_{top} \approx 1.5$ GeV and $\Gamma_W \approx 2.1$ GeV. Practically, the dominant are resolutions of measured variables – for the assumed variable $p_z^{t\bar{t}}$, we have Gauss distribution given by simulation, while measured variables we know with a given measurement errors (which are also Gaussians). Therefore, it is necessary to find a robust procedure to solve the given set of equations and to minimize the effect of experimental errors.

Moreover, we have multiple solutions, because we need to choose only 2 jets for mass reconstruction while we can have more than 2 jets in final state due to initial and final state radiation. Another solutions come from two possible combinations of jets and leptons to form W boson ($\ell^+ + jet_1$ or $\ell^+ + jet_2$). Also, our equations can have multiply kinematically possible solutions (four, as we mentioned above).

Therefore, using reconstructed top mass from data events as direct estimate of top quark mass would lead to considerable shifts from the true values. However, this can be overcome.

How we deal with all of the above mentioned complications will be described in the following section.

6.2 Practical usage of the method

As it was stated above in previous section, there exist quite a lot factors which shift the reconstructed top mass from true top mass.

Therefore, we use well known procedure of using simulated Monte Carlo (MC) templates, which removes the possible bias in reconstructed variable by comparing such variable between MC simulation and the data. By doing so, you will get the bias in reconstructed value from true value, however it will be the same in data and MC simulation. Of course, it holds only in the case the MC simulation describes the data appropriately, i.e. the distributions of kinematic variables agree between MC simulation and data. We will show later in the chapter, that this is the case for CDF MC simulation.

Moreover, such reconstructed variable needs to be as much as possible dependent on true top mass. It's clear, that our reconstructed top mass is such sensitive variable and we will show that this is indeed the case later in the next chapter (see Fig. 7.3). However, there can be used other variables as the top quark mass sensitive variables. For more information, see Sec. 5.1.

How we actually perform the comparison between data and MC simulation will be shown in Sec. 6.2.3.

However, first we need to reconstruct top mass for a given event.

6.2.1 Reconstruction of top mass for an event

In principle, we deal with two kind of complications:

- values of measured variables are not exactly determined, rather they have some precision given by detectors resolutions
- there exist multiply (up to 8) solutions for top mass in a given event. Moreover, there exists ambiguity in choosing two b-jets, in case there are more than two jets in the final state.

We will show in next sections how we overcome these problems.

Uncertainties of measured variables

The measured quantities, i.e. momentum of the leptons, energy of the jets, and $\not\!\!E_T$ have experimental resolutions. We need somehow to include these resolutions into our measurement. There exist more possibilities how to do it:

minimizing χ^2 : the resolutions are taken into account by allowing the reconstructed value (e.g. $M_{\ell\nu}$) to differ from true value (e.g. $M_W = 80.4$ GeV) within uncertainty on this variable (e.g. $\Gamma_W \approx 2.1$ GeV). The terms in the χ^2 have the form $(M_{rec} - M_{true})^2/\sigma^2$, so the values of variables, for which the χ^2 has minimal value, are constrained by the corresponding uncertainties.

This procedure of using χ^2 have been applied in a few different top mass measurements, see "Neutrino ϕ weighting method" (Sec. 5.1.1) or template method in lepton+jets channel (Sec. 5.1.2). However, it can be used only when the errors on variables follow Gaussian distributions [7].

smearing variables : The values of kinematic variables together with their resolutions defines parameter phase-space for a given event. By smearing the statistical variables, i.e. randomly generating these quantities many times (say, 10000) around measured values within appropriate uncertainties, we will obtain smeared events ("trials"). By kinematic reconstruction of such "trials", a distribution of

---

[7] even this is not the case for the e.g. width of W boson Γ_W, in this case, it can be used as an approximation.

possible top masses for a given event is obtained (consistent with measured kinematic characteristics of the events and the measurements errors).

The second possibility is more general as the smearing of the variables can be done according any distribution, not necessary only Gaussian. Other that that, the two approaches are equivalent. We choose to use second procedure. Even we now use the Gaussian distributions for variables, the one of the possibilities to improve measurement method would be to use non-Gaussian distributions for variables (e.g. the uncertainties on jet energies are non-symmetrical).

We smear jet energies, $\not\!\!E_T$ and $p_z^{t\bar{t}}$ for a given event. By kinematic reconstruction of "trials", a distribution of top masses for a given event is obtained, see Fig. 6.11. A most probable value from this distribution is considered as the top mass estimate for a given event, we call it "**raw top quark mass**". This is our top mass sensitive variable which we will compare between simulated and data events.

In order to perform smearing, we need to know what are the resolutions of individual kinematic variables. We assume the jet E_T resolution to be $\sigma(E_T) = 1.8 + 0.09E_T$ and for $\not\!\!E_T$ resolution we assume $\sigma(\not\!\!E_T) = 1.3\sqrt{\not\!\!E_T}$. These jet and $\not\!\!E_T$ resolutions were appropriate in Run I [4]. However, it was shown within CDF collaboration that such jet energy resolution is appropriate also for Run II conditions [84]. The study was performed on dijet events. Both jets were required to have pseudo-rapidity within range $|\eta| < 2$ and were corrected for all jet corrections described in Sec. 4.1.4. In the estimate, there was calculated the uncertainty σ_D on the vector sum of both jets $\vec{E}_T$ and the jet energy resolution for one jet is taken as $\sigma_D/\sqrt{2}$. The comparison between data, MC optimized to CDF data ("Pythia Tune A" – tuned to fit underlying event) [85] and some possible resolution curves for σ_D as a function of average of both jets E_T is in Fig. 6.5. It can be seen that our curve reasonably describes the data and MC. However, this dijet resolution was obtained for light-quark jets. Therefore, using the b-jets for deriving the uncertainties can be one possible improvement of the reconstruction method.

For $\not\!\!E_T$, we show that the above mentioned resolution also holds. We performed the study on $t\bar{t}$ MC events with generated mass 175 GeV in the following way. For the events which have sum of generated neutrino transverse momenta $\sum_{i=\nu,\bar{\nu}} E_T^i$ in certain range (e.g. 20 to 40 GeV), we create distribution of difference between measured $\not\!\!E_T$ and a such sum: $\not\!\!E_T - \sum_{i=\nu,\bar{\nu}} E_T^i$. As a uncertainty of such distribution, we take the error σ of the fitted Gaussian function. Then, we plot the resolution as a function of $\not\!\!E_T$, see Fig. 6.6. Here, $\not\!\!E_T$ corresponding to each point is defined as mean value of $\sum_{i=\nu,\bar{\nu}} E_T^i$ distribution plus mean value of $\not\!\!E_T - \sum_{i=\nu,\bar{\nu}} E_T^i$ distribution for the bin in

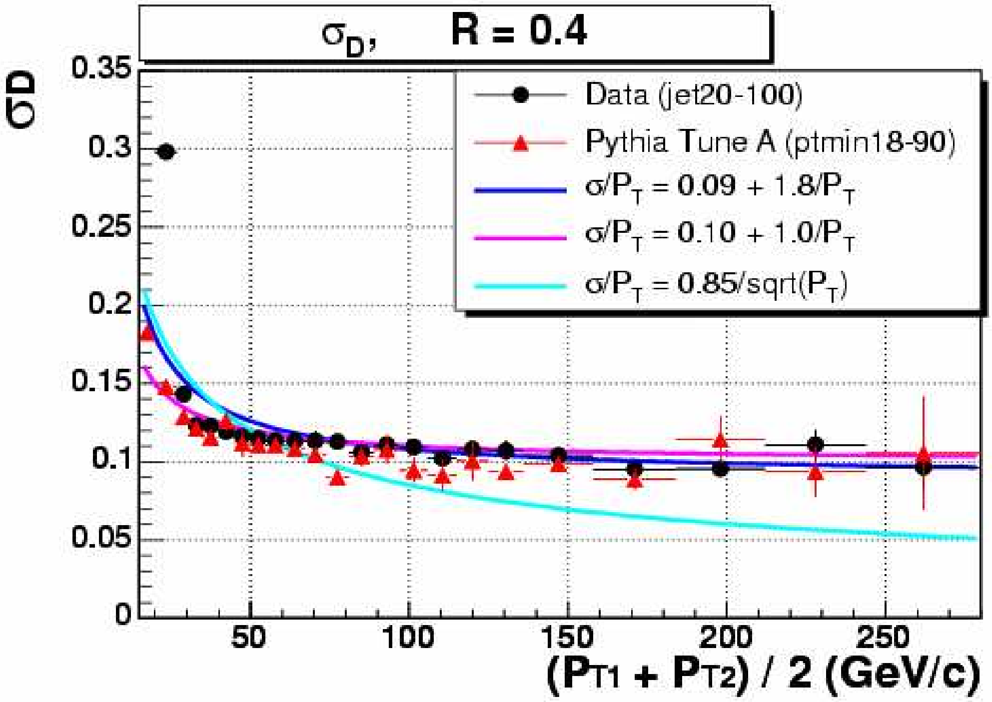


Figure 6.5: Dijet resolution as a function of average of two jet E_T. Compared are data, MC (Pythia Tune 'A') and a few possible parametrizations.

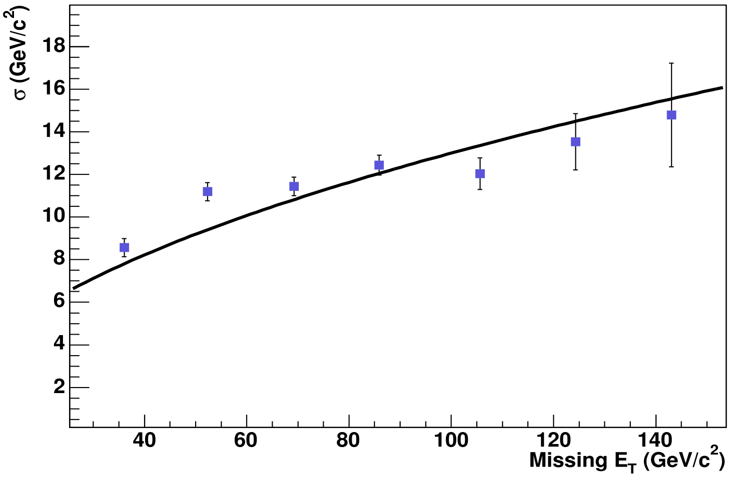


Figure 6.6: $\not\!\!E_T$ resolution as a function of measured $\not\!\!E_T$. The curve corresponds to $\sigma(\not\!\!E_T) = 1.3\sqrt{\not\!\!E_T}$.

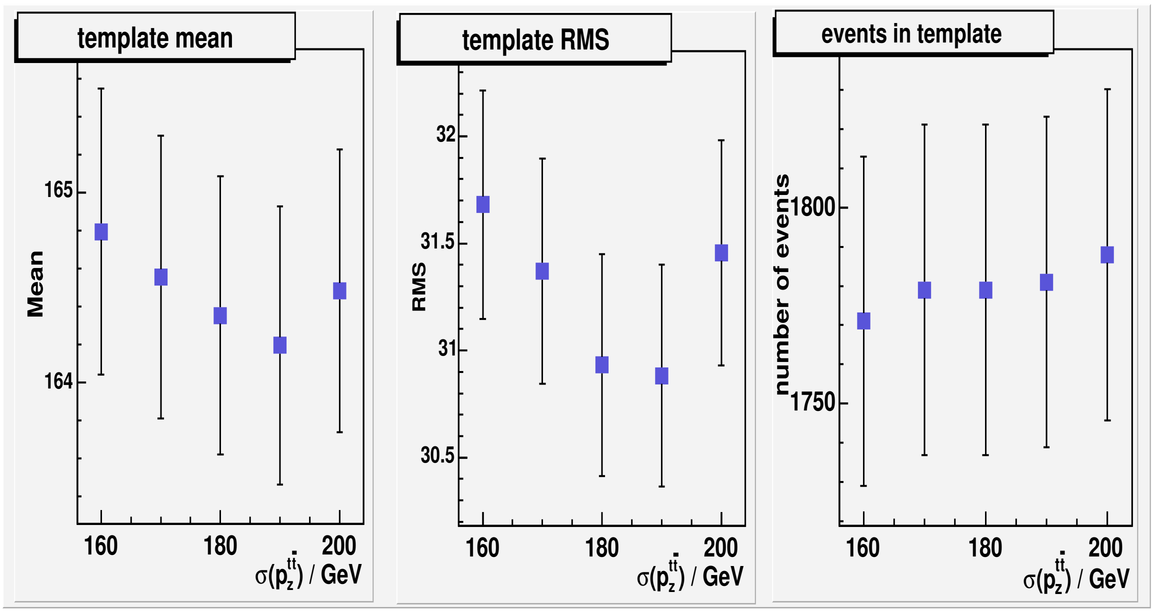


Figure 6.7: The dependence of template characteristics (mean value, RMS and number of events) as a function of $\sigma(p_z^{t\bar{t}})$ for one MC sample with generated top mass 170 GeV.

question. The curve in the plot corresponds to resolution $\sigma(\not{E}_T) = 1.3\sqrt{\not{E}_T}$ and we can see that it roughly describes the points.

As can be seen, we smear the $\not{E}_T$ and jets independently. However, because the jet E_T's are directly included in $\not{E}_T$, it would be more appropriate to smear only jets and use these smeared E_T's in the $\not{E}_T$, rather than smearing the $\not{E}_T$ alone. We made such simple exercise. The results in "raw top mass distribution" we obtained, were not any better comparing to original. This is because after such change, the $\not{E}_T$ resolution is unrealistic. Therefore, it would be required to use some more advanced technique to have smeared jets included, but also the overall uncertainty on $\not{E}_T$ kept.

As the "theoretical" constraint on $p_z^{t\bar{t}}$, we can consider a "measurement" of $p_z^{t\bar{t}} = 0$, with a $\sigma = 195$ GeV of Gaussian distribution (see Sec. 6.1). However, in the real calculations, we still used old value from Run I calculations $\sigma = 180$ GeV [86], but we prove the effect of this is negligible, see Fig. 6.7. For the study, we used MC $t\bar{t}$ sample with input mass 170 GeV.

We checked resolutions of leptons energies, whether we can neglect them comparing to jet E_T and $\not{E}_T$ resolutions in our smearing procedure.

Lepton p_T resolution was studied separately for electrons and muons. The comparison was only performed for measured leptons, which were associated to a lepton at

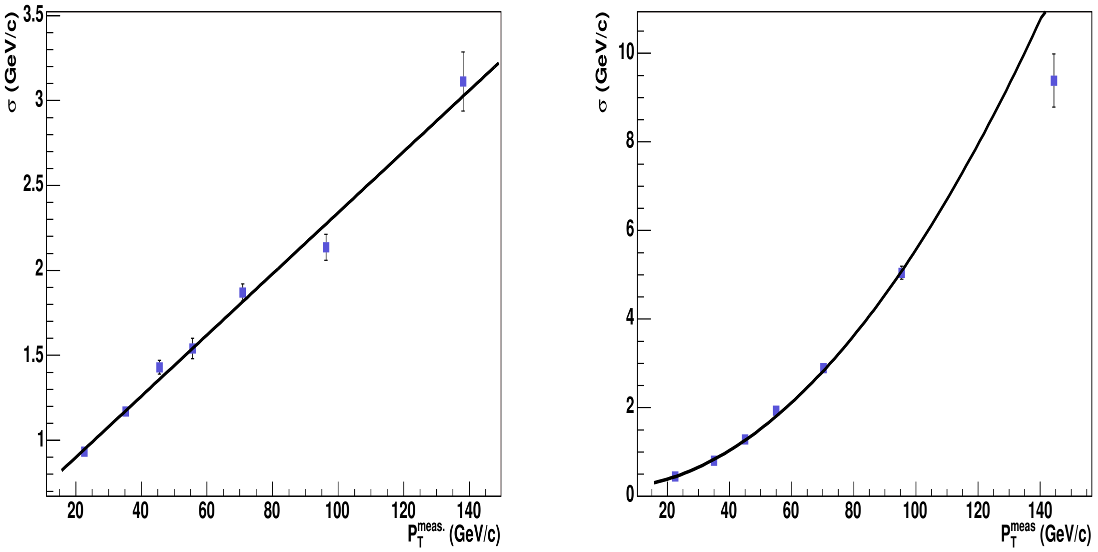


Figure 6.8: Electron (left) and muon (right) p_T resolutions as a function of measured electron and muon p_T, respectively. The shown curves are the fits to the points.

generation level within a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$. The procedure for estimating the lepton resolutions was similar to one used for $\not\!\!E_T$. For different generated lepton transverse momentum p_T^{hepg} bins, we create distribution of difference $p_T^{meas} - p_T^{hepg}$. As a uncertainty of such distributions, we take the error σ of the fitted Gaussian function, and plot the resolution as a function of measured lepton transverse momentum p_T^{meas}, see Fig. 6.8. Here, p_T^{meas} corresponding to each point is defined as mean value of p_T^{hepg} distribution plus mean value of $p_T^{meas} - p_T^{hepg}$ distribution for the bin in question. The curves on the plots are the fits to the corresponding points:

$$\sigma(p_T^{electron}) = 0.018 * p_T + 0.54 \tag{6.7}$$

$$\sigma(p_T^{muon}) = 0.00054 * p_T^2 + 0.17. \tag{6.8}$$

Comparing these leptons resolutions to the resolutions of jets and $\not\!\!E_T$ and realizing what are the typical values (means) of particular kinematic variables (see Fig. 6.18, 6.19, 6.20), we conclude that leptons resolutions are much smaller. $\sigma_\ell \sim$1.5 GeV for typical leptons momenta $P_T \sim 50$ GeV, while $\sigma \sim 7$ GeV for jets and $\sigma \sim 11$ GeV for $\not\!\!E_T$ for their typical values of 60 GeV and 75 GeV, respectively. Therefore we neglect the lepton resolutions in our reconstruction procedure and assume the leptons momenta are measured precisely.

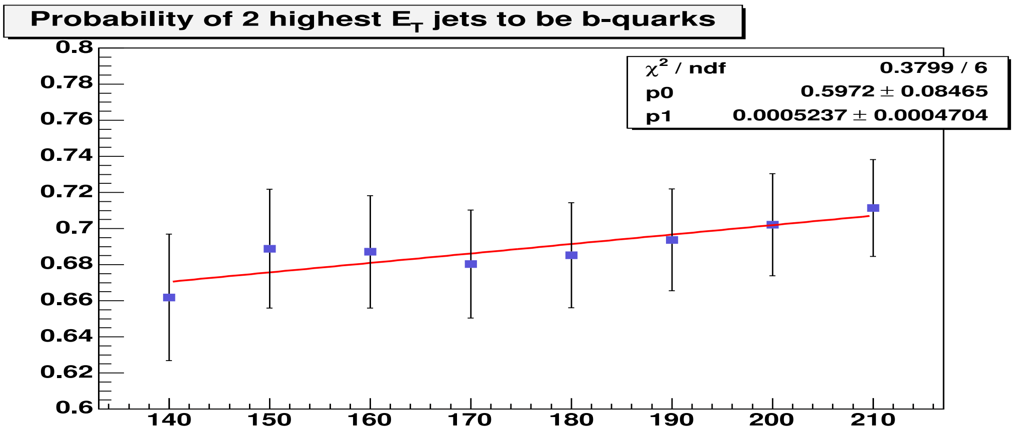


Figure 6.9: The probability that two highest E_T jets are coming from hadronization of b-quarks as a function of top mass.

Multiply solutions

As was already pointed out in previous section, there exists ambiguities in choosing the correct solution from more possible ones. Here we will describe, how we overcome them. We choose only one of all possible solutions. There exist also other approaches like using all possible solutions weighted by some factors. Examples of such weighting methods are two other template methods used in $t\bar{t}$ dilepton channel at CDF, see Sec. 5.1.1.

First thing which we need to decide on even before we start mass reconstruction itself is which two jets to use in the procedure. There could be more than two jets in final state due to ISR, FSR. We consider as b jets the **two jets with the highest** E_t. We made a study how often this is the case. From events which passed the $t\bar{t}$ dilepton event selection, we looked at the events where two highest E_T jets are b-jets. As a b-jet, we consider the jet which matches the original b-quark within cone $R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.4$ around b-quark direction. The probability that 2 highest E_t jets are b-quarks as a function of input top mass is plotted in Fig. 6.9. It corresponds to about 68% for 170 GeV mass sample. After the above selection of two jets, we have all particles selected which we need for mass reconstruction. No explicit identification of b jets ("b-tagging") is performed as of now. Finding adequate treatment for the events with more than 2 jets should improve the measurement. One of the possibilities is to use identified jets. Another possibility is to merge the jets according some criteria.

However, there still exists multiply solutions possible for each event. All together, there are 8 different possible solutions. We can pair two leptons and two jets by two

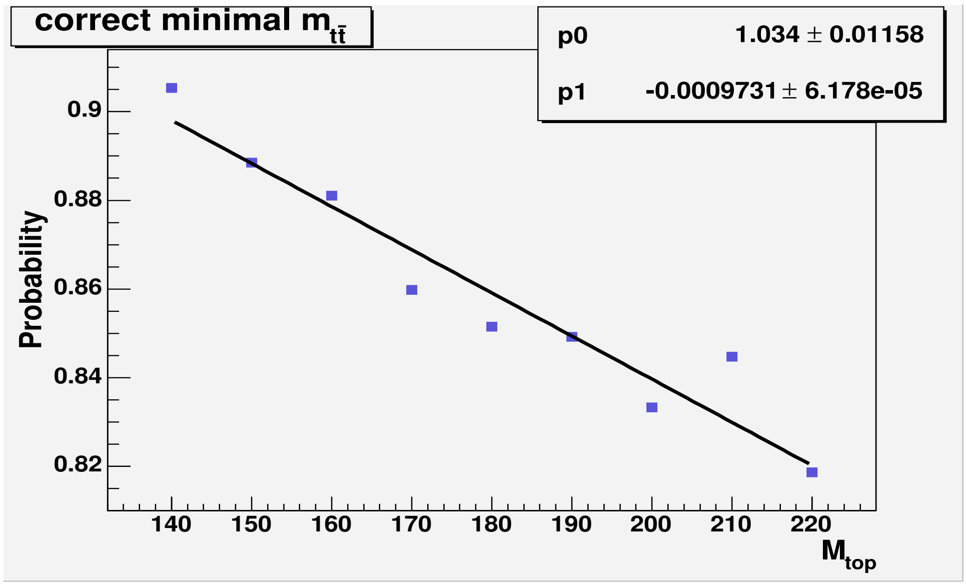


Figure 6.10: The probability of selecting correct kinematic solution from multiply solutions according $m_{t\bar{t}}$ criteria as a function of input top quark mass.

different ways (($\ell_1 - jet_1; \ell_2 - jet_2$) or ($\ell_1 - jet_2; \ell_2 - jet_1$)). For each of these two configurations, there may exists up to 4 kinematically allowed solutions coming from (6.2).

From **kinematically possible solutions** we take the one with smallest invariant mass of whole $t\bar{t}$ system [8]. The motivation for such selection comes from Standard Model. The cross-section of $t\bar{t}$ production by $q - \bar{q}$ annihilation (which is dominant at Tevatron) is inversely proportional to center-of-mass energy squared $\hat{s}$ of interacting partons, see Fig. 2.8. Therefore, from two kinematically possible solutions is more probable the one with smaller $\hat{s}$, i.e. smaller invariant mass of $t\bar{t}$ system.

We made check, how often the solution with the lower $m_{t\bar{t}}$ is the correct solution. We considered the solution to be the correct one, if it had reconstructed $m_{t\bar{t}}^{recon}$ closest to the generated $m_{t\bar{t}}^{HEPG}$. Using all smeared events ("trials") which have solution for all MC events in a given sample, we obtained the decreasing linear dependence of the probability on input top mass, see Fig. 6.10. For input top mass 170 GeV, it means about 87 % probability for solution with the lower $m_{t\bar{t}}$ to be the correct solution.

For each of the **two different lepton+jet configurations** (($\ell_1 - jet_1; \ell_2 - jet_2$) or ($\ell_1 - jet_2; \ell_2 - jet_1$)) in given event, we obtain a distribution of masses for the trial events. We start with the same number of trial events (10000) for each lepton+jet con-

---

[8] defined as invariant mass of the four-vector sum of 2 leptons, 2 b-jets and 2 neutrinos

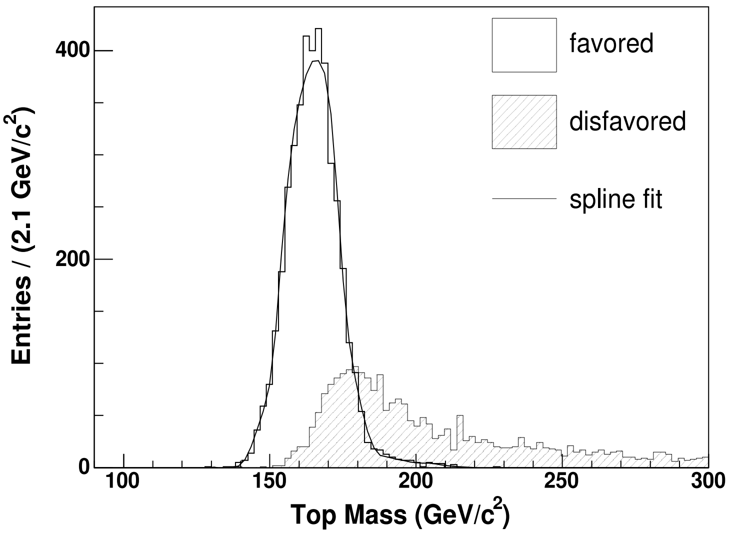


Figure 6.11: The mass distributions coming from smearing of one data event. There are two histograms for two different combinations of lepton+jet. The curve corresponding to favored lepton+jet combination is a spline fit to the distribution.

figuration. However, because sometimes it is impossible to kinematically find solution at all, i.e the set of equations (6.2) don't converge according required conditions (see page 113), we end up with different number of reconstructed trial masses in histograms for two different lepton–jet configurations. Therefore, we use only the combination of lepton+jet which is more probable to reconstruct, i.e. the distribution with larger number of entries in histogram. In Fig. 6.11, there is an example of one event with histograms for both combinations of lepton+jet. The most probable value from the selected histogram is "raw top mass" and it will be used as top mass sensitive variable in further mass measurement procedure. More precisely, we perform the spline fit to this distribution in order to minimize the fluctuation effects of distribution with finite statistics. The value of "raw top mass" is then the center of the histogram bin with the maximal value of the spline fit to the distribution.

Moreover, we require that event to pass the mass reconstruction procedure should have at least 10 % probability of reconstruction, i.e. at least 10 % of trials should be reconstructed for a given event. This cut was introduced to have decent statistics in histogram in order to be able to perform spline fit to such distribution. The studies about reconstruction probability are summarized in Appendix A.

We performed a check, how often the lepton–jet combination with more entries in

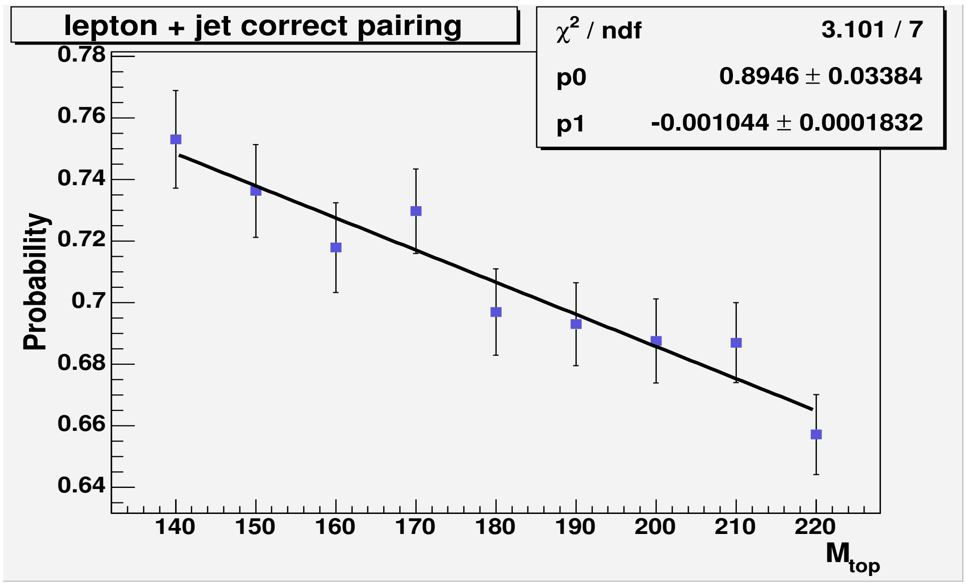


Figure 6.12: The probability of selecting correct pair of lepton + jet by our selection procedure as a function of input top quark mass.

histogram is the correct one. We have used Herwig 178 GeV sample and there were considered only the events which have both jets matched to original b-quarks (within cone $R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.4$) and also both leptons are matched to both b-jets (here, matching means, that each lepton has corresponding, i.e. opposite charged b-jet). Only in this case it is possible to conclude, whether the selected combination is the correct one. The probability of selecting correct lepton+jet pair by our selecting procedure as a function of input top mass is plotted in Fig. 6.12. It shows that the probability is linearly decreasing with increasing top mass and that it is about 72 % for 170 GeV input top mass.

We also checked how different are the number of entries in histograms (i.e. number of reconstructed trials) for selected vs. disfavored combination, in other words, how the selected vs. disfavored combination differ and whether it is not unimportant which to choose. This was done for the 178 GeV mass sample. We show the histogram of the reconstruction probability of trials for favored and disfavored solution in Fig. 6.13. Note, in roughly 45% of events, no single trial was reconstructed for disfavored combination. It can be also seen that in average there are almost 3 times more reconstructed trials in selected combination comparing to disfavored combination. If we would choose disfavored combination (and still require at least 10 % reconstruction probability for the event to be considered as reconstructed), we would reconstruct mass in only about 56 % of events (comparing to about 75 % for favored solution, see Sec. 7.1). This

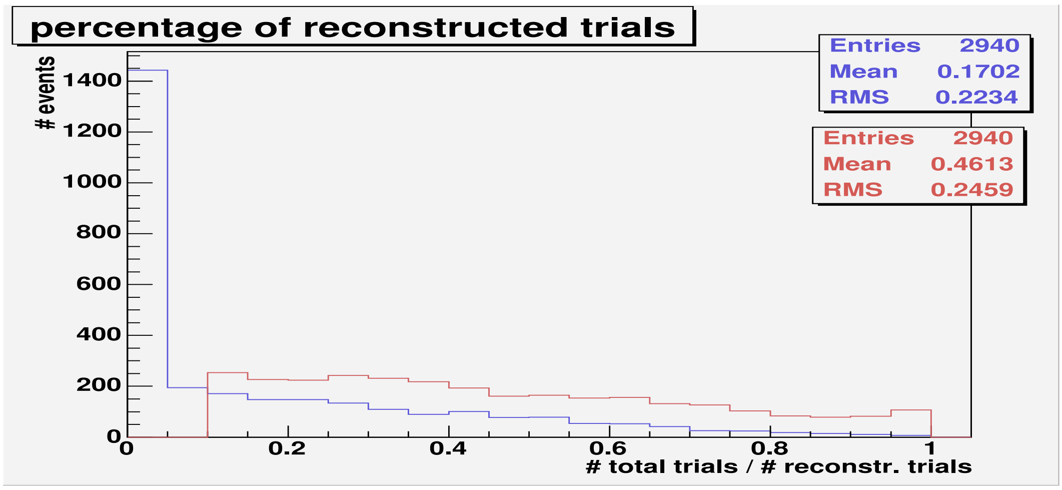


Figure 6.13: The histogram of the reconstruction probability of trials for favored (red) and disfavored (blue) lepton + jet combination.

means dramatical reduction of event statistics.

Summary of reconstruction procedure for a given event

Here, we summarize the whole procedure of top mass reconstruction for a given event. Here, we assume that all the kinematic properties of the event are reconstructed. First, we select two jets in the event with the highest E_T. After that, we have all the kinematic variables for all the particles (2 jets, 2 leptons, $\not\!\!E_T$) which we need in mass reconstruction.

We perform the next steps for both possible lepton–jet configurations. We smear the measured variables according the resolutions. For each smeared event (trial), we reconstruct top quark mass using kinematic equations. Sometimes we get no solutions, sometimes we get multiply kinematic solutions. If there are multiply solutions, we choose the one with smallest value of $m_{t\bar{t}}$. At the end of smearing procedure, we obtain the mass distribution for a given physics event where the reconstructed masses for trial events are included.

From two trial mass distributions (one for each lepton–jet combination), we choose the one with more trial masses reconstructed. From the chosen distribution, we obtain the reconstructed mass (“raw top mass”) for a given event as a most probable value of the distribution.

Plotting in one picture the “raw top mass” for each reconstructed event of particular sample, we obtain distribution we call “**raw top mass distribution**”.

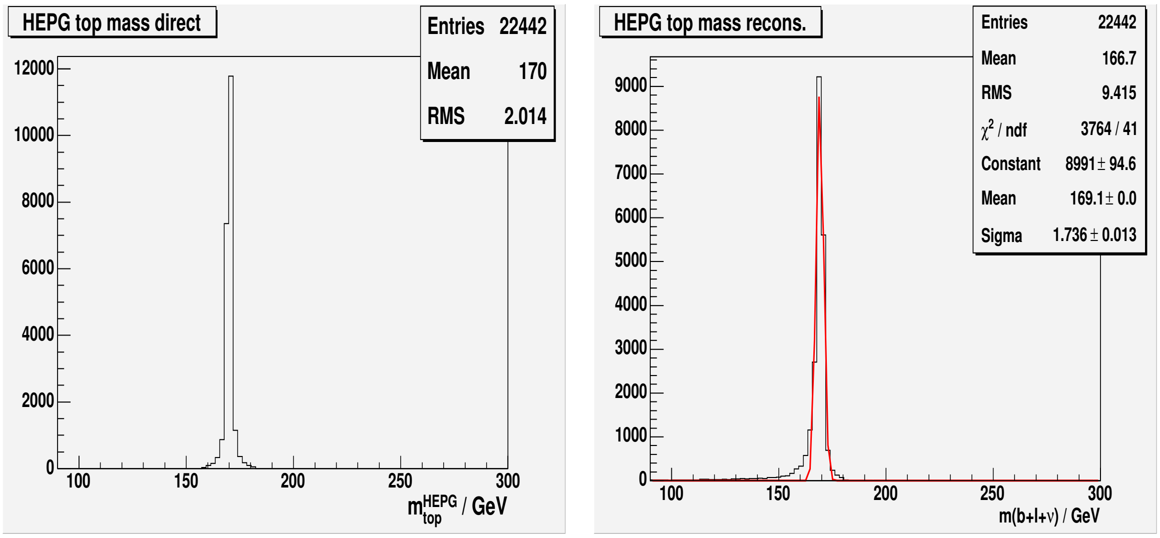


Figure 6.14: Generated top quark mass in MC sample with input top mass 170 GeV (left). The invariant mass of corresponding generated lepton, neutrino and b-quark for the same MC sample (right plot).

The normalized "raw top mass distribution" for each sample is called a **template** (see Fig. 7.3 for $t\bar{t}$ MC templates). The template is the probability density function (P.D.F.) for a given input top mass, i.e. it gives the probability that for given true (input) top mass m_t^{orig}, we will get reconstructed top mass with value m_t^{rec}.

6.2.2 Checks of reconstruction procedure

We performed a few checks, how the raw mass distribution is affected by all of our selection criteria in mass procedure.

First, we show the distribution of generated (HEPG) top mass in Fig. 6.14. We can see sharp peak (left plot) at generated value which is consistent with the input value. The right plot shows the invariant mass of lepton, neutrino and b-jet. Still, the similar sharp peak can be seen. However, there is now the tail on lower side of the peak which can be due to radiation of gluon from b-quark.

Next, we reconstructed the top quark mass by our procedure for events at parton level (no simulation and reconstruction of kinematics was performed). However, in this case, we did not performed smearing ($p_z^{t\bar{t}}$ was calculated using HEPG p_z's of neutrinos). We perform two studies. Once, we used the correctly paired lepton and jet. In second case, we paired the lepton and jet according our selection, i.e. we choose this lepton–jet combination which was more probable to reconstruct. The plots are in Fig. 6.15. We can see now the broader peaks and tails on both sides of the peak. The tails are due

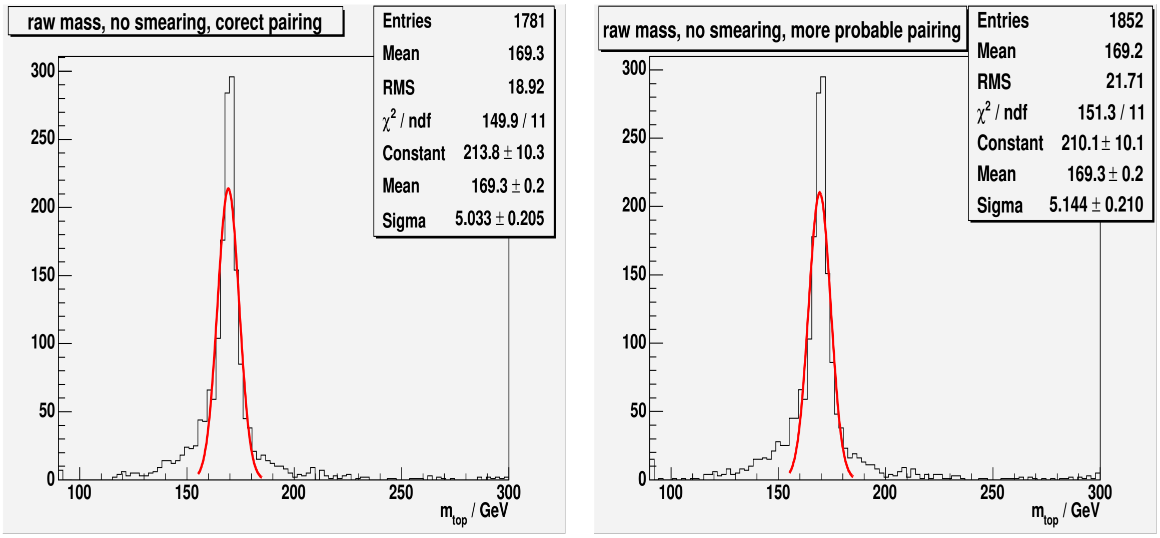


Figure 6.15: Reconstructed top mass for parton level events. No smearing was performed in this case. On left plot, the correct lepton + jet combination was used, while on right plot the combination was chosen according our selection criteria (more probable combination). The fit is performed for the range of roughly $\pm 3\sigma$ coverage (170 ± 15 GeV) of the main peak.

to different assumptions used in our method. Choosing the lepton–jet combination according our criteria makes the tails larger, however, the peak still remains at the same position and the width of the peak is roughly the same.

In next study, we still used parton level events. However, now we used complete reconstruction procedure, i.e. we included smearing in the process. We can see in Fig. 6.16 that this have drastic effect on reconstructed top mass. There is no longer sharp peak, rather quite broad distribution of reconstructed top mass. However, the peak is still in right position.

As a last study here, we performed the mass reconstruction for a events, which pass full chain of generation, simulation and reconstruction. We can see in Fig. 6.17 that still the distributions are wide. However, now even the peak of distribution is shifted from true value to lower values.

6.2.3 Comparison between MC simulation and data

The final step in top mass reconstruction is comparison of the values of sensitive variable (in our case, it's "raw top mass") between data candidates and Monte Carlo simulated events. The comparison is performed by creating MC templates, parameterizing them, reconstructing raw top masses for data events, and finally comparing MC templates

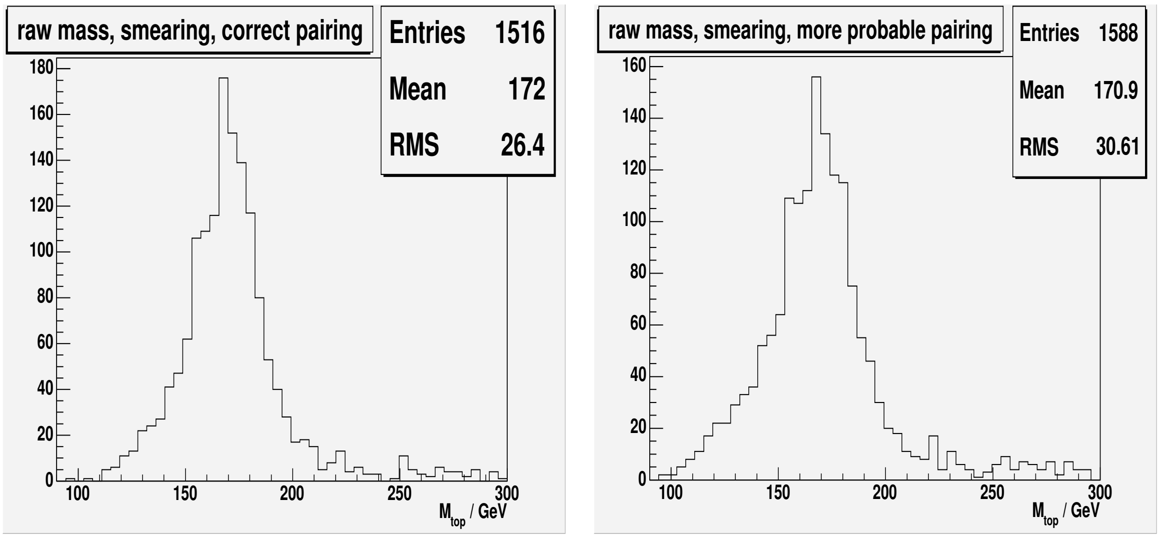


Figure 6.16: Top mass reconstructed by the method for HEPG events. The smearing procedure was used in this case. On left plot, the correct lepton – jet combination was used, while on right plot the combination was chosen according our criteria (more probable combination).

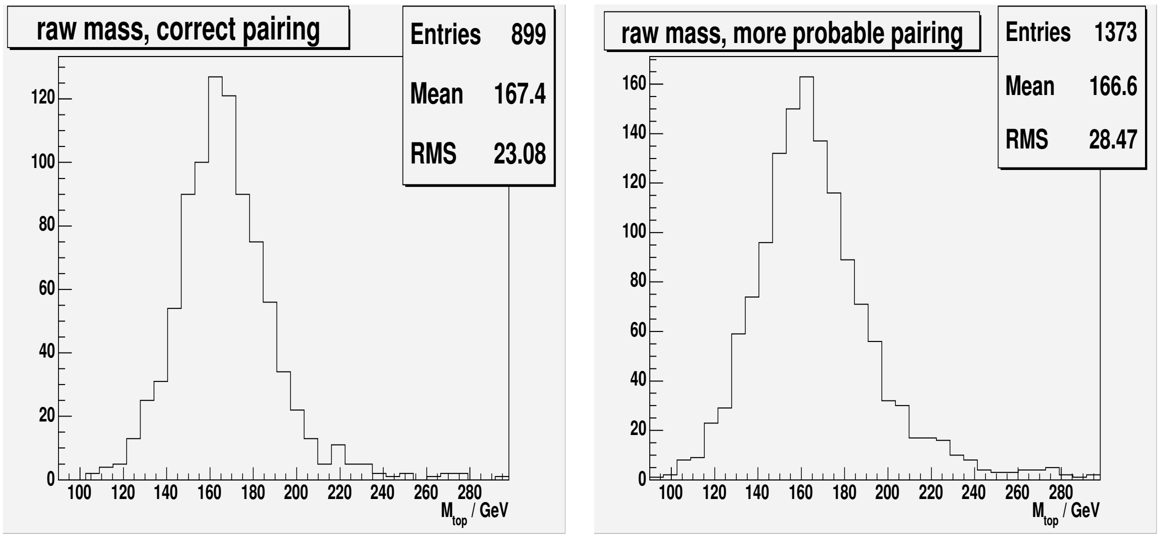


Figure 6.17: Reconstructed top mass for simulated events. The smearing procedure was used in this case. On left plot, the correct lepton – jet combination was used, while on right plot the combination was chosen according our criteria (more probable combination).

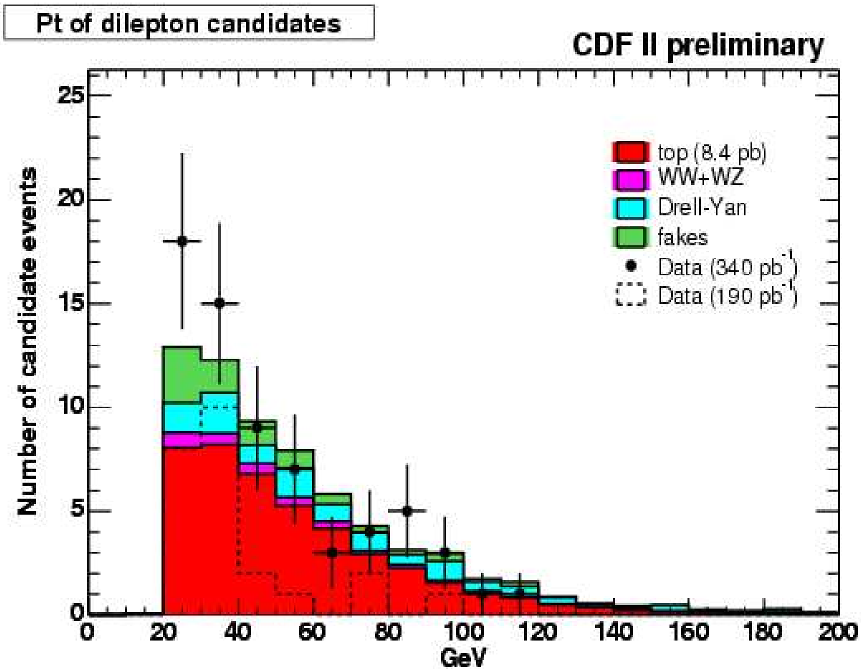


Figure 6.18: Comparison between data and MC simulated events for lepton transverse momentum P_T. The events which pass dilepton event selection are compared. We use MC sample with generated top mass 170 GeV. The number of $t\bar{t}$ events corresponds to CDF measured $t\bar{t}$ cross-section in dilepton channel using 190 pb^{-1} [63].

and raw top masses from data by likelihood minimization.

However, before we can do the comparison of reconstructed top mass between MC and data, we need to make sure that MC models correctly the kinematic variables which have impact on reconstructed top mass. Here, we compare some basic kinematic characteristics between MC and data. Most of them we use in our mass reconstruction procedure. The comparison is shown in Fig. 6.18, 6.19, 6.20 and 6.21 for the following kinematic variables: lepton P_T,$\not{E}_T$, jet E_T and number of jets, respectively. These plots were produced by dilepton cross section group [87]. As number of data $t\bar{t}$ candidate events is so low (33), the comparison between data and MC is limited mainly by data statistics. However, even here we can see that shapes of kinematic variables agree between data and MC and there is no statistically significant disagreement between data and MC. The more precise comparison will be possible with more data acquainted.

Of course , we need to take into account the background events. When reconstructing mass for data events, you don't know which event comes from background process and which is really signal $t\bar{t}$ event. Therefore, you must somehow include the

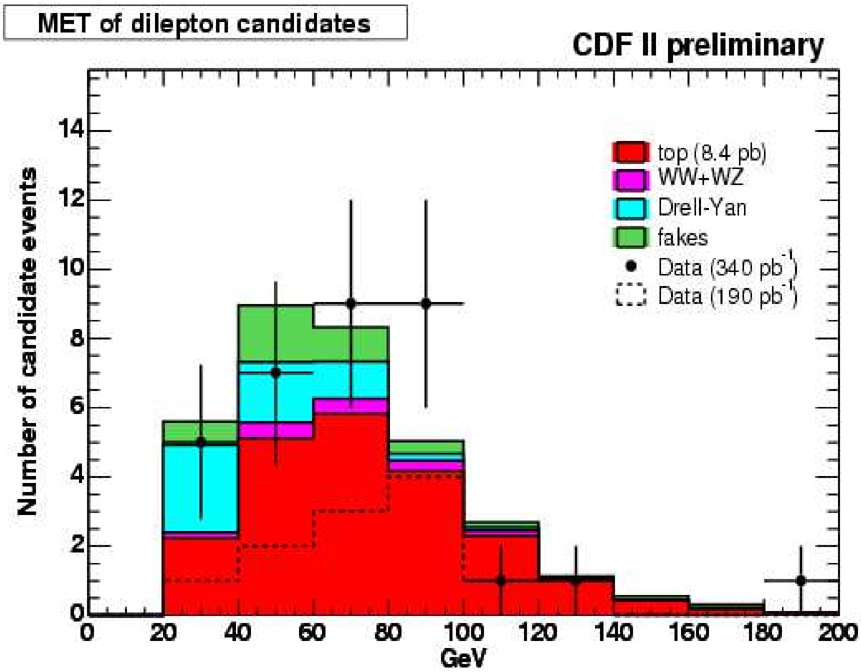


Figure 6.19: Comparison between data and simulated events for missing E_T. We use MC sample with generated top mass 170 GeV.

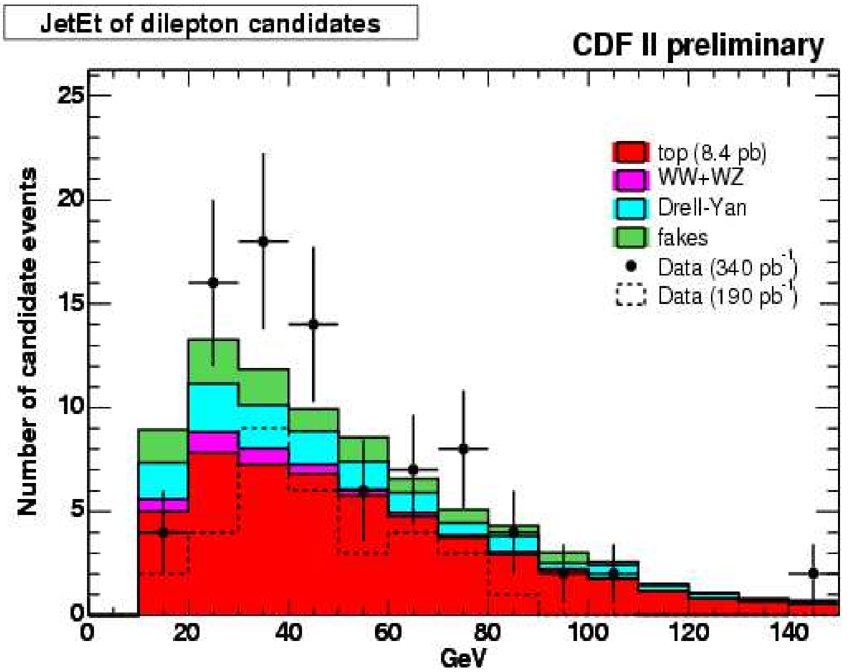


Figure 6.20: Comparison between data and simulated events for jets transverse energy E_T. We use MC sample with generated top mass 170 GeV.

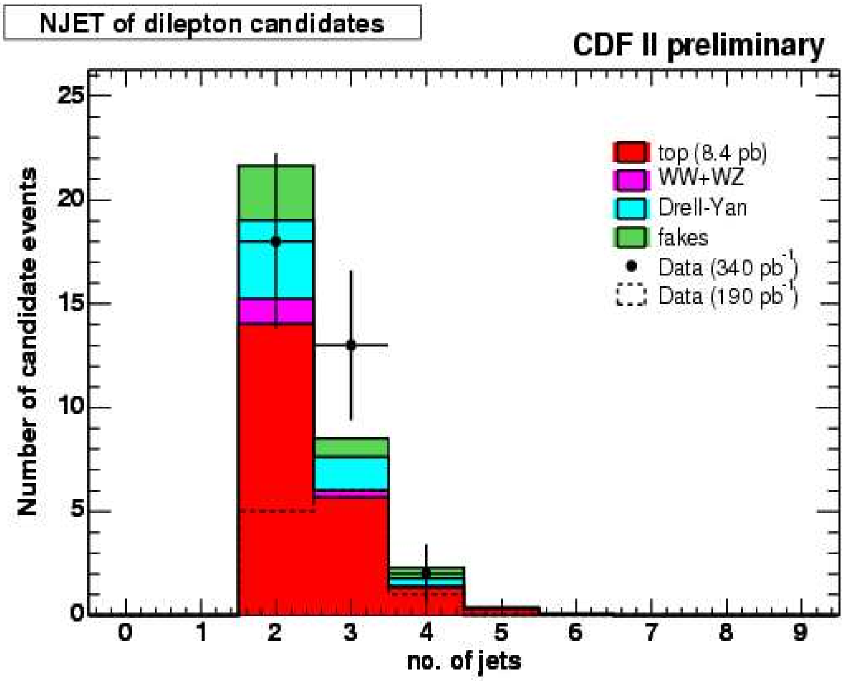


Figure 6.21: Comparison between data and simulated events for number of jets. We used sample of 170 GeV top mass simulated events.

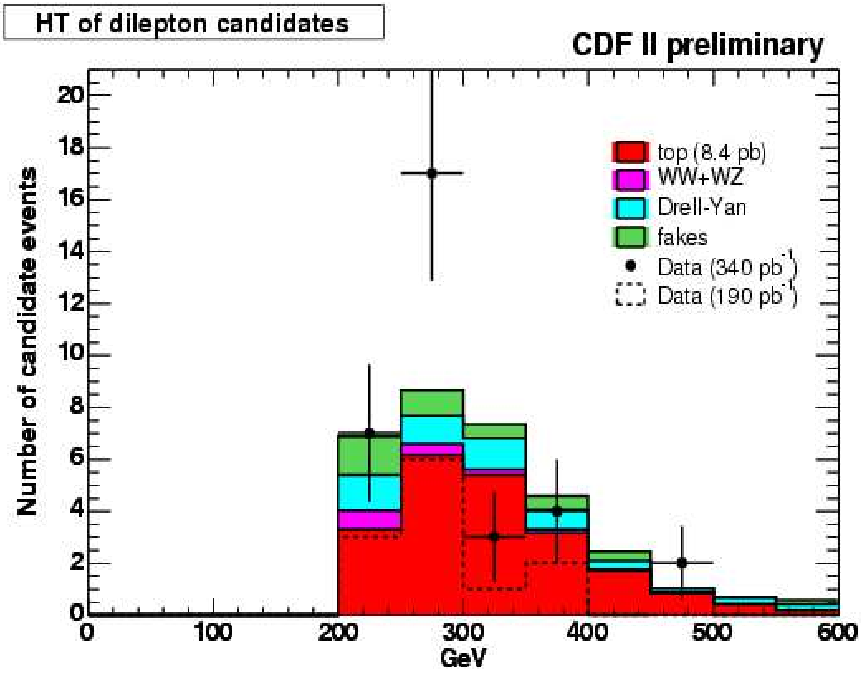


Figure 6.22: Comparison between data and simulated events for H_T variable (scalar sum of total transverse energy in the events). We use sample of 170 GeV top mass simulated events.

background in the calculation of the top mass. We include the influence of background processes on the top mass estimate very similarly as we do it for the signal events. We create the templates for each type of background (similarly as we create the templates for signal $t\bar{t}$ events), combine different types of background together and finally incorporate the combined background template into the likelihood function.

The likelihood function finds the probability that our data are described by an admixture of background events and dilepton $t\bar{t}$ events with a certain top mass. The total likelihood takes the form

$$\begin{aligned}
\mathcal{L} &\equiv \mathcal{L}_{shape} \times \mathcal{L}_{nev} \times \mathcal{L}_{bg} \\
\mathcal{L}_{shape} &\equiv \prod_{i=1}^{N} \frac{n_s \times f_s({m_{t_i}}^{rec}, m_t^{orig}) + n_b \times f_b({m_{t_i}}^{rec})}{n_s + n_b} \\
\mathcal{L}_{nev} &\equiv \frac{e^{-(n_s+n_b)}(n_s + n_b)^N}{N!} \\
-\ln \mathcal{L}_{bg} &\equiv \frac{(n_b - n_b^{exp})^2}{2\sigma_{n_b}^2},
\end{aligned} \tag{6.9}$$

where the sum of the number of signal and background events is constrained by Poisson distribution in $\mathcal{L}_{nev}$ and $\mathcal{L}_{bg}$ was introduced to constrain the number of background events by Gaussian. In this manner, the likelihood-estimated number of signal events is independent of the expected number of signal events based on an assumed $t\bar{t}$ cross-section. The top mass m_t^{orig}, the number of background events n_b and the number of signal events n_s are free parameters in a likelihood function and are returned as a result of the fit. The mass m_t^{orig}, which minimize $-\ln \mathcal{L}$ will be taken as the final top mass estimate for a given sample. The statistical uncertainty on m_t^{orig} is given by the difference between the minimization mass result and the mass at $-\ln(\mathcal{L}/\mathcal{L}_{max}) = 0.5$.

In order to use the likelihood formula, we need to know:

- N - number of events in data sample and their reconstructed raw masses $m_{t_i}^{rec}$
- n_b^{exp} - estimate of number of expected background events in the sample and also error on this estimate σ_{n_b}
- $f_s({m_{t_i}}^{rec}, m_t^{orig})$ - probability density function for $t\bar{t}$ samples, i.e. function which will parametrize the signal $t\bar{t}$ templates
- $f_b({m_{t_i}}^{rec})$ - function which will parametrize the background templates

Number of events N is directly measured in the data. We reconstruct their "raw masses" $m_{t_i}^{rec}$ by the procedure introduced in the previous sections. Number of expected background events and its error (n_b^{exp} and σ_{n_b}) we obtained by MC and/or data study (see Tab. 4.6).

Note, variables in likelihood, which correspond to number of events (N, n_s, n_b, n_b^{exp}), mean the number of events after top mass reconstruction! Therefore N is not the total number of $t\bar{t}$ candidates we see in the data, rather just the number of events in the data which the "raw mass" was reconstructed for, i.e. the events which pass the mass reconstruction procedure. Similarly, the n_b^{exp} is the number of expected background events after the mass reconstruction, i.e. number of expected background events (according Tab. 4.6) multiplied by the probability of reconstruction (see Fig. 7.2 and Tab. 7.4).

So, finally what remains are the probability density functions f_s,f_b which describe the templates. How the templates and functions which describe them are obtained will be shown in the following chapter.

Chapter 7

Simulated Events Results

As we described in previous chapter, the essential parts of our method are the templates for Monte-Carlo (MC) simulated events. How we produce the templates for signal $t\bar{t}$ events for different generated top mass, will be described below. Moreover, we will show, how we produce the templates for different types of background and how we combine them into final background template.

7.1 Signal $t\bar{t}$ templates

We used all together 21 CDF officially generated Herwig Monte-Carlo samples with the generated top mass within the wide range of 140 GeV to 220 GeV. The events in these samples already passed all the chain of generation, full simulation of passing through CDF detector and reconstruction of physical quantities.

The complete list of input top masses used in MC samples is in Tab. 7.1. There are also listed, for each sample, the number of events which pass $t\bar{t}$ dilepton selection criteria and the number of events which top mass was reconstructed for. In each sample, there were originally generated roughly the same number of events ($\sim$ 200000 of $t\bar{t}$ events). For higher top mass, the top quark decay particles have in average higher energies, and consequently higher chance of passing the selection cuts. Therefore, there are more events passing the $t\bar{t}$ selection cuts when going to higher top mass, as can be seen in table. Using these samples we create signal $t\bar{t}$ templates (see Fig. 7.3), i.e. we reconstruct the "raw top mass" for each event in a given sample according procedure described in previous chapter and make the normalized distribution of "raw top masses" for each MC sample with a given input top mass. The number of reconstructed events for each MC sample is listed in Tab. 7.1 and is roughly between 1500 and 2000 events.

The dependence of mean, median and most probable value (MPV) of template on input top mass is plotted in Fig. 7.1. It can be seen that the dependence of these

top mass [GeV]	N_{selec}	N_{recon}
140	1446	1119
145	1485	1132
150	1652	1255
155	1607	1217
160	1841	1406
165	1826	1424
170	1999	1549
175	1977	1492
177.5	2113	1600
180	2103	1571
182.5	2188	1676
185	2169	1630
187.5	1879	1395
190	2243	1690
192.5	2164	1600
195	2177	1598
200	2239	1645
205	2383	1740
210	2498	1793
215	2324	1684
220	2511	1815

Table 7.1: The list of input top masses used for generation of MC $t\bar{t}$ samples. There are also listed the number of events, N_{selec}, which pass $t\bar{t}$ selection criteria and number of events, N_{recon}, for which the method reconstruct "raw top mass".

characteristics is pretty linear.

Using the numbers from Tab. 7.1, we calculated the probability of reconstruction for a $t\bar{t}$ event of given input top mass. We plot this reconstruction probability as the function of input top mass in Fig. 7.2. We can see that this probability is essentially constant on the level of $74.75 \pm 0.42\%$ over the wide mass range. The reasons why the probability is not 100 % are discussed in Appendix A.2. Note, this reconstruction probability for the physics event differs from the probability of reconstruction for trial (smeared event), see Sec. 6.2.1.

After we obtained the templates, we need somehow to describe them by some functions which will be dependent on input top mass, or more precisely their parameters will be dependent on top mass. Therefore we simultaneously fit all the templates at

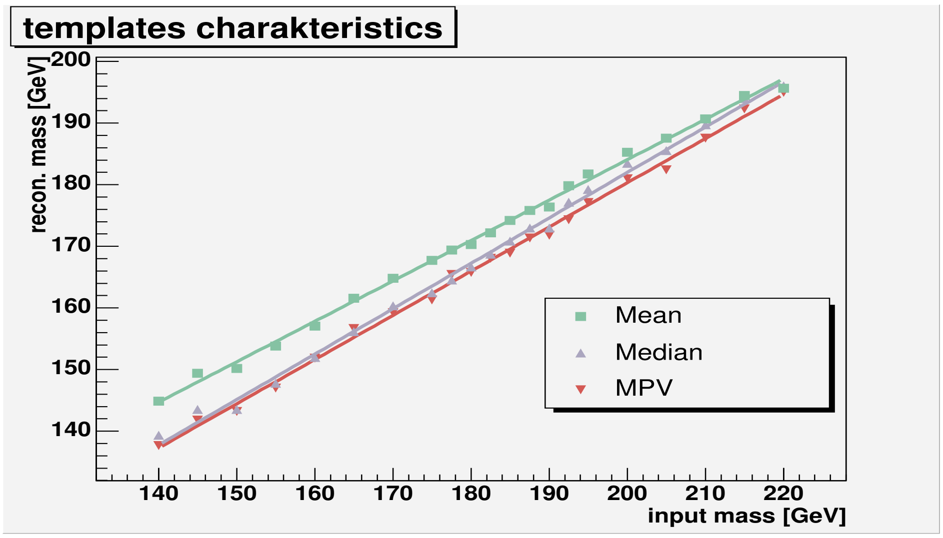


Figure 7.1: Mean, median and most probable value (MPV) of template as a function of input top mass.

once (we say, we do "global fit") [1]. As there is not known theoretically, what are the distributions the templates are following, it is up to us to choose the fitting functions. We chose to use Gaussian (to describe main peak of the templates) and Landau-like distributions. As we can see in Fig. 7.3, the templates have long tails, therefore we chose to use Landau distribution (known to have long tails) in order to describe these tails of the templates. However, what we actually use is just Landau – like function which is frequently used as an approximation to Landau function, but it is not "orthodox Landau". Some other methods use combination of Gaussian and Gamma distributions to describe the templates, see Sec. 5.1.1. As long as the fits appropriately describe the templates, it does not matter. So, our fitting function have the following form:

$$\begin{aligned} f(m_t^{rec}, m_t^{true}) &= \tfrac{p_3(m_t^{true})}{I_1} \cdot e^{-0.5(\lambda_{Landau} + \, e^{-\lambda_{Landau}})} + \\ &+ \tfrac{(1.-p_3(m_t^{true}))}{I_2} \cdot e^{-0.5\lambda^2_{Gauss}} \end{aligned} \tag{7.1}$$

where:

- λ_{Landau} and λ_{Gauss} are defined as follows

$$\lambda_{Landau} = \tfrac{m_t^{rec} - p_1(m_t^{true})}{p_2(m_t^{true})}$$

$$\lambda_{Gauss} = \tfrac{m_t^{rec} - p_4(m_t^{true})}{p_5(m_t^{true})}$$

---

[1] One of other possibilities, is to use Neural Networks to describe the shapes of the templates. Actually, we played such game, however without obtaining an improvement.

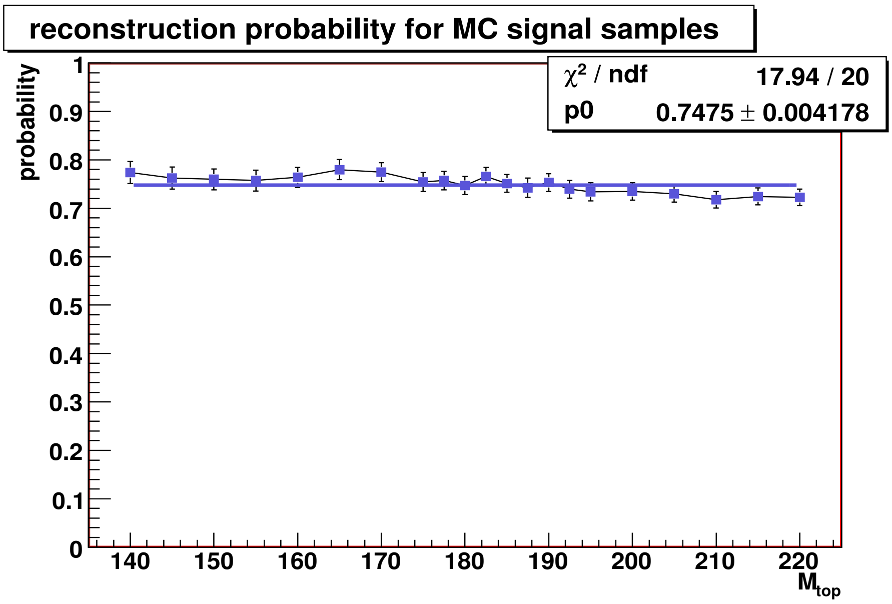


Figure 7.2: Probability of top mass reconstruction for signal $t\bar{t}$ event as a function of generated top quark mass.

- I_1, I_2 are normalization integrals in the region $m_t^{rec} \in (90 \text{ GeV}, 300 \text{ GeV})$ [2]

  $$I_1 = \int_{90}^{300} e^{-\frac{1}{2}(\lambda_{Landau} + e^{(-\lambda_{Landau})})} \, \mathrm{d}m_t^{rec}, \qquad I_2 = \int_{90}^{300} e^{-\frac{1}{2}\lambda^2_{Gauss}} \, \mathrm{d}m_t^{rec}$$

  Thus, the integral of $f(m_t^{rec}, m_t^{true})$ function is normalized to one by definition in the region of $m_t^{rec} \in (90. \text{ GeV}, 300. \text{ GeV})$.

- m_t^{true} is the input (true) top mass which was used in process of generation of the sample (as already said, we used MC signal samples with generated masses in range 140 GeV to 220 GeV)

- m_t^{rec} is the kinematically reconstructed top mass, i.e. it's our "raw top mass"

The $f(m_t^{rec}, m_t^{true})$ is the probability density function and $f(m_t^{rec}, m_t^{true})\mathrm{d}m_t^{rec}$ gives the probability that given event from sample generated with the top mass m_t^{true} will have reconstructed mass in the interval $(m_t^{rec}, m_t^{rec} + \mathrm{d}m_t^{rec})$.

As you can see, we are using five fitting parameters, four of them being the mean and the sigma of Landau and Gauss function, respectively. The fifth parameter, denoted here as p_3, controls the ratio of Landau function in total fitting function.

The fitting parameters $p_1, .., p_5$ which we try to find, depend on the input top mass m_t^{true}. Ideally, the reconstructed top mass would be linear dependent on input top

---

[2] we are looking for solutions only in this range of masses

mass with the offset zero and the slope one. So, naively, one would expect that even the offset and the slope can change for real case, the linear dependence will remain. Moreover, the linear dependence of means, most probable values and medians of the templates (see Fig. 7.1) suggests, that the dependence of parameters of the fit should be linear. We prove that this is indeed the case in Appendix B).

Therefore, we can write these parameters as follows:

$$p_i = a_i + b_i * m_t^{true},\ \ i = 1, 2, 3, 4, 5 \tag{7.2}$$

So, in total we have 10 parameters ($a_i, b_i, i \in (1, 2, 3, 4, 5)$), which we need to find the values for. Moreover, we assume the p_3 parameter to be constant (for more discussion, see Appendix B), i.e the fraction of the Landau function in total function is not dependent on input top mass ($b_3 = 0 \Rightarrow p_3 = a_3$).

The global fit to MC templates was performed by the "unbinned fit", i.e. by using actual values of raw top masses from all the events and all MC samples rather than using histograms (templates). We minimize the negative logarithm of the following likelihood using MINUIT package [88]:

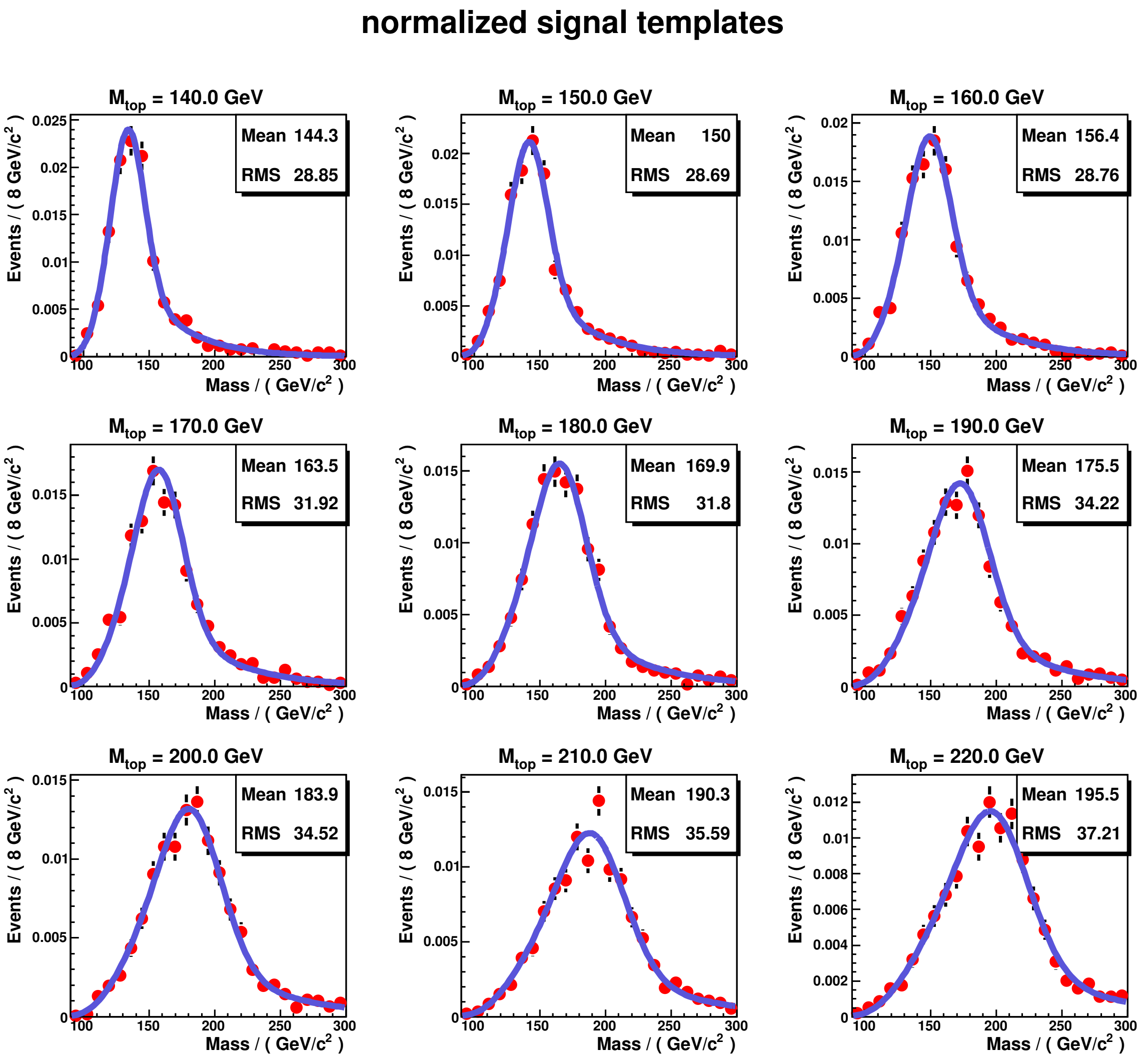
$$\mathcal{L} = \prod_{i=140}^{220} \prod_{j=1}^{N_{ev_i}} f(m_{t_{ij}}^{rec}, m_{t_i}^{true}) \tag{7.3}$$

where the product goes through all mass sample i from the range 140 GeV to 220 GeV and each event j from each sample. N_{ev_i} means the number of events in the i-th sample.

The normalized distributions of reconstructed raw top mass for the input top mass in the interval 140 to 220 GeV/c^2, together with the global fit, are presented in Fig. 7.3. Not all the samples we used in the fit are shown, just the samples with 10 GeV step. The parameters of the global fit are given in Tab. 7.2. Nice plot of "raw top mass" templates is in Fig. 7.4. This presents the two-dimensional visualization of probability density function in plane of m_t^{true} and m_t^{rec}.

We made a few checks to make sure the fit to the templates is appropriate. The quality of the fit can be characterized by $\chi^2/NDF = 2101.0/(2100 - 9 - 21) = 1.01$, where number of degrees of freedom (NDF) is calculated as follows, 2100 is the total number of bins in all histograms, 9 is the number of parameters and 21 is the number of constraints (21 normalized templates). This value of χ^2/NDF is expected to be close to one, so seems our fit is of good quality.

We also made the comparison by Kolmogorov-Smirnov(KS) test between each histogram and corresponding fit. The KS probabilities vary between 0.073 (190 GeV mass


normalized signal templates

M_{top} = 140.0 GeV, M_{top} = 150.0 GeV, M_{top} = 160.0 GeV, M_{top} = 170.0 GeV, M_{top} = 180.0 GeV, M_{top} = 190.0 GeV, M_{top} = 200.0 GeV, M_{top} = 210.0 GeV, M_{top} = 220.0 GeV

Figure 7.3: Signal templates for generated top mass from range 140 GeV to 220 GeV together with global fit.

sample) and 0.996 (187.5 GeV) with the average KS probability of 0.62. The 0.073 KS probability is extreme low case and all other KS probabilities are higher than 0.19. These values also show that our fit to templates was good.

7.2 Background templates

The main background processes to $t\bar{t}$ dilepton process are diboson, Drell-Yan process and FAKE processes, as already described in Sec 4.2.

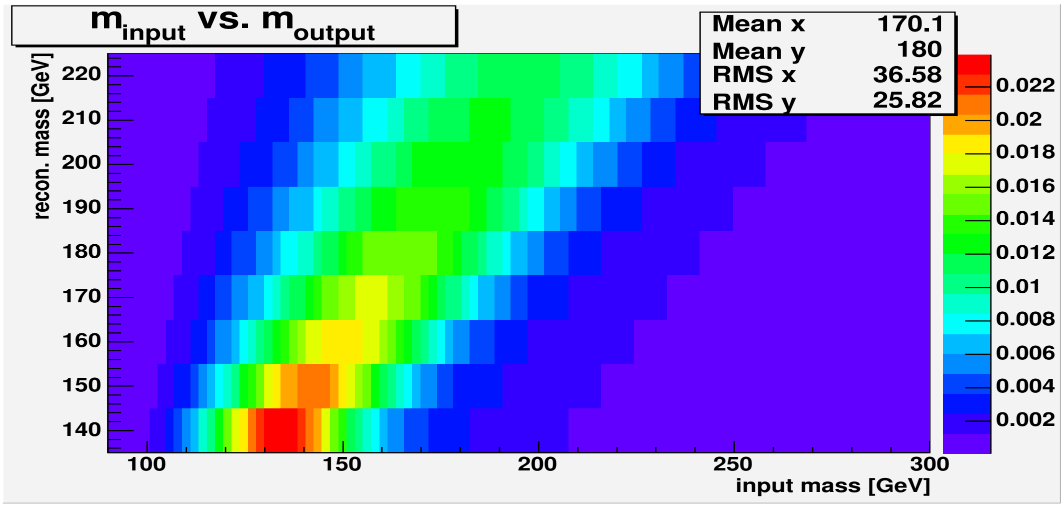


Figure 7.4: Distribution of probability density function of "raw top mass" in plane of m_{true} and m_{rec}.

parameter	a	b
p_1	69.659 ± 4.872	0.454 ± 0.025
p_2	-8.791 ± 0.547	0.184 ± 0.003
p_3	0.462 ± 0.023	0.0 ± 0.0
p_4	22.789 ± 2.018	0.797 ± 0.012
p_5	-12.523 ± 1.052	0.178 ± 0.006

Table 7.2: The slope a and offset b corresponding to parameters $p_i, i \in \{1, .., 5\}$ of the global fit to signal templates.

Once again, we used CDF official MC samples for getting the background events. These were created using ALPGEN generator and they are summarized in Tab. 7.3. There are also listed the types of generated samples used for particular background events. Because of small $t\bar{t}$ acceptance of background samples, there were requested particular number of partons (or particular types of partons) already at the generator level. This was done in order to increase acceptance in MC sample and consequently to decrease amount of generating mostly "useless" events (because they are rejected by $t\bar{t}$ selection criteria).

However, even in this case we ended with quite low statistics of background events for Drell-Yan and FAKE samples (statistics of WW and WZ samples was high enough). This is because only inside WW events we've got generated all basic objects needed to pass $t\bar{t}$ event selection (2 leptons, 2 jets and missing E_T coming from neutrino). On the other side, Drell-Yan and FAKE events can pass the $t\bar{t}$ event selection only in case there was mismeasurement or bad reconstruction, see Sec. 4.2. They are missing one of

background	comment on generated samples
WW	WW + 2 partons
WZ	WZ
Drell-Yan	$DY \rightarrow ee$ + 2 partons $DY \rightarrow \mu\mu$ + 2 partons $DY \rightarrow \tau\tau$ + 2 partons
FAKES	W$\rightarrow e$ + (2,3,4) partons W$\rightarrow \mu$ + (2,3,4) partons W$\rightarrow e$ + $(c\bar{c}/b\bar{b})$ + 2 partons W$\rightarrow \mu$ + $(c\bar{c}/b\bar{b})$ +(1,2) partons

Table 7.3: The samples used for obtaining background templates. In comments on samples, there are listed also number of additional partons generated within a given sample.

the signatures of $t\bar{t}$ events (Drell-Yan events don't have neutrinos in final state, while FAKE events have only one real lepton in final state).

Because of low statistics in these official samples, we were not able to create reasonable templates (we obtained < 100 events in templates), so we have released some of the event selection cuts for Drell-Yan and FAKE samples in order to get higher statistics for templates in question [3] Of course, we needed to perform the comparison of kinematic variables between events selected with original selection and events chosen by the released selection criteria to prove that the events are kinematically same. The whole procedure of releasing the cuts and the comparison is described in Appendix C.

The templates for $WW+WZ$, Drell-Yan, and FAKE events are in Fig. 7.5. We can see that the reconstructed mass (mean value of distribution) for background samples is in the following order $M_{Fakes} \simeq M_{Diboson} < M_{Drell-Yan}$. This is expected, since their sum of final state transverse energy (denoted as H_T) is also in the similar order (see Fig. 6.22 or [89]) and usually the events with higher H_T have higher reconstructed mass (see Fig. 5.15).

7.2.1 Combining backgrounds

When combining backgrounds together, one has to take special care of the way in which one combines different contributions. The number of expected background events in 340 pb^{-1} of data together with probability of reconstruction for each type of background is shown in Table 7.4.

---

[3] Of course, there is always possibility of producing more and more MC samples. However, as we mentioned above, both FAKE and Drell-Yan events pass event selection due to mismeasurements. Such mismeasurements are rare, and thus the number of generated events would need to be huge.

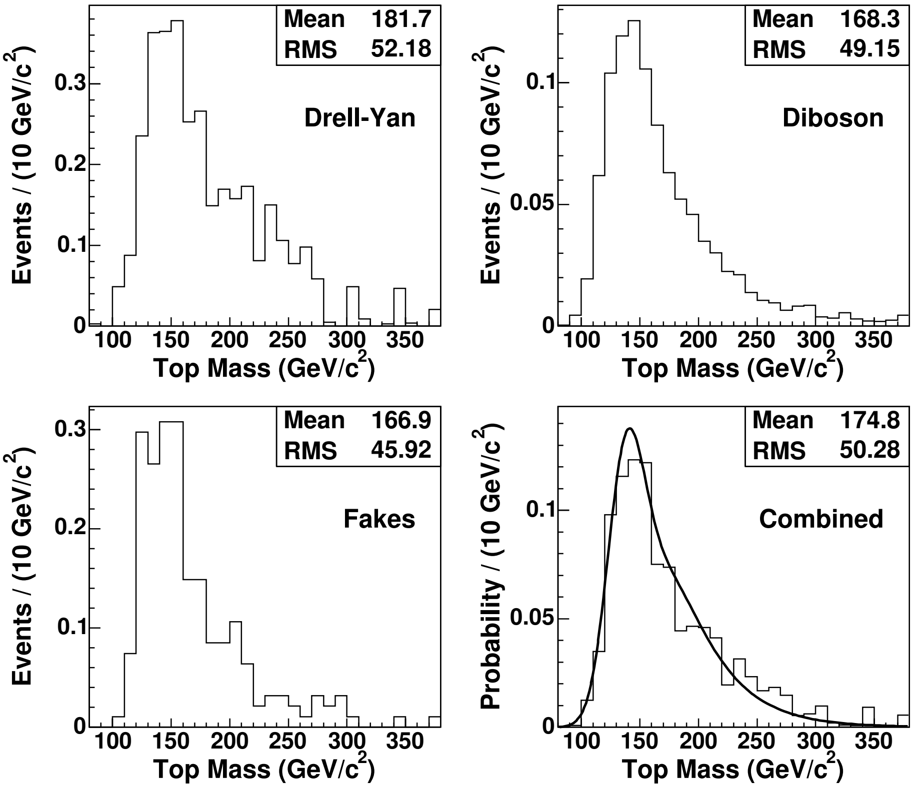


Figure 7.5: Reconstructed top quark mass templates for the Drell-Yan, Diboson, and Fakes background sources along with the combined background template and associated fitted probability density function. Background sources are normalized to the expected contribution in 340 pb^{-1} sample.

Each of the backgrounds have to be included proportionally to number of expected events but also to the probability of reconstruction of each type of background. The reason is the following: Consider for simplicity just two types of backgrounds. Let's say WW has 100% probability and DY has 0% probability of reconstruction. Then, as background shape you would use just WW shape , because in data you will have reconstructed only WW events and no DY events, so you want fit WW data to WW template and not to some combination of WW and DY. In case, DY has non zero probability of reconstruction you must do weighting according to the probabilities. Extension to more than two types of background is trivial.

After scaling background contributions, adding them together and normalizing the combined distribution we obtain the template for combined background (Fig. 7.5). The combined background template is also fitted by a combination of Landau + Gauss distribution, the similar way as $t\bar{t}$ signal templates are fitted, see Eq. (7.1). However, linear dependence on top mass has no sense in this case, so the fitted parameters are

background	expected num. of events	recon. prob.
WW/WZ	1.6 ± 0.3	0.65 ± 0.02
Drell-Yan(e, μ, τ)	5.5 ± 1.2	0.62 ± 0.23
FAKES	3.5 ± 1.4	0.60 ± 0.07
total	10.5 ± 1.9	0.61 ± 0.12

Table 7.4: The number of expected events in 340 pb^{-1} and reconstruction probability for each background.

parameter	value
p_1	138.4 ± 2.4
p_2	13.5 ± 1.8
p_3	0.7 ± 0.1
p_4	200.4 ± 15.2
p_5	-67.1 ± 15.2

Table 7.5: The parameters of the fit to the combined background template.

just constants, i.e. $p_i = a_i$ and are shown in Tab. 7.5. The Kolmogorov-Smirnov test of compatibility between distributions gives probability of about 85 % which means very good agreement.

7.3 Summary of expected number of reconstructed events

As it was mentioned in previous chapter, not all events pass mass reconstruction procedure, i.e. for not all events is mass reconstructed. Here, we just summarize the expected number of reconstructed signal and background events in data. As we saw in Tab. 4.6, there are 17.2 ± 1.4 $t\bar{t}$ signal events expected[4] in 340 pb^{-1} of CDF data sample. Altogether, 10.5 ± 1.9 background events is expected to pass the event selection. The reconstruction probability for signal events is about 75 %, see Fig. 7.2 and for combined background, it is about 61 % (Tab. 7.4). In Tab. 7.6, we summarize the expected number of events passing event selection, reconstruction probabilities and expected number of events passing mass reconstruction. The numbers for total background in last column of table will be used in final top mass estimate by log-likelihood minimization, see Sec. 6.2.3.

---

[4] this is expected in case top quark mass is 175 GeV

event type	N_{selec}	P_{recon}	N_{recon}
WW/WZ	1.6 ± 0.3	0.65 ± 0.02	1.0 ± 0.2
Drell-Yan(e, μ, τ)	5.5 ± 1.2	0.62 ± 0.23	3.4 ± 1.5
FAKES	3.5 ± 1.4	0.60 ± 0.07	2.1 ± 0.9
total background	10.5 ± 1.9	0.61 ± 0.12	6.4 ± 1.2
$t\bar{t}$ ($M_{top} = 175$ GeV)	17.2 ± 1.4	0.75 ± 0.004	12.9 ± 1.1
total	27.7 ± 2.3	0.70 ± 0.05	19.3 ± 2.0

Table 7.6: The number of expected events in 340 pb^{-1} after event selection (N_{selec}), reconstruction probability for each type of events (P_{recon}) and expected number of events which pass mass reconstruction (N_{recon}).

So, finally we have templates for signal and background events together with the functions which describe them. Moreover, we also estimated the expected number of events, for which will be reconstructed top mass. This is everything which was missing (see Eq. (6.10)) in order to get final top mass estimate using data sample collected at CDF experiment.

However, before jumping on data, we will test our method whether it gives reasonable outcomes. These tests are subject of next chapter.

Chapter 8

Testing the method by pseudo-experiments

We have put together whole machinery for top mass reconstruction in previous two chapters. We can now use real CDF experiment physics data to estimate top quark mass. However, before we will do that, we should prove that our method gives unbiased estimates of both top quark mass and uncertainty on top mass. We can perform such tests on MC events, where we know what is the input mass. Moreover, if we do the testing for the number of the events we expect (or see) in data, we can obtain the estimate of top mass uncertainty which we should expect in the data events. Later, we can compare it with the actual value obtained in data.

We do the testing of our mass reconstruction procedure by performing the "pseudo-experiments" (PEs). The principles of pseudo-experiments together with its strengths and weakness will be presented in next section. Later, we will show the actual results we obtained using PE on MC samples.

8.1 Principles of pseudo-experiments

We choose certain number of events from the MC sample and for this set we reconstruct the top mass the same way as we will do it for the set of data events. This is one pseudo-experiment (PE), analogue of hardly built physics experiment. The top mass for the set of events is estimated by likelihood minimization, see Sec. 6.2.3. As already mentioned above, to be able to compare PEs to result from data, we will use in one set such number of events which is expected in data (this is ~ 13 signal $t\bar{t}$ events and ~ 6 background events, as we will see below). We can repeat this procedure many times, always with different set of events and thus perform many pseudo-experiments. The actual testing can be done by, e.g. obtaining the distribution of reconstructed mass which should be consistent with the generated mass (e.g. Most Probable Value - MPV is consistent with generated mass).

In principle, if we would have big enough MC sample, we would divide the whole sample into set of dis-joined set of events, so each event in all PEs is unique. This would be the fair testing by using the PEs. However, doing the PEs this way would require, in our case, approximately $13 \times 10000 = 130000$ events in one $t\bar{t}$ MC sample (number of expected $t\bar{t}$ events multiplied by the number of PEs we usually use). However, the available MC samples have much less events. In one MC sample, there is ~ 2000 events, see Sec. 7.1. Therefore, we perform the pseudo-experiments other way, using the re-sampling method [90], which is widely used in high energy physics.

Using re-sampling method, we choose randomly the set of events for one PE from the pool of all events in a given MC sample (in our case, ~ 2000 events, see Tab. 7.1). In principle, we can repeat this procedure arbitrary many times and thus create arbitrary many PEs. Clearly, the advantage of the re-sampling procedure is that with statistically relatively low MC sample, we can create many PE, so we don't need to generate huge MC samples. However, there are some limits of usability of re-sampling method. In such procedure, some events can be present in pseudo-experiments more than once, even when we make sure there are all events different in one particular set of events used in one PE (we do this). Therefore, performing the PEs in such way, we will have some correlation between the sets of events, because some PEs will share the same events (or at least some of them).

The three important numbers regarding the PEs are the total number of events N_{sample} in a given MC sample from which will be randomly chosen the events for PEs, number of events in one PE N_{PE}^{ev} and the total number of pseudo-experiments N_{PE}. In our case, the $N_{sample} \sim 1100 - 1800$, see Sec. 7.1, so we are interested what are possible limitations on N_{PE}^{ev} and N_{PE}.

Clearly, increasing number of PEs after certain number of PEs performed should have no effect. Otherwise, we would be able to infinitely improve the precision of our parameter estimation with the definite number of events in the MC sample. It was shown in [90], that we can use up to $(N_{sample}/N_{PE}^{ev})^2$ PEs. Since in our case, $N_{sample} \sim 1100 - 1800$ while $N_{PE}^{ev} \sim 13$, we can use about 10000 PEs. Using more PEs have no advantage in better parameter or error of the parameter estimation, so its useless.

We performed the tests with different number of events in one PE. The MC sample used for these tests had input mass 170 GeV. We used $N_{PE}^{ev} = 20$ up to $N_{PE}^{ev} = 500$ events in one PE. The dependence of pull width (for definition, see below) on number of events in one PE is plotted in Fig. 8.1. We can see that the pull width have

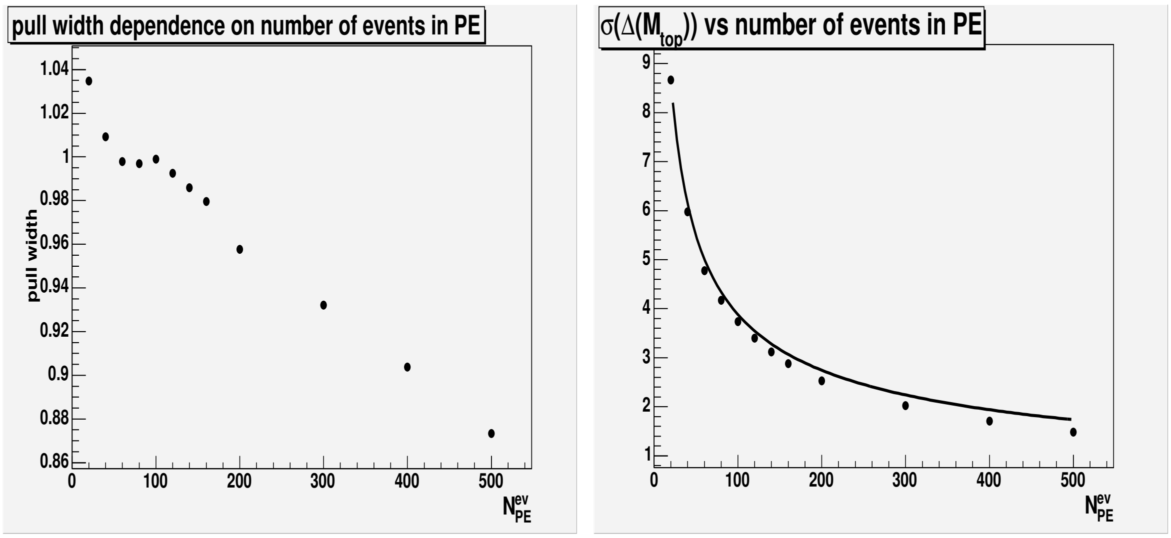


Figure 8.1: The pull width (left) and the width of difference between input and output mass distribution (right) dependence on number of events in one pseudo-experiment.

tendency to decreases more and more from correct value (~ 1) as the N_{PE}^{ev} increases. This could mean that the errors estimated within PE are overestimated or the spread of the reconstructed top mass values is underestimated. The dependence of width of $\Delta M_{top} \equiv M_{input} - M_{output}$ distribution on number of events in one PE is shown on second plot in Fig. 8.1. From this plot, we can see that there is once again tendency that decreasing of the mass spread goes faster than one would expect according statistical prediction ($\sigma \sim 1/\sqrt{N_{PE}^{ev}}$, which is shown as curve on the figure, where the curve is normalized for uncertainty obtained for $N_{PE}^{ev} = 20$). Therefore, we conclude that the effect of smaller pull width is due to artificially smallness of the spread. The natural explanation is that the correlations between particular PEs start to play role. It seems, we can safely use up to $\sim 100-150$ events in one PE. After that, the correlation seems to have effect. In such cases ($N_{PE}^{ev} \sim 100$), one particular event is used in average in about 5 % of all PEs (assuming 10000 PEs), so at such level, the correlation between PEs can start to show up.

The actual way how we perform PEs is as follows. Average number of events in PEs is given by expected number of signal and background events after mass reconstruction, see Tab. 7.6. These are 12.9±1.1 signal events and 6.4±1.2 background events. For one given PE, we use the number of events N_{PE} generated according Poisson distribution with mean value of expected number of events. We do this for signal and background events. Then, we randomly choose $N_{PE}^{t\bar{t}}$ signal events from pool of all events in given

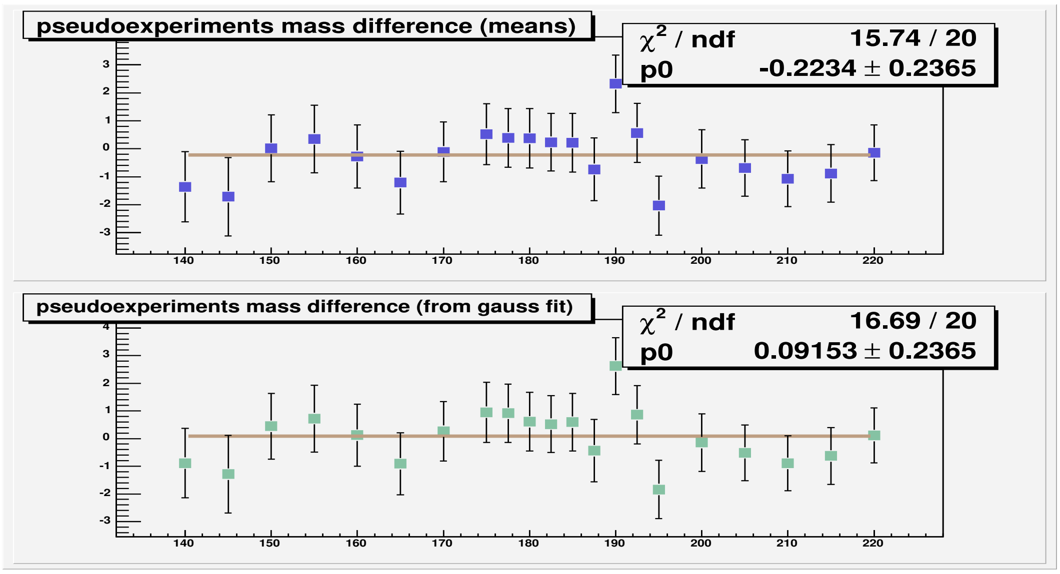


Figure 8.2: The difference between generated (input) mass and reconstructed mass from pseudo-experiments as a function of input top mass. Here, only signal events were used. In the upper plot, the mean values of mass distributions are used, in the lower one the mean of Gauss fit to the distribution is used.

MC sample. For background, we randomly choose the mass value from combined background template, see Fig. 7.5. We reconstruct the mass for such selected set of events (one PE) using likelihood minimization, see Sec. 6.2.3. We get estimate on both top mass and its uncertainty (positive and negative) from likelihood minimization for a given PE. We repeat the above described procedure 10000 times, so we get 10000 PEs. We create distributions of top mass estimates, errors on top mass estimates and also the pull[1]. As a top mass estimate from PEs we use the most probable value (MPV) of the mass distribution (more precisely, we use mean value of Gaussian fit to the mass distribution).

8.2 Testing on MC samples

First, we performed the pseudo-experiments for $t\bar{t}$ signal events only. The difference between reconstructed and generated mass as a function of generated mass is shown in Fig. 8.2. The plot shows that we are unbiased in average over whole range of input mass.

---

[1] The pull is defined as $\frac{m_t^{fit}-m_t^{orig}}{\sigma(m_t^{fit})}$, where $\sigma(m_t^{fit}) = (\sigma_+ + \sigma_-)/2$, with σ_+,σ_- being positive and negative error on top mass, respectively.

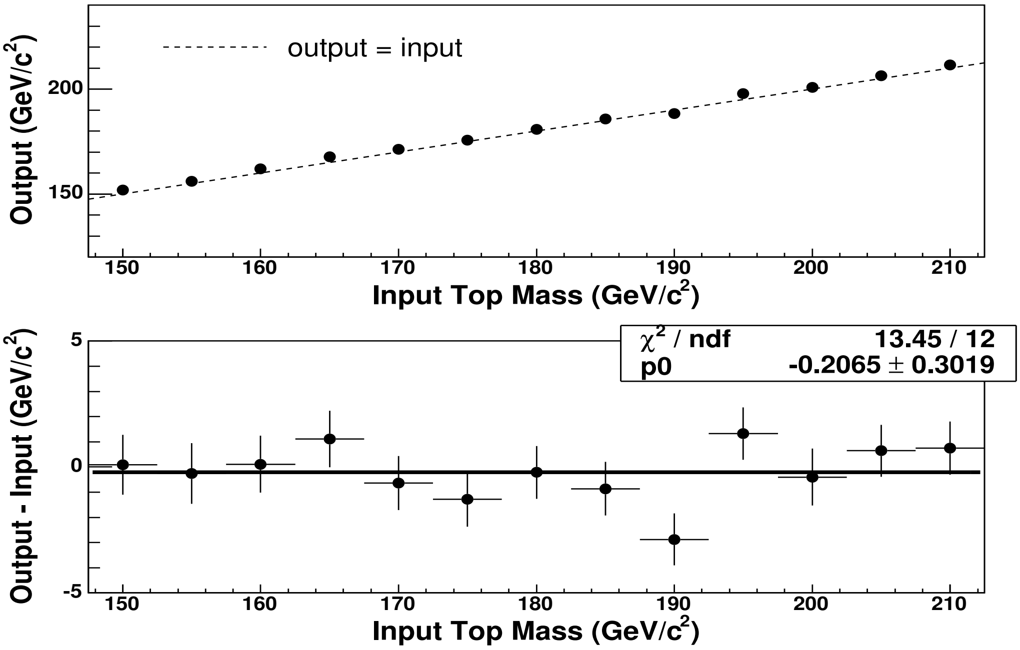


Figure 8.3: The upper plot shows mean of the output (measured) top mass as a function of the input (generated) mass, while the lower plot gives the difference between output and input top mass as a function of the input mass (the constant fit is also shown). The background was included in these pseudo-experiments.

After that, we perform the pseudo-experiments with background events included. The reconstructed mass as a function of the input top mass is summarized in Fig. 8.3. It can be seen that the reconstructed mass follows the input mass. The linear fit returns offset statistically consistent with zero (1.30 ± 2.09) and slope consistent with one (0.993 ± 0.011). Moreover, there is shown the difference between input mass and reconstructed mass as a function of input mass. It can be seen, that we are consistent with no bias within statistical uncertainty.

The error on each point in Fig. 8.2 and Fig. 8.3 was determined in the following way, in order to avoid oversampling of the MC sample. We used one sample with generated top mass of 175 GeV. We estimate the error of top mass median value by executing PEs with no re-sampling (randomly chosen unique events), i.e. one MC event is used once and only once in all PEs. As we have in MC sample with 175 GeV generated top mass only about 1500 reconstructed events (see Tab. 7.1) and in average, there are about 13 signal events in one PE, we can create only about 110 different sets, i.e. only about 110 PEs without doing re-sampling. From distribution of PE top mass, we take median as the estimate of top mass for a given MC sample. Repeating this procedure 100 times

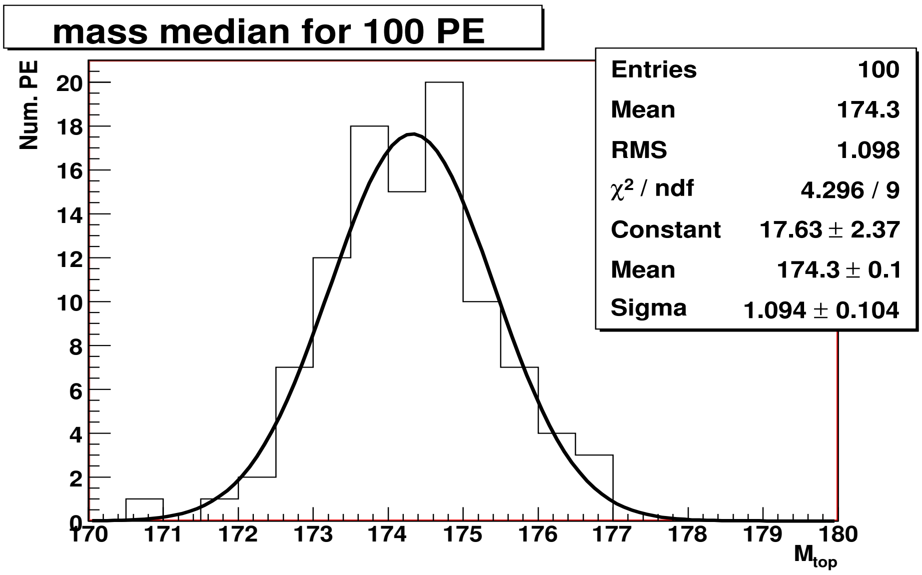


Figure 8.4: The distribution of mass distribution medians from 100 different sets of pseudo-experiments performed without re-sampling.

(always with different seed for random generator,so we always create different sets of events), the histogram of medians is obtained (with 100 entries), see Fig. 8.4. The width of this distribution is an estimate of the uncertainty on median, i.e. 1.09 GeV. This is our uncertainty on mass estimate due to limited statistics of MC sample. The uncertainties on mass estimated for samples with generated mass other than 175 GeV were determined by scaling the 1.09 according number of events in other templates, i.e. by a factor $\sqrt{(N_{175}/N_{mass})}$, where N_{175} is the number of events in MC samples with generated top mass of 175 GeV and N_{mass} is the number of events in MC sample with generated top mass of "mass" (see Tab. 7.1).

As an example, we show the pseudo-experiments estimated top mass, estimated error (average of positive and negative error estimate), and the pull distributions for some of the input masses in Fig. 8.5. Both mass and error distributions look reasonable which is also proved in pull distributions. Pull distributions have correct Gaussian shape and the width around one, as it should be according the definition. There are a little larger tails in pull distributions for a large input masses.

The pull mean and pull width dependences on input top mass are in Fig. 8.6. The pull mean values looks reasonable well (they are distributed around zero). The pull widths seems to be smaller at higher mass which is consequence of not very good Gaussian fit to the pull distributions for higher mass (see pull distribution for 220 GeV input top mass in Fig. 8.5). Overall, pull widths are a slightly shifted from one (in

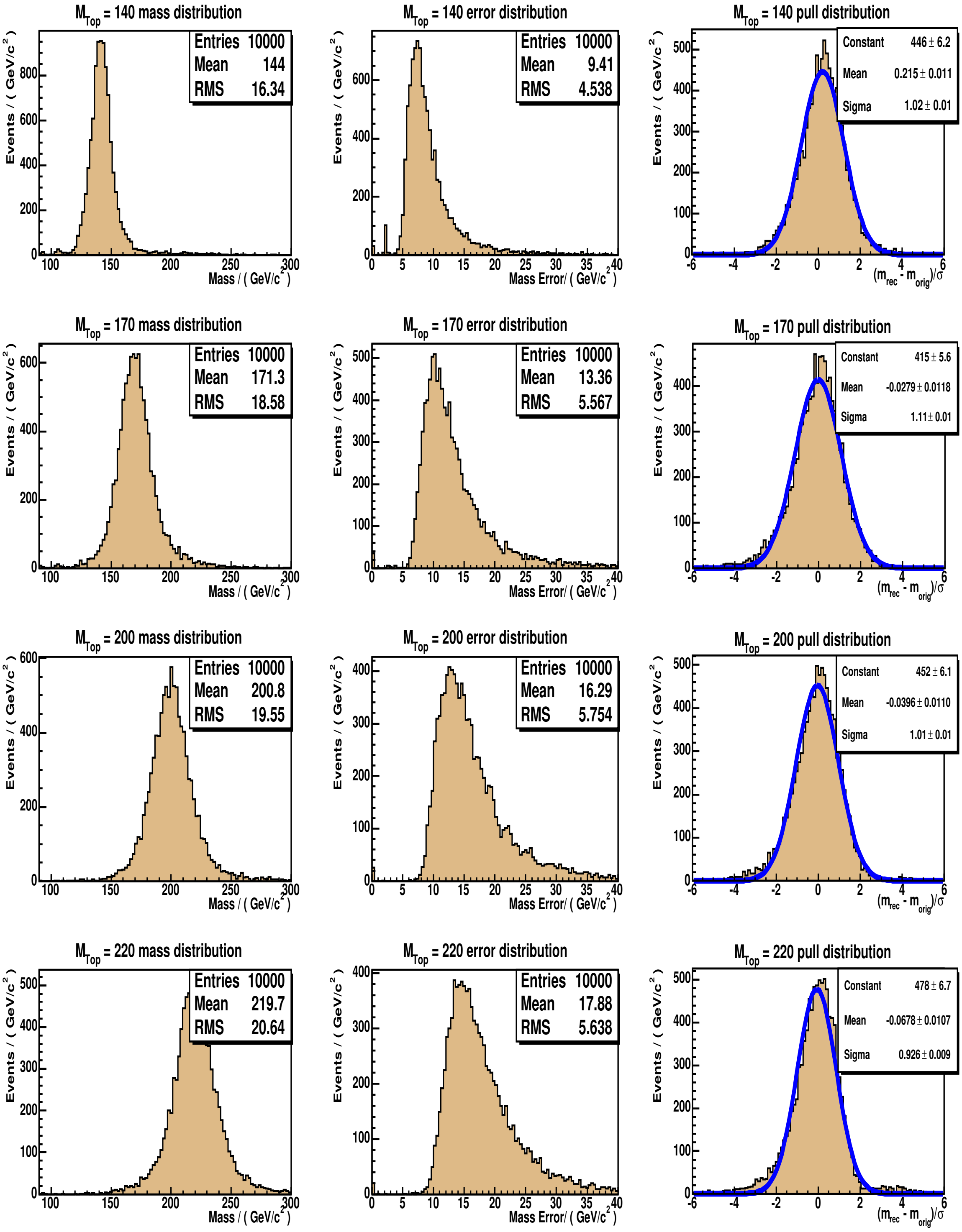


Figure 8.5: Pseudo-experiments for MC samples with input top mass 140, 170, 200 and 220 GeV. The mass, error and pull distributions are presented.

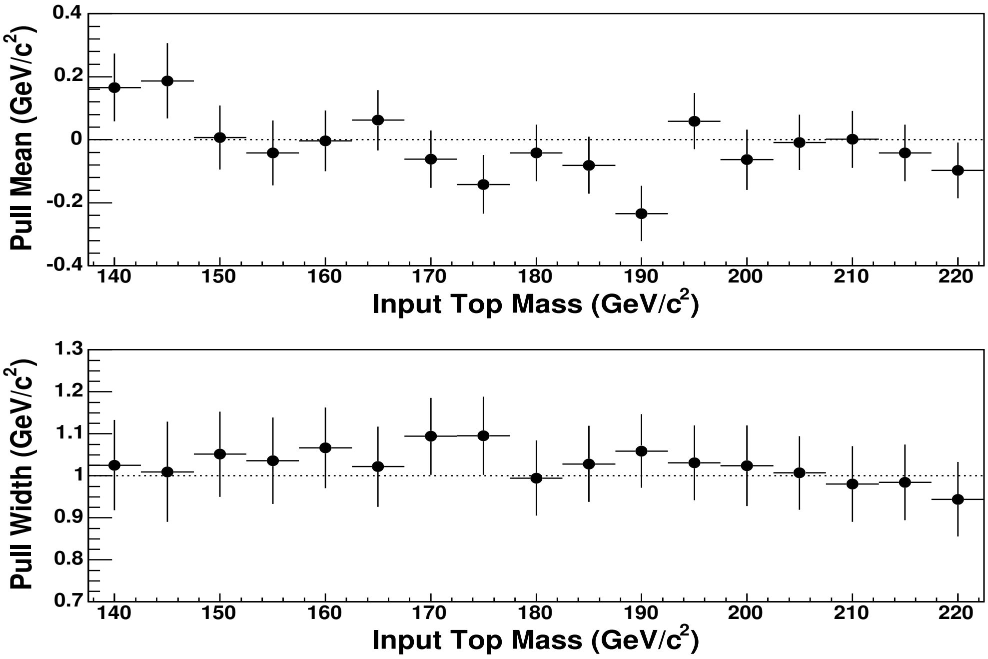


Figure 8.6: Summary of pull distributions for the pseudo-experiments, showing the pull mean (upper) and width (lower) as a function of generated (input) top mass, compared with zero mean and unity width (horizontal lines).

ideal case, the distributions should have width of one by definition). In average, it is 1.033, so this scale factor we will apply also to our error estimate we will obtain in data events.

We calculated the “a priori” statistical uncertainty of our method for the M_{top} = 178 GeV Herwig sample[2]. The expected statistical uncertainty using our method for number of events corresponding to 340 pb^{-1} of integrated luminosity is 15.1 GeV (mean value of error distribution multiplied by correction 1.033), see Fig. 8.7. Using MPV of the distribution, the error estimate is about 12.5 GeV.

Until now, we have tested our method and saw it gives reasonable estimates. However, all these tests very not completely independent. It’s because, we performed the tests on the same samples which we used for creating the templates. To be really sure that our method is correct, we need to do the tests on some other, independent samples. We performed also such kind of test. Within CDF collaboration, there were generated five, so called, “blind samples” [3]. We reconstructed the “raw mass” and performed the

---

[2] as already mentioned, the value 178 GeV was world top mass average prior to start of this measurement

[3] We did not know what were the original top masses of these samples. Actually, we still don’t know them, as they were generated by CDF top mass group conveners in order to test each method

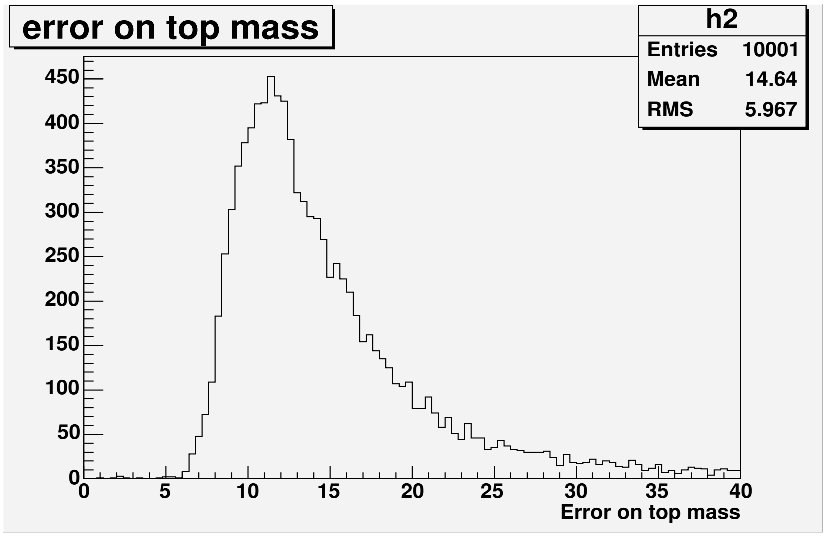


Figure 8.7: The uncertainty on top mass from pseudo-experiments for MC sample with input mass 178 GeV.

$M_{fit} - M_{true}$ (GeV)
$0.2 \pm 1.2\ P$
$1.6 \pm 1.2\ P$
-2.1 ± 1.2
-1.1 ± 1.2
-0.3 ± 1.2

Table 8.1: The difference between reconstructed and generated top mass for the "blind samples". The first two samples marked as 'P' are Pythia samples, the rest are Herwig samples.

pseudo-experiments for these samples using the original templates parametrization, see Tab. 7.2. The results are presented in Tab. 8.1. It shows, that our results are unbiased within the errors.

Finally, we have all the method ready and also tested that it gives correct results. Now, we can go and put our hands on data, to see what comes from it.

---

of top mass reconstruction at CDF.

Chapter 9

Systematic uncertainties on top quark mass determination

Apart from the statistical uncertainty on the measured top quark mass due to the limited size of our data sample, there are several sources of systematic uncertainty. Each measurement which use some assumptions about any input variable and/or which use any correction to measured quantities is a subject to systematic errors.

Top mass determination by our method is sensitive to the Monte Carlo templates as well as to the energy scale of jets. We have studied the effects of these sources of the uncertainties on the top mass.

We've done this the following way. We changed particular source of the uncertainty (e.g. correction to the jet E_T, amount of the initial state radiation, type of parton distribution function) within reasonable range. Where it was appropriate, we changed the value by its uncertainty. However, this was not always possible. Then, we looked at the change in reconstructed top mass, which was invoked by this shift from the original value of particular variable. At the end, we considered this shift in reconstructed mass as an estimate of the systematic uncertainty on top mass due to the particular source of uncertainty.

Technically, we performed these estimates by pseudo-experiments studies. How we do the pseudo-experiments is described in detail in Chap. 8. First, we performed the mass reconstruction for the events in MC sample, where the systematic shift of some kind of variable was applied (e.g. events where jet E_T was shifted by uncertainty on jet E_T). Then, we performed large number (10000) of pseudo-experiments for these "shifted" events. In these pseudo-experiments, we always used the global parametrization of templates we obtained from nominal MC signal and background samples, see Tab. 7.2, Tab. 7.5. After we performed pseudo-experiments, we obtained distribution of reconstructed mass. In these studies, we use the most probable value of mass distribution as the top mass estimate for a given sample. Moreover, we made also the cross-check using median values and the results were consistent. At the end, we compared the reconstructed top mass between the sample of original events and the sample of events where some kind of systematic shift was applied. Because of the shift of some kind of input, we get also shift of the output (i.e. top mass) comparing to the nominal non-shifted reconstructed top mass. Such shift in reconstructed top mass from pseudo-experiments we then consider as estimate of systematic uncertainty.

Below, we will describe all the sources of systematic uncertainty which we considered and will present the results we obtained.

9.1 Jet Energy Scale Corrections

The energy deposited in calorimeter which is by some algorithm assigned to jet (so called "raw jet energy"), must be corrected for a few reasons, see Sec. 4.1.4. These corrections scale the "raw jet energy" to the particle level and eventually to parton level. Therefore, we are talking about jet energy scale(JES). Of course, the corrections have their own uncertainties. Not knowing precisely what our corrections really should be (for MC and for data), we are dealing with systematic effects which we try to quantify.

There exists a few different corrections to JES as described in Sec. 4.1.4. These are η dependent corrections, multiply interactions corrections, absolute jet energy scale corrections, underlying event corrections and out of cone energy corrections. We estimated the top mass uncertainty due to each of these corrections.

We performed the studies using Herwig $t\bar{t}$ sample with the generated top mass of 178 GeV. From the nominal MC sample, we created new MC samples, where the jet energies were shifted by the ± 1 standard deviation (uncertainty) of each JES correction. For such "shifted" samples, we reconstructed raw mass and obtained "shifted" template. Then, we performed pseudo-experiments (using the default parametrization of templates) for these "shifted" samples. The systematic error due to each individual JES correction is taken to be the half of the reconstruction mass difference between the sample with $+1\sigma$ ($M_{top}^{+1\sigma}$) and -1σ shift ($M_{top}^{-1\sigma}$) in jet energy correction applied[1].

The summary of the systematic uncertainties due to different JES corrections is in Tab. 9.1. As expected, largest uncertainties are due to absolute scale corrections and out-of-cone corrections, because these are the largest corrections to JES, see Sec. 4.1.4.

---

[1] we use half of the difference, because the difference between $+1\sigma$ and -1σ mass estimate corresponds to 2σ difference in reconstructed mass

source	$M_{top}^{+1\sigma}$	$M_{top}^{-1\sigma}$	$(M_{top}^{+1\sigma} - M_{top}^{-1\sigma})/2$ (GeV)
η-depend. calib.	178.21	176.89	0.7 ± 0.3
absolute scale	179.65	175.53	2.1 ± 0.3
underlying event	177.98	177.24	0.4 ± 0.3
out-of-cone	180.11	175.55	2.3 ± 0.3
splash-out	178.07	177.24	0.4 ± 0.3
Total			3.2 ± 0.7

Table 9.1: The difference in reconstructed mass between sample with increased ($M_{top}^{+1\sigma}$) and decreased ($M_{top}^{-1\sigma}$) jet energy scale for each jet energy scale correction. The total systematic uncertainty on top quark mass due jet energy scale is the sum of all contributions in quadrature.

The total uncertainty of 3.2 GeV is obtained by summing partial uncertainties in quadrature, assuming they are uncorrelated.

The above jet energy corrections are developed from studies of samples dominated by light-quark and gluon jets. Since top mass also depends on energy response to b quark jets, additional uncertainty occurs from extrapolating this procedure to b-quarks The resulting systematic effect on jet energy is considered to stem from three main sources of uncertainty: uncertainty in energy response arising from uncertainties in the decay properties of bottom quarks (semileptonic decays), uncertainty arising form the imperfect knowledge of the fragmentation properties of bottom quarks, and uncertainty arising from the different color flow associated with bottom quark jets produced in top quark decay. It was found before in [91], that these sources are responsible for about 1 % change in b-jet energy scale.

We evaluated the b-jet energy scale uncertainty by using only events, which have two jets matched to original b-quarks. The jet matches b-quark if its direction is within cone $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ of the direction of b-quark. Applying the 1% increase of the jet energy to such events and making the reconstruction, we obtain the shift in reconstructed raw mass and thus a new template. Once again, performing the pseudo-experiments with "shifted" events while still using nominal parametrization of templates cause the bias in reconstructed mass in pseudo-experiments. We obtained the shift 0.6 ± 0.6 GeV comparing to PE with original events. This shift we take as systematic uncertainty due to b-jet energy scale.

9.2 MC generators

We use Herwig to generate most of the signal $t\bar{t}$ mass samples and some other generators (Pythia, ALPGEN) to generate background samples. However, each generator use different hadronization model in order to calculate the particular processes and the results (e.g. cross-section estimate, jet P_T distributions,...) between generators for the given process slightly differ! Therefore it is needed to estimate what is the effect of using different generators on our top mass measurement.

For this purpose, we use 2 generators, Pythia and Herwig. The $t\bar{t}$ samples with input top mass 178 GeV are generated by both MC generators. The difference between reconstructed top mass from both samples using pseudo-experiments studies is assigned to be systematic uncertainty due to using different MC generators. The difference is 0.4 ± 0.6 GeV (see Tab. 9.2). To be conservative, we use the 0.6 GeV as the systematic uncertainty.

9.3 Initial state radiation

Extra jets originating from the incoming partons and/or outgoing partons affect the measurement of top mass. These jets can be misidentified as jets coming from b-quarks or they can simply change the kinematics of the final state partons.

The initial state radiation (ISR) and final state radiation (FSR) are controlled by the same DGLAP evolution equation that tells us the probability for a parton to branch [92].

ISR is studied using Drell-Yan events because they have advantage of no FSR and they are produced by the $q\bar{q}$ annihilation process, as are most ($\sim 85\%$) of $t\bar{t}$ pairs. The amount of ISR is measured as a function of the Drell-Yan mass scale and shows logarithmic dependence on the Drell-Yan mass squared, as shown in Fig. 9.1.

Based on this measurement, two ISR systematic $t\bar{t}$ Monte Carlo samples ($+1\sigma_{ISR}$ and $-1\sigma_{ISR}$ with input top mass 178 GeV) are produced using PYTHIA, by changing the QCD parameters for parton shower evolution, in order to cover the region given by uncertainties in measured average p_T of Drell-Yan events (the corresponding curves of Drell-Yan dilepton p_T vs. invariant mass squared are shown in Fig. 9.1). Specifically, the value of Λ_{QCD} and scale factor, K to the transverse momentum evolution scale $k_{\perp}^2$ for ISR showering are varied. The values of parameters used are Λ_{QCD} (5 flavors) $=$ 292 MeV, $K = 0.5$ for sample with more ISR ($+1\sigma_{ISR}$) and Λ_{QCD} (5 flavors) $= 73$ MeV, $K = 2$ for sample with less ISR ($-1\sigma_{ISR}$). The values of parameters used in nominal sample were Λ_{QCD} (5 flavors) $= 146$ MeV and $K = 1$.

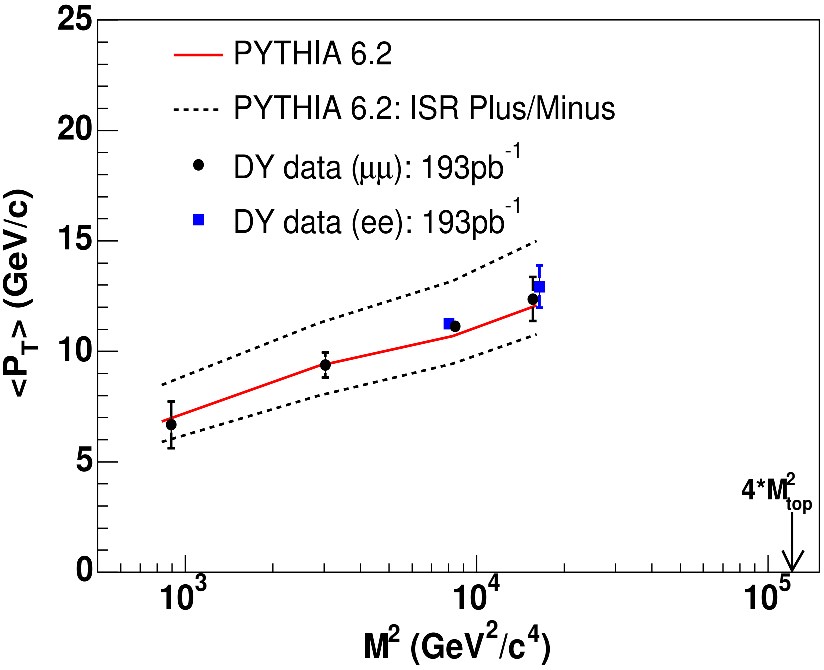


Figure 9.1: The average p_T of the dilepton system, which corresponds to the level of ISR activity, shows a logarithmic dependence on the dilepton invariant mass M_{ll}^2. The data are compared with the predictions of PYTHIA 6.2 and of the $+1\sigma_{ISR}$ and $-1\sigma_{ISR}$ samples [91].

By doing top mass reconstruction on these "shifted" samples, we obtain shift in reconstructed mass comparing to the mass from nominal sample. The results we obtained are summarized in Tab. 9.2. We use half of the difference between results obtained for samples with less and more ISR, 0.6 GeV, as the systematic uncertainty due to ISR.

9.4 Final state radiation

Since ISR and FSR shower algorithms are the same, the same variations in Λ_{QCD} and K as in ISR are used to generate FSR systematic samples by varying a set of parameters specific to FSR modeling.

Naively, more FSR means more radiated gluons out of final state b quarks. This means, the reconstructed energy of b-jets will be smaller and consequently will be smaller also reconstructed top mass.

Once again, we used two samples (generated by Pythia), one with more FSR and one with less FSR. The results are presented together with previous systematics in

Source	sample	M_{top} (GeV)	ΔM_{top} (GeV)
Generators	Herwig	178.1	-0.4 ± 0.6
	Pythia	178.6	
ISR	more ISR	177.3	-0.6 ± 0.3
	less ISR	178.5	
FSR	more FSR	177.8	-0.3 ± 0.3
	less FSR	178.1	

Table 9.2: The reconstructed mass from pseudo-experiments for different MC samples together with assigned systematic uncertainties on the top mass due to different generators, ISR and FSR.

Tab. 9.2. We can see that the reconstructed top quark mass is smaller for sample with more FSR. This is in agreement with our expectation mentioned above. As an estimate of systematic uncertainty due to FSR, we use the difference between reconstructed mass for sample with more and less FSR, respectively. This is 0.3 GeV.

We examine also the effects of higher order corrections to $t\bar{t}$ production using MC@NLO [93], a full NLO Monte Carlo generator. We found a shift 0.2 GeV in reconstructed mass when using MC@NLO sample comparing to our original LO sample. We consider this effect as negligible and that it is already covered by the ISR/FSR systematics.

9.5 Parton distribution functions

The calculation of the top quark invariant mass does not depend directly on the choice of input parton distribution function (PDF). However, changing the PDF changes also the top quark η and p_T distributions as well as the size of ISR. This results in a change in the jet p_T distributions and in the probability of selecting the correct jets, both of which affect the reconstructed top quark mass.

For generating MC samples, we use a parameterization of quark and gluon PDFs. There exists two main sets – CTEQ [94] and MRST [95] which are based on fits to data from a number of experiments (including CDF). Therefore, the systematic uncertainty due to PDF is composed from uncertainty within CTEQ set (this does not include uncertainty in Λ_{QCD}), difference between CTEQ and MRST and uncertainty from different Λ_{QCD} by comparing two different MRST sets. The uncertainty in the PDF for a CTEQ set have been parameterized as a set of 20 independent eigenvectors (with a positive and a negative variation).

The effect of different PDFs was studied by using a re-weighting method [96]. Instead of generating a different MC samples with the different PDF sets (which would be quite time consuming), we instead use just a single simulated MC sample (178 GeV Pythia sample). The mass templates are then generated for the different PDF sets by weighting events according to the probability for that event to proceed (probability to observe incoming partons) according to the given PDF (relative to the default PDF). This technique also removes most of the uncertainty due to limited Monte Carlo statistics.

We generate 46 templates, where the PDFs used are:

0 - CTEQ5L Default PDF. The relative weight for every event is 1.

1 - MRST72 Leading order PDF using more or less the same data as our default PDF but fitted by different group. No significant difference is expected relative to the default, but if difference is found, it should be taken as a systematic.

2 - MRST75 Same as previous, but using different value of α_s, corresponding to $\Lambda_{QCD} = 300$ MeV (comparing to $\Lambda_{QCD} = 228$ MeV in previous PDF). The difference between these two PDFs is taken as a systematic.

3 - CTEQ6L More recent fit from CTEQ group, still leading order.

4 - CTEQ6L1 Same as CTEQ6L, but using LO running α_s (CTEQ6L use NLO)

5 - CTEQ6M This is the next-to-leading order PDF from the CTEQ group. Using a leading order matrix element, as the one found in HERWIG, in conjunction with this PDF will not give a correct description of event rates. However, we assume in the following that the *relative* behavior of various NLO PDFs is accurately modeled even through we use a LO matrix element.

6-45 - CTEQ6M uncertainty PDFs These PDFs encode the uncertainty in the CTEQ6M PDF. The possible variations in CTEQ6M PDF are separated into independent contributions from 20 eigenvectors. For each eigenvector there exists "+1σ" and "-1σ set (for example, the sets 44 and 45 belong to eigenvector 20).

We perform pseudo-experiments with each event weighted accordingly for each PDF in order to estimate the systematics. We produced pseudo-experiment mass distributions (e.g. see Fig. 8.5) and take the MPV as the estimate of reconstructed mass for a given PDF sample.

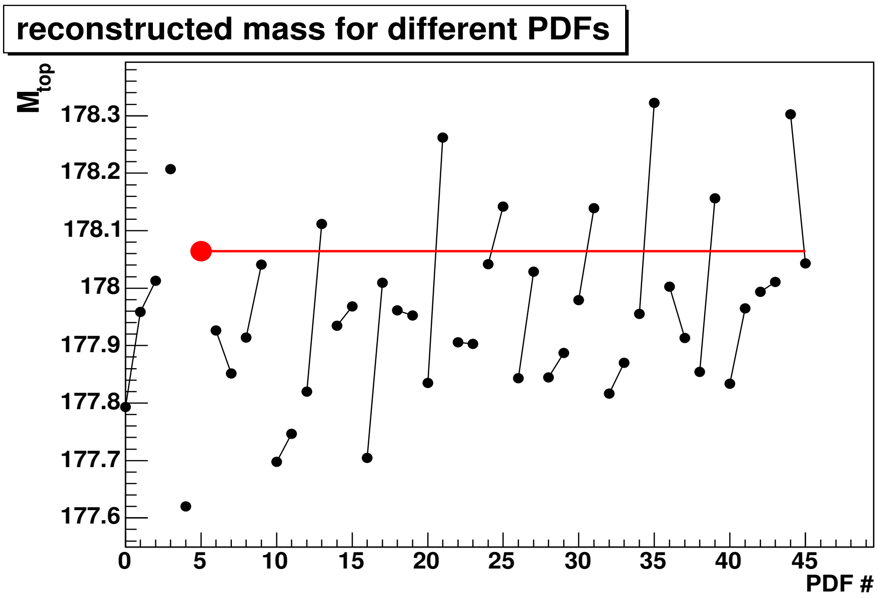


Figure 9.2: The reconstructed mass as a function of PDF sample. The PDF numbers are the ones we assigned to the samples in this section (see text). The big bullet corresponds to CTEQ6M. Paired are the points which corresponds to $+1\sigma$ and -1σ variation in a given CTEQ6M eigenvector.

These reconstructed masses for each PDF are shown in Fig. 9.2. Adding half the difference between each of the 20 CTEQ6M pairs in quadrature[2] gives a 0.441 shift in the reconstructed top mass due to the uncertainty in CTEQ PDF. However, the total statistical uncertainty on this sum of 20 pairs is 0.483. This statistical uncertainty is calculated the following way. The typical RMS of the mass distribution in the pseudo-experiments is 15.3 GeV (for input top mass 178 GeV), so we expect an uncertainty on the mean of that distribution to be $15.3/\sqrt{10000} = 0.153$ GeV (we perform 10000 pseudo-experiments). Therefore the uncertainty on half of the difference of the means is $1/2{*}\sqrt{2}{*}0.153 = 0.108$ GeV and this is the statistical error quoted for each eigenvector. Because, we sum the uncertainties due to all eigenvectors in quadrature, the total statistical uncertainty is $\sqrt{20 \times 0.108^2} = 0.483$. So, the systematic uncertainty due to CTEQ PDF is 0.441 ± 0.483 GeV. To be conservative, we quote the systematic error on the top mass due to the CTEQ6M PDF to be 0.483 GeV.

Finally we estimate the remaining PDF uncertainties. The systematics due to different PDF group (CTEQ5L vs. MRST) was estimated using samples 0 and 1 (first two points in Fig. 9.2) and was measured to be MRST72 - CTEQ5L $= 0.165 \pm 0.153$ GeV. The systematics due to different α_s (MRST72 vs. MRST75) was estimated

---

[2] because, these contributions are independent according the definition

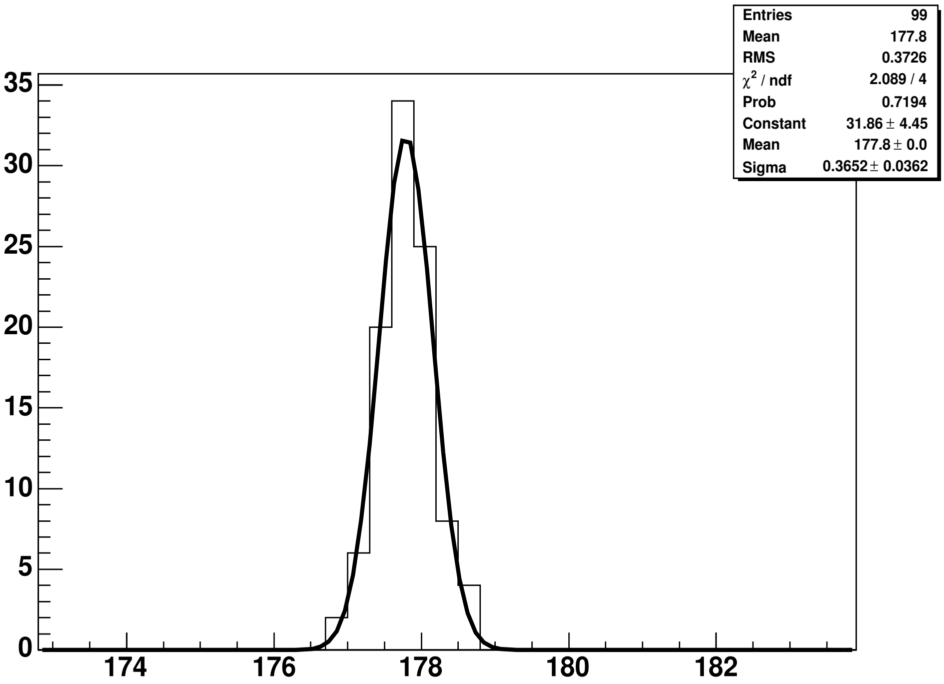


Figure 9.3: Most probable values of mass distributions obtained from pseudo-experiments where Poisson fluctuated signal templates were used.

using samples 1 and 2 and was measured to be MRST75 - MRST72 = 0.054±0.153 GeV.

Adding all three source in quadrature gives the total systematic uncertainty due to PDF of 0.533 ± 0.453 GeV.

9.6 MC limited statistics

Limited number of events in the MC signal and background templates results in an uncertainty in the templates parameterizations (see Tab. 7.2, 7.5) used in the top mass estimate by the likelihood function (see Eq. 6.10). Therefore it is source of systematic error.

We estimate this uncertainty the following way. For each nominal signal template, we fluctuate the number of events in each bin of the template according Poisson distribution and create the new template. Then we obtain the signal parametrization function (see Eq. 7.1) to this fluctuated templates and perform pseudo-experiments by drawing events from the nominal template. The whole procedure is repeated 100 times. From each set of pseudo-experiments, the most probable value (MPV) of reconstructed top mass distribution is extracted. As a systematic uncertainty, we take the width of the distribution of MPVs which is shown in Fig. 9.3.

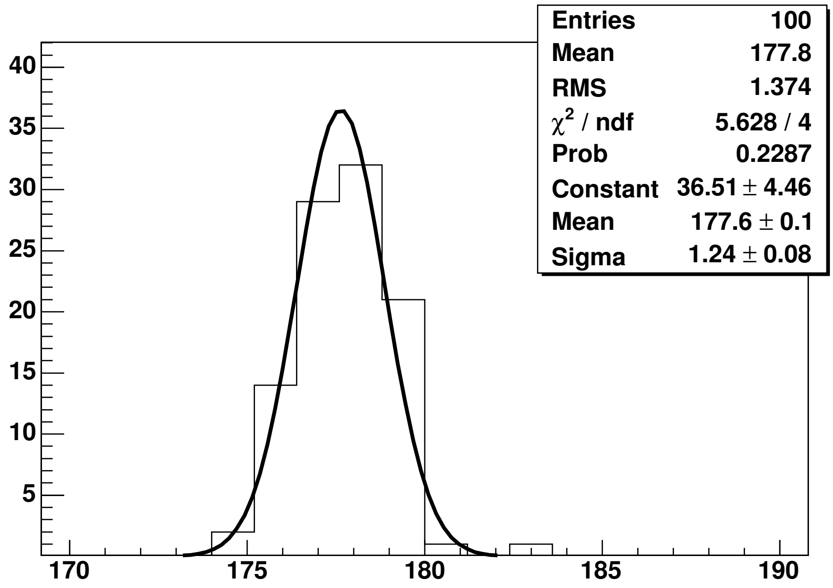


Figure 9.4: Most probable values of mass distributions obtained from pseudo-experiments where Poisson fluctuated background templates were used.

Similarly, we also estimated systematic uncertainty due to limited background MC statistics. We fluctuated according Poisson distribution the number of events in each bin for each background sub-sample and then combine the templates in the final background template. Once again, we repeat it 100 times and perform the pseudo-experiments drawing the events from combined fluctuated background template while the background parametrization in the likelihood fit remains always the same. From each set of pseudo-experiments, the MPV of mass distribution was extracted. The width of the MPVs distribution is 1.2 GeV (see Fig. 9.4) and it is used as the estimate of the systematic uncertainty due to limited MC background statistics.

9.7 Background shape

We evaluated the systematic uncertainty due to different kind of background shapes. We performed this by varying the amount of background components within their uncertainties. The two largest and most different backgrounds are Drell-Yan and FAKE events, see Sec. 7.2. So, we changed the total background composition in the way that we increased the amount of Drell-Yan events by the uncertainty on the number of expected Drell-Yan events (see Tab. 7.4) and at the same time we decreased the amount of FAKE background events by the uncertainty on number of FAKE events. Then, we did it in the opposite way, i.e. we decreased number of Drell-Yan events and increased number of FAKE events. As a systematic error, we take the half of the difference between the reconstructed mass in pseudo-experiments using these two different background cocktails (as a signal we used Herwig 178 GeV sample). The result is 1.2 ± 0.3 GeV.

Another source of systematic uncertainty is an incorrect missing E_T simulation, mis-modeling of lepton P_T or mis-modeling of jet E_T. This could lead to a significant change in the background shape. Once again, we consider only the largest sources of background events (Drell-Yan and FAKE events) which are at the same time also the sources of events which are hard to model in MC.

We compared MC FAKE events to fake-able data events[3]. Because, the FAKE events pass the $t\bar{t}$ selection criteria only when the jet is misidentified as a lepton, we are mainly interested in lepton P_T and jet E_T variables. The mean value of the MC lepton P_T distribution is 4.0% lower than the mean value of lepton P_T distribution seen in fake-able data events. The mean value of the MC jet E_T distribution is 2.1% higher that the mean value of jet E_T distribution from fake-able data events. In order to study the uncertainty related to the difference in lepton P_T, we increased the lepton P_T in MC by 4.0%. A new FAKE template and then a new combined background template were created. We performed pseudo-experiments with the new background template, and we see 0.1 GeV shift in top mass. When we increased jet E_T in MC by 2.1% we saw shift of 0.2 GeV compared to the nominal value.

Drell-Yan MC was compared to dilepton data events which are inside Z mass peak ($m_{\ell\ell} \in (76, 106)$) and which pass the cut on $\not\!\!E_T$. The Drell-Yan events can pass $t\bar{t}$ selection criteria only when there is mismeasurement in jet E_T, so the $\not\!\!E_T$ will appear in the event. Therefore, we are interested here in jet E_T and $\not\!\!E_T$ comparison between data and MC. On average, the MC has 18.5% higher $\not\!\!E_T$ than the data events and 24.7% higher jet E_T than the data events. Therefore, similarly as for FAKE events, we change the $\not\!\!E_T$ and jet E_T in MC by these values and perform the mass reconstruction. The difference in reconstructed top mass with respect to nominal value is taken as a systematic uncertainty. It is 0.3 GeV due to Drell-Yan $\not\!\!E_T$ modeling and 1.0 GeV due to Drell-Yan jet E_T modeling.

The summary of the systematic uncertainties due to background shape is in Tab. 9.3. The total systematic uncertainty due to background shape was calculated by adding the individual contributions in quadrature and it is 1.6 ± 0.5 GeV.

---

[3] for definition of fake-able data event, see Appendix C.1.1

Source	M_{top} (GeV)	system. uncert. (GeV)
Relative composition		1.2 ± 0.3
DY+1σ,FAKES-1σ	179.5	
DY+1σ,FAKES+1σ	177.1	
FAKE lepton P_T	178.2	0.1 ± 0.2
FAKE jet E_T	177.9	0.2 ± 0.2
Drell-Yan $\not{E}_T$	178.5	0.3 ± 0.2
Drell Yan jet E_T	179.1	1.0 ± 0.2
Total		1.6 ± 0.5

Table 9.3: The summary of the systematic uncertainties on the top quark mass due to background shape sources. We use the Herwig sample with the reconstructed mass value 178.1 GeV (see Tab. 9.2) as the nominal mass for all sources (except relative composition). The total uncertainty is obtained by adding the individual contributions in quadrature.

9.8 Summary of the systematic errors

The summary of the systematic uncertainties on the top mass measurement is in Tab. 9.4. The largest source of systematics is due to uncertainty in the jet energy scale. The other large sources of uncertainty are coming from background shape and MC statistics. They will be decreased as larger MC samples will be available and when we will better understand differences between MC and data. We can also see that lots of systematic uncertainties are actually within the errors consistent with zero. However, to be conservative, we consider these systematic uncertainties as real, even when the uncertainties are probably smaller and we include them into estimation of total uncertainty.

The total systematic uncertainty on top quark mass is obtained by assuming all the above mentioned sources of systematic errors are independent. Hence, by adding the individual contributions in quadrature the total systematic error on top quark mass measurement becomes 4.0 GeV.

9.9 Expectations of the top mass systematic errors at Tevatron

With increasing sample size of the data at Tevatron, the statistical error on top mass measurement will become smaller and later at LHC even negligible comparing to the systematics errors (in lepton+jet channel at CDF, the systematic error is already higher

Source	ΔM_{top} (GeV)
Jet Energy Scale	3.2 ± 0.7
b-jet energy scale	0.6 ± 0.6
Generators	0.4 ± 0.6
ISR	0.6 ± 0.3
FSR	0.3 ± 0.3
PDFs	0.5 ± 0.5
Background shape	1.6 ± 0.5
Template statistics	
Signal	0.4 ± 0.0
Background	1.2 ± 0.1
Total	4.0 ± 0.6

Table 9.4: Summary of the systematic uncertainties on the top quark mass measurement. The total uncertainty is obtained by adding the individual contributions in quadrature.

than the statistical error). Therefore, it is good to have projections of systematic errors for the next years, as this will be essentially the expected uncertainty on top mass measurement.

The two uncertainties which scale with the data statistics are statistical uncertainty and uncertainty from jet energy scale. All other uncertainties (due to Generators, ISR, FSR, PDF) depend on improvement in theoretical understanding of underlying physics.

The statistical error decrease with the increasing amount of data by $\sim 1/\sqrt{N}$, where N is number of events.

The same way can be improved also systematic error due to jet energy scale. Jet energies can be scaled to appropriate level by using the decay $W \rightarrow jj$ and requiring that the invariant mass of jets is equal to W mass. Because the W mass is very well known, we will know the jet energy scale very well. With more and more data, we will know the JES with better precision and we can make top mass uncertainty due to this source smaller. Therefore, this uncertainty is also scaled as $\sim 1/\sqrt{N}$. Note, scaling JES using $W \rightarrow jj$ can be performed in-situ only in "lepton+jet" (or all-hadronic) channel because only there the W boson is decaying to pair of jets. In dilepton channel, you can not do this, because the W decays in lepton and undetectable neutrino.

However, most of the systematic uncertainties (ISR, FSR, PDF, generators) do not scale with luminosity, because of their theoretical nature. Even it is assumed that the large data sets can be used to constrain some of the model parameters, it is expected that we can ultimately reduce the total systematic uncertainty due to all systematic

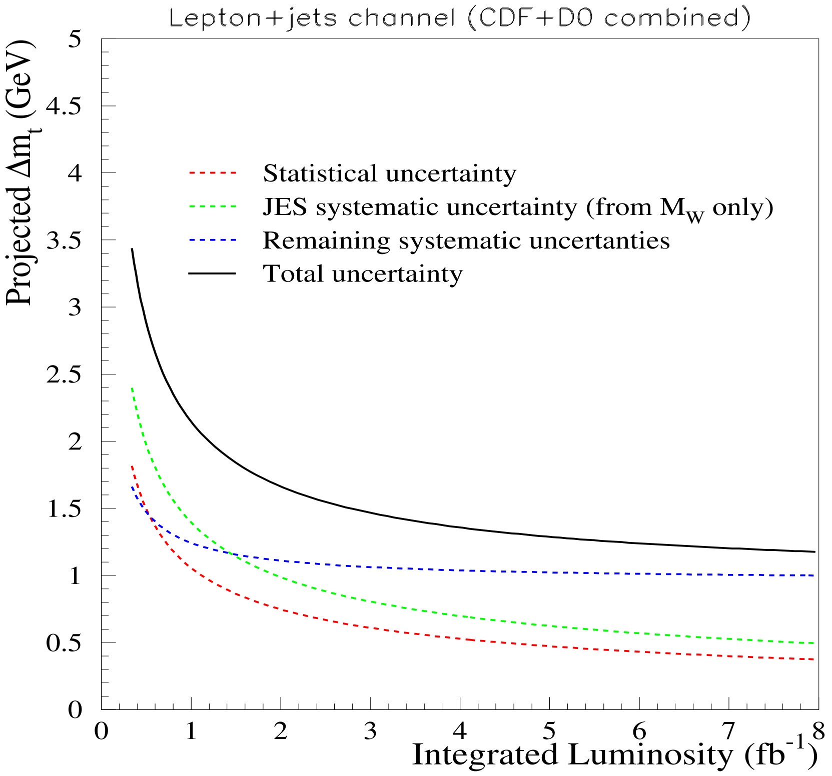


Figure 9.5: Projections of statistic and systematic uncertainties on top mass measurement as a function of integrated luminosity at Fermilab [24]. The combination of CDF and D0 results and only lepton+jet channel is assumed.

sources except JES to about 1 GeV at the end of Run 2 (after 8 fb^{-1} of data collected).

In Fig. 9.5 there is the summary of the expected systematic and statistical uncertainties on top mass measurement in lepton+jet channel combined from both CDF and D0. We can see that the statistical error will decrease below 0.5 GeV and similarly will do the systematics due to JES. The dominant contribution will come from other sources of systematic errors like PDFs, ISR, FSR. The total expected uncertainty on top mass at the end of Run II is expected to be below 1.5 GeV.

At the following experiments at the LHC, the picture from Tevatron will still remain approximately valid. Of course, with so many available $t\bar{t}$ events, it will be possible to use different kind of algorithms and also to select events which are as little as possible affected by different systematic effects. Therefore the total uncertainty on top quark mass is expected to be a little bit smaller than at Tevatron, around 1 GeV, see Sec. 5.3.1.

As already mentioned in Sec. 5.3.2, the top quark mass can be measured much more precisely in e^+e^- collisions using $t\bar{t}$ production near the threshold. This is considered to be the most accurate method and it appears that an uncertainty of $\delta m_t \approx 0.15$ GeV can be achieved for the top quark $\overline{MS}$ mass.

Chapter 10

Data Results

After we have tested our method on simulated events (see Chap. 8), we are comfortable that our method returns appropriate top quark mass and also error estimate. We can finally apply our reconstruction procedure on CDF data sample and measure the top quark mass and its statistical uncertainty.

First, we will describe in detail the reconstructed "raw top quark mass" in data events. Then, we will perform the comparison between the CDF data and MC by likelihood fit, in order to obtain our final estimate of the top quark mass.

10.1 Data events

In the CDF Run II data sample corresponding to luminosity 340 pb^{-1}, we found 33 events which pass the selection criteria for $t\bar{t}$ event described previously in Chap. 4. The kinematic characteristics of these events are summarized in Tab. 10.1, while the comparison between the CDF data and MC events have been already shown in Sec. 6.2.3. In the table are listed run and event number, dilepton category, transverse momenta of both leptons ($p_T^{\ell_1}$,$p_T^{\ell_2}$), number of jets (N_{jets}), the transverse energies of two highest E_T jets ($E_T^{jet_1}$, $E_T^{jet_2}$) and missing transverse energy ($\not{E}_T$) for a given event.

We can see that there are 5 $e-e$, 12 $\mu-\mu$ and 16 $e-\mu$ events.

We applied the mass reconstruction procedure described in Chap. 6 on these 33 $t\bar{t}$ data candidates. We were able to reconstruct top quark mass for the 30 out of 33 events[1]. Thus, the probability to reconstruct a top mass for a given data event is 30/33 = 90.9%. For MC simulated events, we get about 70 % for expected number of signal and background events, see Tab. 7.6. We discuss more the reconstruction probability difference between MC and data in Appendix D.

---

[1] Not always it's possible to find a kinematic solution. Moreover, recall that the event must pass the cut 10 % on event reconstruction probability in order to be considered as reconstructed, see Sec. 6.2.1.

run/event number	dilepton-category	$p_T^{\ell_1}$	$p_T^{\ell_2}$	N_{jets}	$E_T^{jet_1}$	$E_T^{jet_2}$	$\not\!\!E_T$
150418 / 960369	NICEM – PHX	89.5	91.5	3	33.7	28.9	68.7
153374 / 2276742	CEM – CEM	72.0	56.9	3	89.3	58.4	58.1
177491 / 3807306	NICEM – CEM	25.9	21.0	3	70.8	45.9	79.2
186598 / 1618142	CEM – CEM	69.4	35.7	3	73.6	72.3	67.0
186598 / 4194951	CEM – CEM	54.5	36.1	2	94.0	61.9	42.6
153325 / 599511	CMX – CMX	37.1	25.2	3	54.0	37.4	46.5
153447 / 2643751	CMUP – CMP	87.4	26.3	3	110.2	23.9	35.5
154654 / 7344016	CMX – CMUP	59.4	52.9	3	39.5	22.9	56.7
160988 / 385505	NICMX – CMX	26.0	21.0	2	90.7	60.2	32.2
162820 / 7050764	CMUP – CMP	33.5	26.2	2	69.3	55.6	95.8
163012 / 1438203	RHOCMX – CMUP	119.1	98.5	2	107.6	77.1	80.6
165198 / 1827962	CMX – CMUP	34.5	23.8	2	66.4	34.8	101.4
166063 / 2833132	CMU – CMUP	41.4	38.0	2	52.9	52.8	65.0
178540 / 2208375	CMUP – CMIO	64.4	34.8	2	99.9	45.7	53.6
178738 / 10340757	CMUP - CMIO	50.0	21.4	2	79.2	39.1	88.4
185037 / 2287335	CMUP – CMU	85.7	56.1	2	69.1	63.3	26.9
185377 / 103906	CMP – CMUP	76.0	20.9	2	72.9	29.3	80.6
143257 / 760520	CMP – CEM	27.3	21.0	5	47.0	39.9	92.2
150431 / 368759	CEM – NICMUP	58.3	36.2	4	58.1	35.6	44.1
150435 / 2896171	NICMU – CEM	45.4	42.9	2	91.8	47.4	64.7
151978 / 507773	NICEM - CMUP	35.9	34.7	2	50.2	48.9	91.4
155114 / 478702	CMX – CEM	34.5	28.5	3	77.0	39.1	90.0
156484 / 3099305	CMX – CEM	35.2	34.3	3	70.3	37.3	75.9
161633 / 963604	CEM – CMIO	40.2	26.5	3	65.6	39.4	60.5
165364 / 592961	CEM - CMU	47.4	21.6	3	82.4	50.5	127.9
167053 / 12011678	CEM – RHOCMX	75.8	50.4	4	90.2	29.3	28.5
167629 / 180103	CEM – CMUP	96.6	78.6	2	80.5	38.5	28.4
167631 / 2058969	CEM – CMX	34.0	31.3	2	41.4	39.7	61.3
168599 / 2964061	CMUP – CEM	40.8	23.7	3	38.8	36.5	96.8
178738 / 1660363	CMUP – CEM	100.0	60.6	3	45.8	39.6	90.3
183963 / 1259645	CMIO – CEM	84.7	27.0	2	146.2	31.9	63.1
184779 / 892809	CMIO – PHX	51.1	46.9	2	49.2	22.9	48.9
185594 / 3002817	CEM – CMIO	80.5	26.1	2	143.1	25.2	199.2

Table 10.1: Kinematic characteristics of $t\bar{t}$ data candidates. Events are grouped according dilepton category into $e-e$, $\mu-\mu$ and $e-\mu$ events in mentioned order. Leptons p_T are listed in the order of mentioned lepton types for a given event.

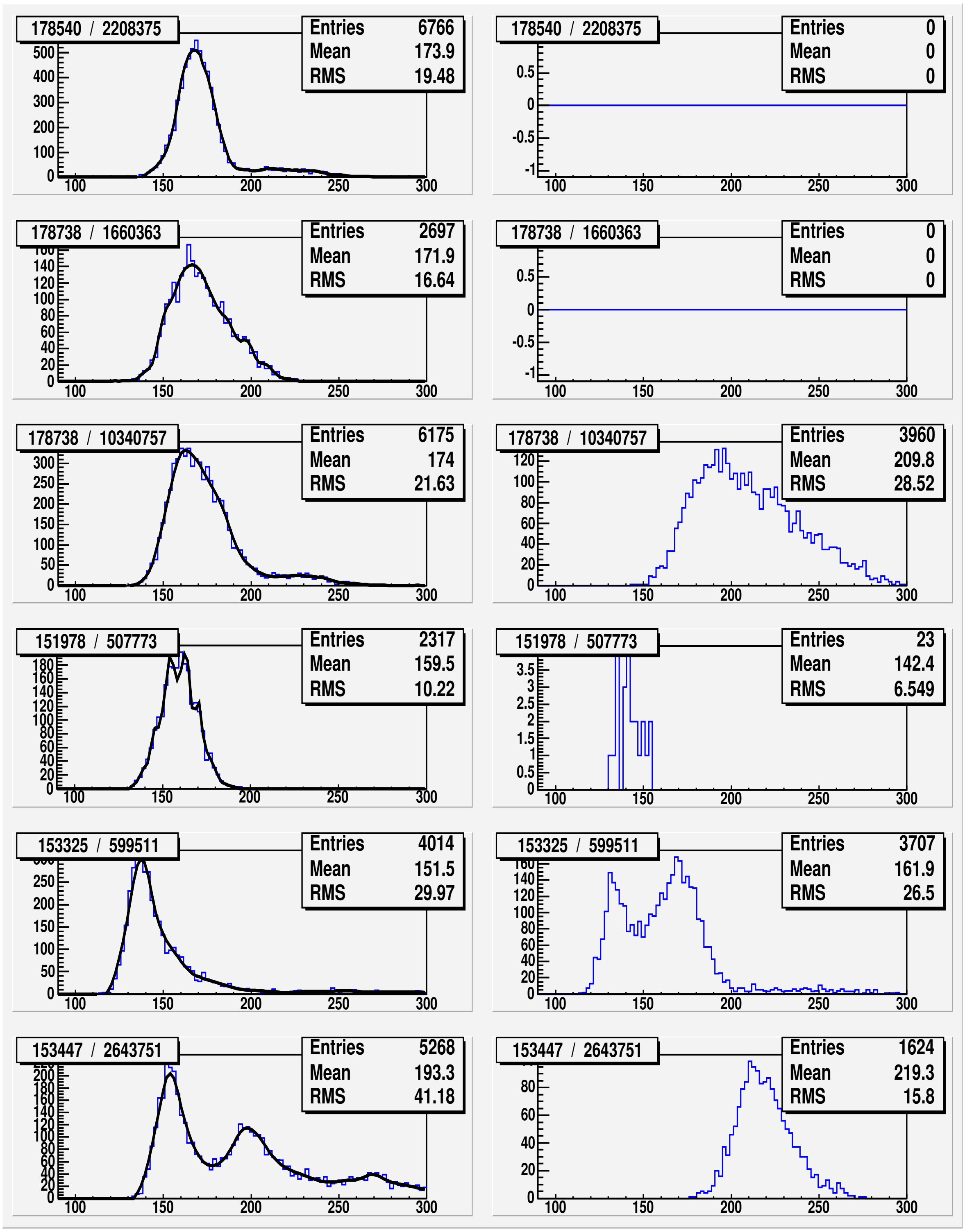


Figure 10.1: Top mass distributions of smeared events for each data $t\bar{t}$ candidate event. On left side is combination of lepton+jet which was selected, on right side is disfavored combination of lepton+jet.

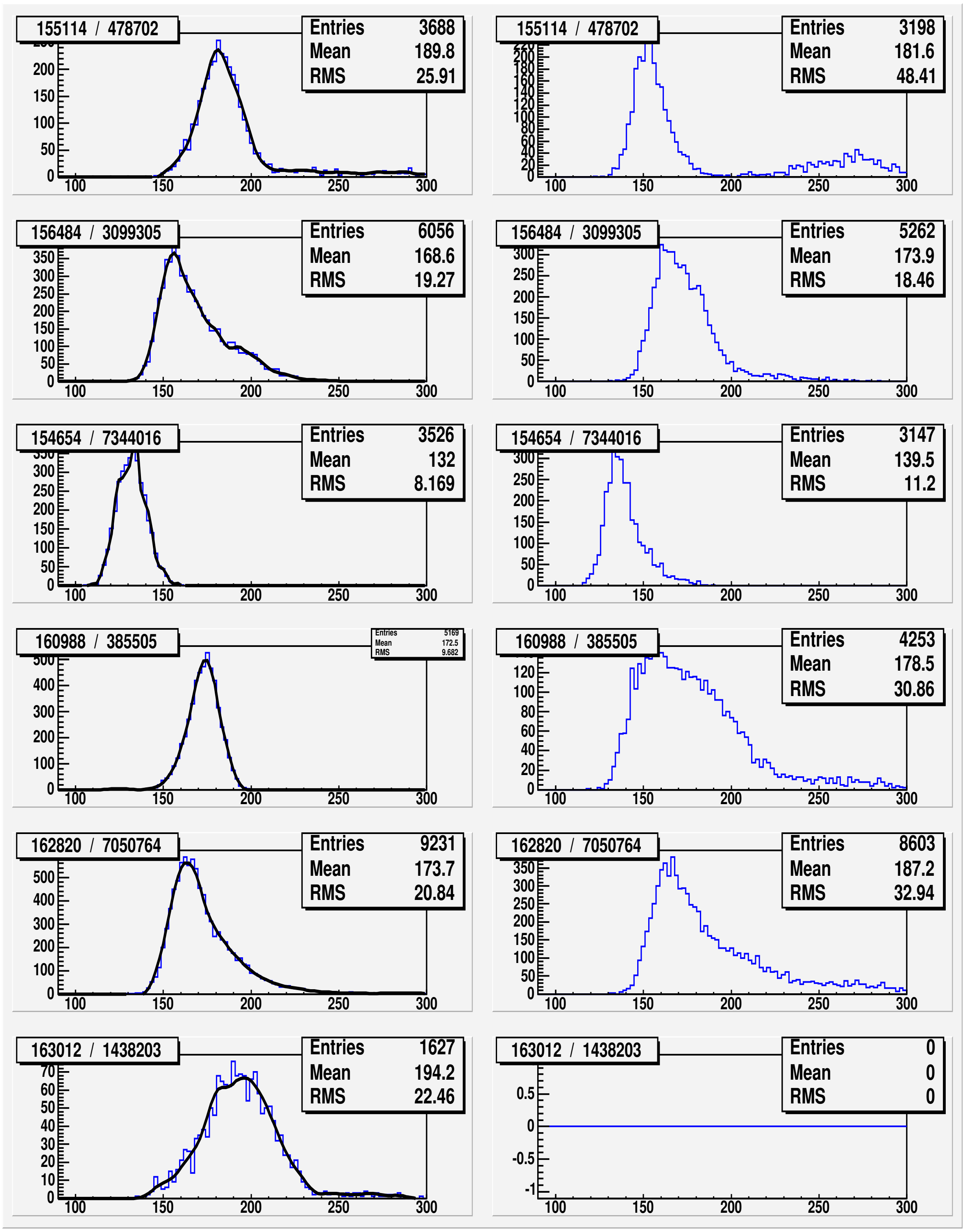


Figure 10.2: Top mass distributions of smeared events for each data $t\bar{t}$ candidate event. On left side is combination of lepton+jet which was selected, on right side is disfavored combination of lepton+jet.

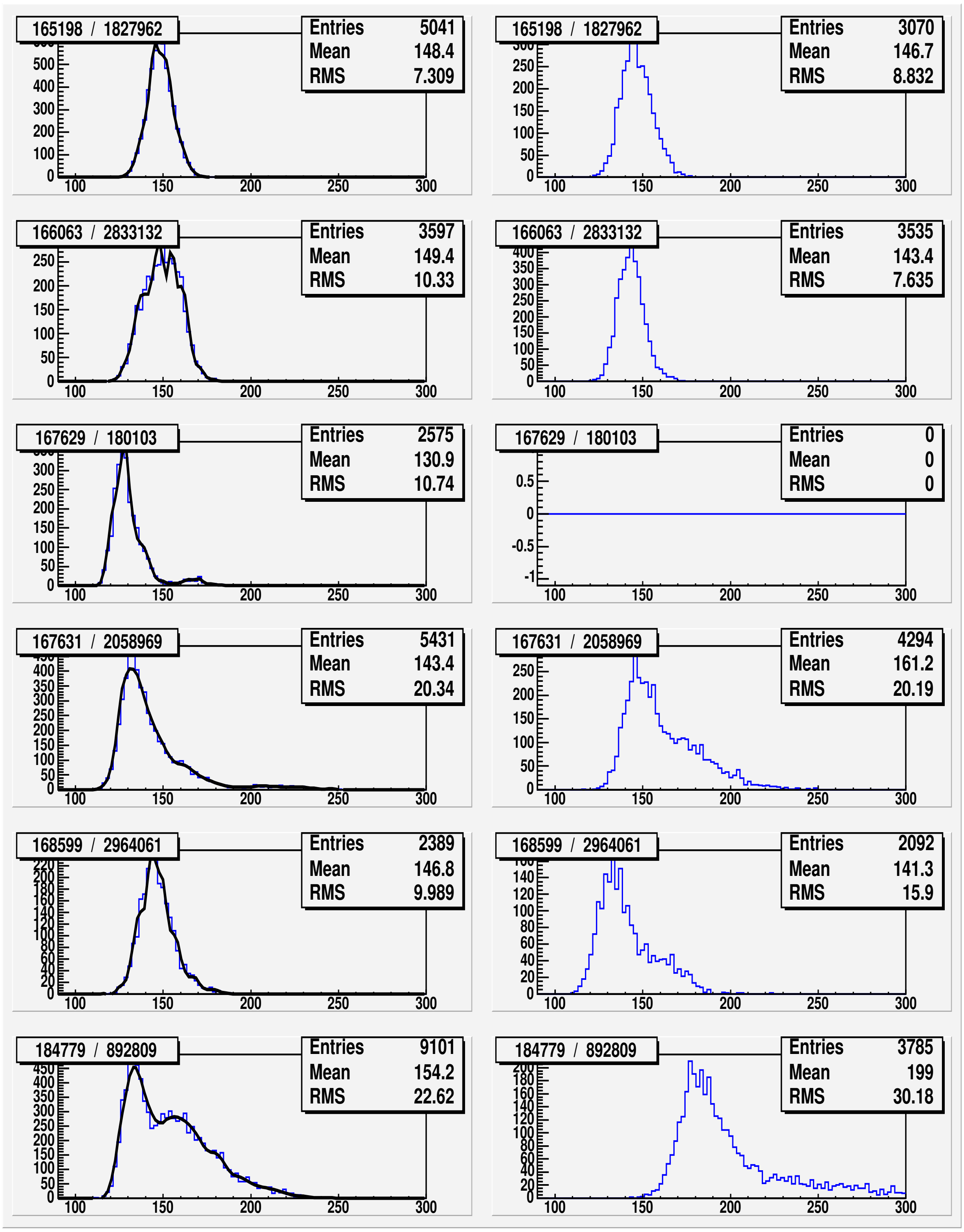


Figure 10.3: Top mass distributions of smeared events for each data $t\bar{t}$ candidate event. On left side is combination of lepton+jet which was selected (together with spline fit), on right side is disfavored combination of lepton+jet.

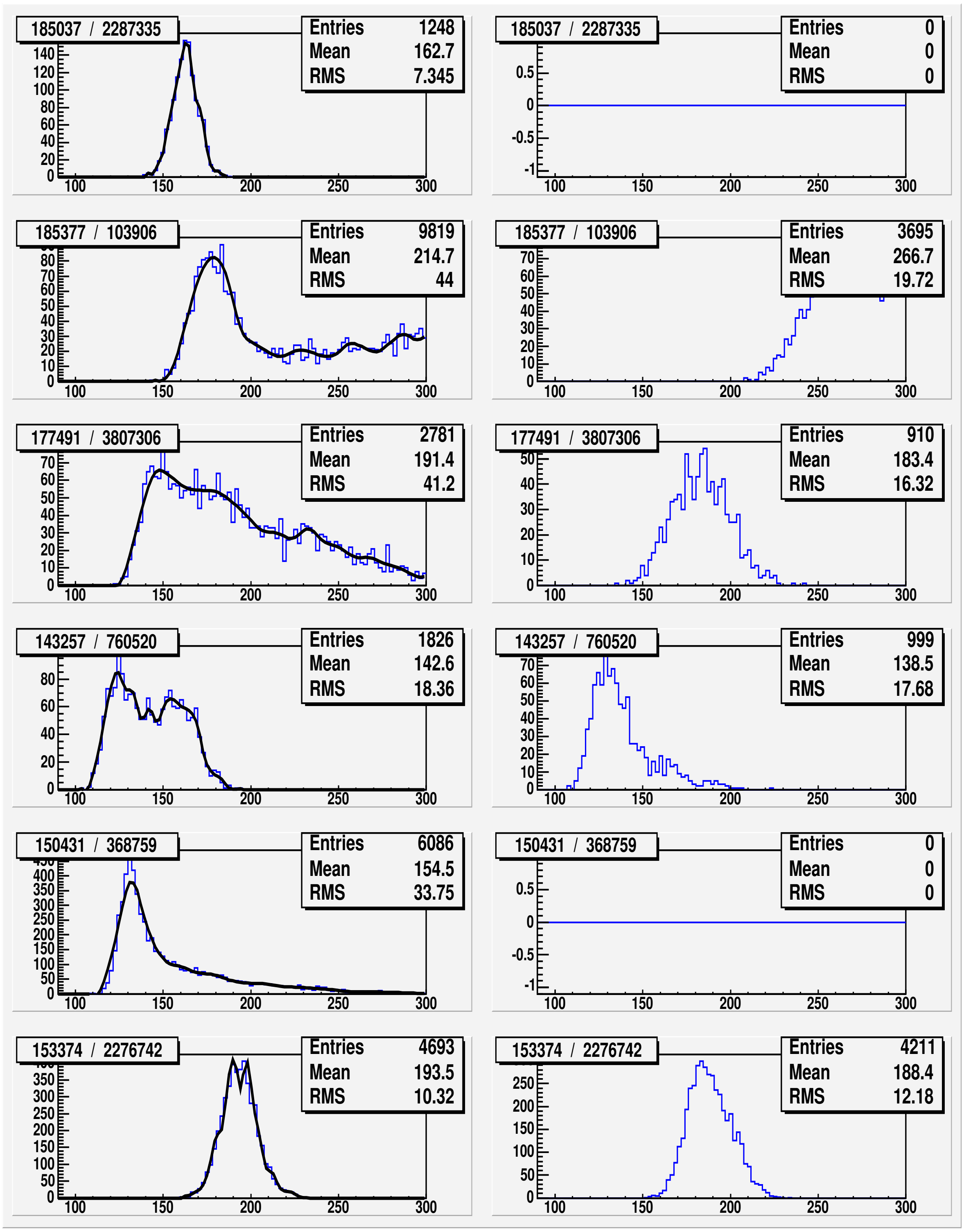


Figure 10.4: Top mass distributions of smeared events for each data $t\bar{t}$ candidate event. On left side is combination of lepton+jet which was selected (together with spline fit), on right side is disfavored combination of lepton+jet.

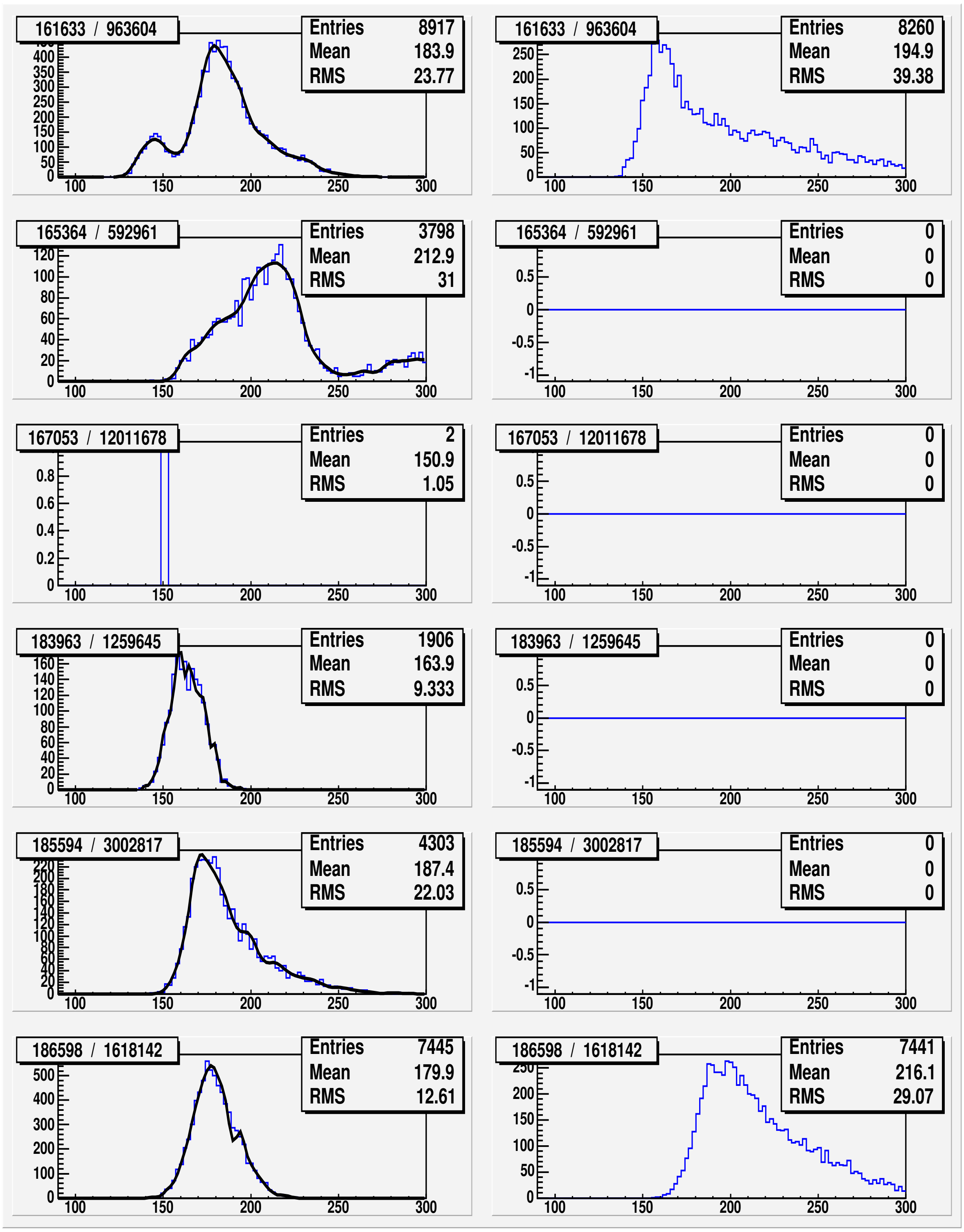


Figure 10.5: Top mass distributions of smeared events for each data $t\bar{t}$ candidate event. On left side is combination of lepton+jet which was selected (together with spline fit), on right side is disfavored combination of lepton+jet.

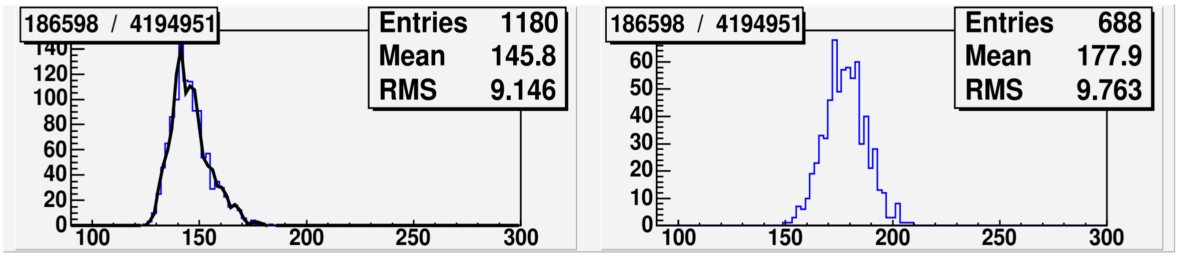


Figure 10.6: Top mass distributions of smeared events for each data $t\bar{t}$ candidate event. On left side is combination of lepton+jet which was selected (together with spline fit), on right side is disfavored combination of lepton+jet.

As have been explained in Sec. 6.2.1, for each event we perform smearing of kinematic variables and reconstruct the mass for these smeared events (trials). Thus, we obtain mass distribution for each physics event. Such top mass distributions for each $t\bar{t}$ data candidate are presented in Fig. 10.1 – Fig. 10.6. There are two histograms for each event. They are corresponding to two possibilities of pairing the lepton and b-jet. As described before, from these two possibilities, the one with higher number of trials with reconstructed top mass is selected, i.e. the one with higher number of entries in histogram. The selected combinations are plotted on the left side, the disfavored ones are on the right side for each event. The most probable value from the histogram of selected combination is "raw top mass" for a given event. Altogether, there are 31 events plotted. These are 30 events where we reconstructed top mass (reconstruction probability is higher than 10 %). Moreover, there is plotted one event where only 2 out of 10000 trials were reconstructed for selected combination and thus this event does not pass the cut on reconstruction probability. Another two events which don't pass mass reconstruction are not plotted, because they don't have even one reconstructed trial, therefore the mass histograms for these events are empty. The list of events together with value of reconstructed "raw mass" for each event is shown in Tab. 10.2 (events which fail the reconstruction procedure are labeled as 'FAILED'). Two out of three events which failed the top mass reconstruction procedure are events with the non-isolated lepton (150418/960369 have electron - NICEM, 150435/2896171 have muon - NICMU). Thus, they have larger probability to be a background event (e.g. FAKE event) and consequently larger probability not to be reconstructed. The last event which failed the mass reconstruction (167053/12011678) is an event with a 4 jets in final state. The 2nd jet has relatively low $E_T = 29.3$ and is not much higher than 3rd and 4th jet E_T of 22.9 and 16.7 GeV, respectively. Therefore, it is quite possible, that we used wrong jet in mass reconstruction procedure.

The distribution of raw mass for data events is presented in Fig. 10.7, where

event category	run/event	raw top mass
ee	153374 / 2276742	189.75
ee	177491 / 3807306	147.75
ee	186598 / 1618142	177.15
ee	186598 / 4194951	141.45
$\mu\mu$	153325 / 599511	137.25
$\mu\mu$	153447 / 2643751	154.05
$\mu\mu$	154654 / 7344016	133.05
$\mu\mu$	160988 / 385505	175.05
$\mu\mu$	162820 / 7050764	162.45
$\mu\mu$	163012 / 1438203	196.05
$\mu\mu$	165198 / 1827962	145.65
$\mu\mu$	166063 / 2833132	147.75
$\mu\mu$	178540 / 2208375	168.75
$\mu\mu$	178738 / 10340757	162.45
$\mu\mu$	185037 / 2287335	162.45
$\mu\mu$	185377 / 103906	179.25
$e\mu$	143257 / 760520	124.65
$e\mu$	150431 / 368759	133.05
$e\mu$	151978 / 507773	162.45
$e\mu$	155114 / 478702	181.35
$e\mu$	156484 / 3099305	156.15
$e\mu$	161633 / 963604	179.25
$e\mu$	165364 / 592961	212.85
$e\mu$	167629 / 180103	128.85
$e\mu$	167631 / 2058969	130.95
$e\mu$	168599 / 2964061	143.55
$e\mu$	178738 / 1660363	166.65
$e\mu$	183963 / 1259645	160.35
$e\mu$	184779 / 892809	133.05
$e\mu$	185594 / 3002817	172.95
ee	150418 / 960369	FAILED
$e\mu$	150435 / 2896171	FAILED
$e\mu$	167053 / 12011678	FAILED

Table 10.2: Reconstructed top mass ("raw top mass") for data candidates. We present here also the dilepton category and the run and event number for each event.

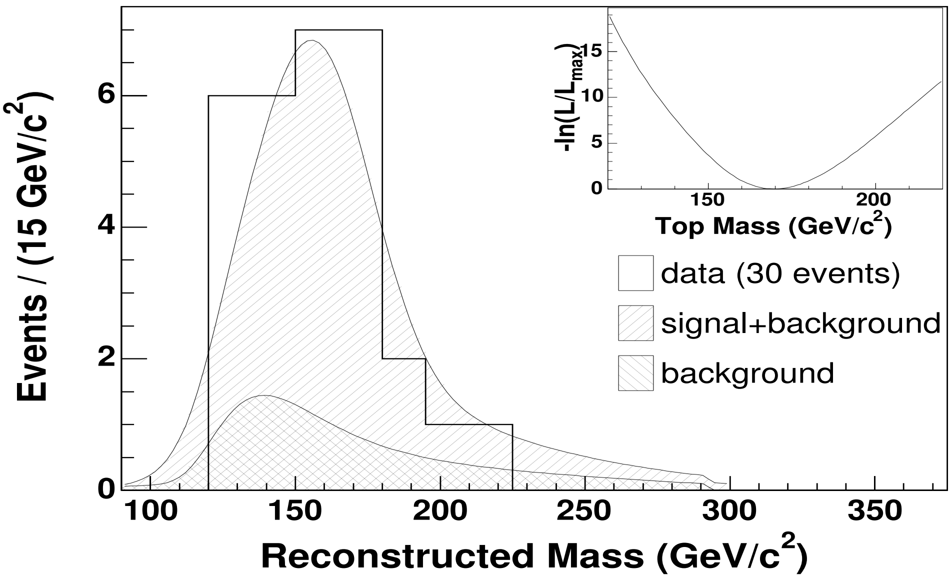


Figure 10.7: Reconstructed top mass ("raw top mass") for the data $t\bar{t}$ candidates (histogram) together with normalized background and signal + background p.d.f. curves. For signal we used here sample with 170 GeV input top mass. There is also shown the likelihood function (inset).

it is also visually compared with the combined signal + background MC template (for signal, we used $t\bar{t}$ 170 GeV MC template). The Kolmogorov-Smirnov test of compatibility between distributions gives probability of about 61 %, which means very good agreement.

The average of raw top mass for data events is 158.9 ± 3.9 GeV. We will see later in Sec. 10.2 that this is roughly 10 GeV lower value comparing to top mass estimate from likelihood fit (see Eq.(10.1) for proper comparison). As it was already pointed out in description of the method (see Sec. 6.1), this value would be our estimate of top mass, in case we would not compare the data to MC (by likelihood fit)[2] and we were already expecting the bias to lower values. The average of raw top mass for a different dilepton categories is 164.0 ± 10.0 GeV for $e - e$ events, 160.4 ± 5.0 GeV for $\mu - \mu$ events and 156.2 ± 6.6 GeV for $e - \mu$ events. All the values are statistically consistent between each other.

Below, we discuss the properties (which seems to be problematic) of all selected events and also top mass distributions for each event.

150418 / 960369 *ee*

This event has non-isolated electron (NICEM), however it only slightly doesn't pass

---

[2] of course, after properly taking into account background

the cut on isolation ($Isol = 0.21 > Isol_{cut} = 0.1$). Moreover, there are 3 jets in the event.

The top mass is not reconstructed for this event. There are no trials reconstructed for both lepton+jet configurations.

153374 / 2276742 *ee*

This is 3 jet event.

The mass distributions of favored and disfavored lepton+jet combination are very similar in number of reconstructed trials (47 % vs. 42 %) as well as in shape (both are sharp peaks).

177491 / 3807306 *ee*

There is non-isolated lepton lepton (NICEM, Isol = 0.2) in the event and also the three jets.

The reconstruction probability (RP) is much better for favored solution ($RP_{fav} =$ 28%) than for disfavored ($RP_{dis} = 9\%$). However, the mass distribution is very wide for favored combination.

186598 / 1618142 *ee*

This is 3 jets event.

Favored and disfavored combinations have quite close probabilities of reconstruction ($RP_{fav} = RP_{dis} = 74\%$). While favord combination has nice sharp peak, disfavored combination has broad peak with of mass distribution with a long tail.

186598 / 4194951 *ee*

This event seems according selection criteria an ideal $t\bar{t}$ event with exactly two jets.

Even for both lepton+jet combinations there are quite small reconstruction probabilities ($RP_{fav} = 12\%$, $RP_{dis} = 7\%$), both peaks are nice sharp peaks.

153325 / 599511 $\mu\mu$

This is 3 jets event.

Reconstruction probabilities for each combination of the lepton and b-jet pairing are very close ($RP_{fav} = 40\%$, $RP_{dis} = 37\%$). However, the disfavored combination has wide distribution (with a double peak) while favored combination has nice sharp peak.

153447 / 2643751 $\mu\mu$

In this event, the invariant mass of two muons is intriguingly close to mass of Z boson ($M_{\mu\mu} = 90.63$ GeV). Moreover, there are 3 jets in the event.

The reconstruction probability is much higher for favored combination ($RP_{fav} =$ 53%, $RP_{dis} = 16\%$). The favored combination's mass distribution has two peaks, which may be related to the loss of momenta of the b-jets by radiation of gluons. However, the disfavored combination have quite a sharp peak in mass distribution.

154654 / 7344016 $\mu\mu$

This is once again 3 jets event.

The both lepton+jet solutions are quite close in reconstruction probability ($RP_{fav} = 35\%$,$RP_{dis} = 31\%$) and also both have nice sharp peaks.

160988 / 385505 $\mu\mu$

This event have non-isolated muon with very high non-isolation ($Isol = 1.1$).

The reconstruction probabilities for favored and disfavored combination are quite close ($RP_{fav} = 52\%$,$RP_{dis} = 43\%$), however the disfavored mass distribution is much wider than favored distribution.

162820 / 7050764 $\mu\mu$

This is once again another "ideal" $t\bar{t}$ event according kinematic variables with exactly 2 jets.

There is very high reconstruction probability for both favored and disfavored combinations ($RP_{fav} = 92\%$,$RP_{dis} = 86\%$). Disfavored combination have long tail in mass distribution.

163012 / 1438203 $\mu\mu$

This is 2 jet event.

Only favored combination has reconstructed events with quite low reconstruction probability ($RP_{fav} = 16\%$) and wide mass distribution.

165198 / 1827962 $\mu\mu$

This seems another "ideal" $t\bar{t}$ event with 2 jets.

The top mass distributions of favored and disfavored combinations are very similar in shapes (sharp peaks), however the reconstruction probability is a quite higher for favored combination than for disfavored ($RP_{fav} = 50\%$,$RP_{dis} = 31\%$).

166063 / 2833132 $\mu\mu$

Another "ideal" $t\bar{t}$ event with 2 jets.

The top mass distributions of favored and disfavored combinations are very similar in both reconstruction probability ($RP_{fav} = 36\%$,$RP_{dis} = 35\%$) and peak of mass distribution (sharp peak).

178540 / 2208375 $\mu\mu$

The invariant mass of two muons is quite close to Z boson mass peak $M_{\mu\mu} = 94GeV$.

Only favored combination has reconstructed events where the reconstruction probability is quite high $RP_{fav} = 68\%$ and it has nice sharp peak.

178738 / 10340757 $\mu\mu$

This seems another "ideal" $t\bar{t}$ candidate, just it has trackless muon (CMIO).

Quite high probability of reconstruction for both favored and disfavored option ($RP_{fav} = 62\%$,$RP_{dis} = 40\%$). However, disfavored combination has wider mass distribution.

185037 / 2287335 $\mu\mu$

Another “ideal” $t\bar{t}$ candidate according kinematic parameters.

Here, only favored combination has reconstructed events with quite small probability of reconstruction $RP_{fav} = 13\%$, but with nice sharp peak.

185377 / 103906 $\mu\mu$

Another “ideal” $t\bar{t}$ candidate according kinematic parameters with two jets.

There is very high reconstruction probability for the favored combination ($RP_{fav} = 98\%$,$RP_{dis} = 37\%$). The favored distribution has quite normal peak, however with an increasing tail. The disfavored distribution has very high reconstructed mass.

143257 / 760520 $e\mu$

This is the only event which has 5 jets.

The mass distributions of both favored and disfavored combination of the lepton and b-jet pairings are quite wide. The reconstruction probabilities are also quite similar and moreover quite low ($RP_{fav} = 18\%$, $RP_{dis} = 10\%$). Selected solution has a double-peak.

150431 / 368759 $e\mu$

This is 4 jets event moreover with non-isolated muon (NICMP with quite high Isol = 0.6).

Only favored combination has reconstructed events (“trials”) with quiet high reconstruction probability ($RP_{fav} = 61\%$) and also sharp peak (with long tail).

150435 / 2896171 $e\mu$

There is quite big non-isolated muon (NICMU,$Isol = 1.05$), so its possible it is jet which is faking a muon.

The mass is not reconstructed at all, i.e. for none of trials.

151978 / 507773 $e\mu$

One lepton which is only slightly non-isolated (NICEM, $Isol = 0.16$).

The reconstruction probability is very small for disfavored combination ($RP_{fav} = 23\%$, $RP_{dis} = 0.23\%$). The peak of mass distribution is quite sharp.

155114 / 478702 $e\mu$

This is 3 jets event.

The probabilities for favored and disfavored combination are quite close ($RP_{fav} = 37\%$, $RP_{dis} = 32\%$). The peaks of mass distributions look quite sharp, only in disfavored combination there is some bump at high top masses.

156484 / 3099305 $e\mu$

This is 3 jets event.

The probabilities for favored and disfavored combination are quite close and high ($RP_{fav} = 61\%$, $RP_{dis} = 53\%$) and both mass distributions are quite usual.

161633 / 963604 $e\mu$

This is 3 jets event with one CMIO muon.

Very high reconstruction probability for both favored and disfavored case ($RP_{fav} = 89\%$, $RP_{dis} = 85\%$), but both mass distributions are quite wide.

165364 / 592961 $e\mu$

This is 3 jets event.

Only favored combination has reconstructed events ($RP_{fav} = 38\%$) and the mass distribution is quite wide.

167053 / 12011678 $e\mu$

This is 4 jet event. The 2nd jet has relatively low $E_T = 29.3$ and is not much higher than 3rd and 4th jet E_T of 22.9 and 16.7 GeV, respectively. Therefore, it is quite possible, that we used wrong jet in mass reconstruction procedure.

The top mass is not reconstructed. The favored combination has only 2 out of 10000 trials reconstructed.

167629 / 180103 $e\mu$

This is another “ideal” $t\bar{t}$ event with 2 jets.

Only favored combination has reconstructed events. with $RP_{fav} = 25\%$. The mass distribution has nice sharp peak.

167631 / 2058969 $e\mu$

This is 2 jet event with H_T almost at the limit ($H_T = 203$).

The disfavored distribution is wider than the favored with sharp peak while the reconstruction probabilities are similar ($RP_{fav} = 54\%$, $RP_{dis} = 43\%$).

168599 / 2964061 $e\mu$

This is 3 jet event.

Favored and disfavored combination has low and very close probabilities of reconstruction ($RP_{fav} = 24\%$, $RP_{dis} = 21\%$). The favored distribution is sharper than the disfavored one.

178738 / 1660363 $e\mu$

This is 3 jet event.

Only favored combination has reconstructed events, but with small reconstruction probability ($RP_{fav} = 27\%$). The peak is quite usual.

183963 / 1259645 $e\mu$

This is 2 jet event with a CMIO muon.

Only favored combination has reconstructed events with quite small reconstruction probability ($RP_{fav} = 19\%$), but wiath a nice sharp peak.

184779 / 892809 $e\mu$

This is 2 jet event with a CMIO muon and PHX electron.

The favored combination has much higher reconstruction probability than disfavored ($RP_{fav} = 91\%$, $RP_{dis} = 38\%$). Both mass distributions are quite wide.

185594 / 3002817 $e\mu$

This is 2 jet event with CMIO muon.

Only the favored combination has reconstructed events with modest reconstruction probability $RP_{fav} = 43\%$). The mass distribution has quite usually wide peak, however with long tail.

According the above descriptions, we have usually nice sharp peaks for the "ideal events" (usually it requires just exactly two jets in the event). Therefore, finding the way of improving the reconstruction for more jets (≥ 3 jets) events (e.g. by combining the jets) should improve the reconstruction of "raw top mass" and consequently also the resolution on top mass.

10.2 Data top mass estimate

The final estimate of the top quark mass and its statistical uncertainty using CDF data sample is performed by a likelihood fit (see Sec. 6.2.3).

Just as a test, we first assume no background, i.e. we fix the $n_b = 0$ in the likelihood formula (Eq. 6.10). The fitted value of the top mass is then:

$$M_{top} = 168.4^{+6.3}_{-5.9}(stat.) \quad \text{GeV}. \tag{10.1}$$

Then, we take into account also the background. We set the number of expected background events to $n_b^{exp} = 6.4 \pm 1.2$ according Tab. 7.6. The uncertainty on number of expected background events ($\sigma_{n_b} = 1.2$) constrain Gaussianlly the amount of background in the fit according Eq. (6.10). The maximization of the likelihood converge to top mass value [3]:

$$M_{top} = 169.5^{+7.5}_{-7.0}(stat.) \quad \text{GeV}, \tag{10.2}$$

where the amount of background as a result of the fit is $n_b = 6.1 \pm 1.7$ events and number of signal events as a result of the fit is $n_s = 24.5 \pm 5.6$ events. We find the probability 32 % that this likelihood-estimate of number of background events n_b is consistent with *a priori* estimate n_b^{exp} and the probability about 7 % that the likelihood-estimate for the total number of signal and background events ($n_s + n_b$) is consistent with the observed number of events N.

The negative log likelihood ($-\ln \mathcal{L}$) dependence on the top quark mass for the 30 data events is presented in Fig. 10.7. We can see that the log likelihood ($-\ln \mathcal{L}$)

---

[3] in fact, we perform minimization of negative logarithm of likelihood function

dependence is already quite nice parabola which is expected in the limit of large number of events.

Note, when we incorporate the background into likelihood fit, the mass estimate from the fit is shifted upwards. It is due to fact that the background events has lower reconstructed top mass (the combined background template in Fig. 7.5 has peak around 140 GeV) than the $t\bar{t}$ events (the peak of MC $t\bar{t}$ template with generated $m_{top} = 170$ GeV is around 160 GeV, see Fig. 7.3). When we assume zero background in the fit, then according Eq. (6.10), we compare the data events only to signal MC templates and we obtain certain top mass estimate from likelihood fit. But, when we perform fit assuming non-zero background, then we are comparing the same data events to the combination of the signal and background template. Because the background template is shifted to lower values, our top mass estimate will be shifted to higher values, comparing to the case with no background template included.

The result we obtained from the log-likelihood fit must be corrected for pull width from from the pseudo-experiments as described on page 151 (we must multiply the statistical uncertainty estimate by 1.033). Thus, the final top mass estimate is:

$$\boxed{M_{top} = 169.5^{+7.7}_{-7.2}(stat.) \quad \text{GeV.}} \tag{10.3}$$

As a cross-check, we performed also unconstrained fit, where we release the Gaussian constraint on number of background events in the fit. In such case, the number of background events converge to zero and the top mass estimate corrected for pull width is then $M_{top} = 168.4 \pm 7.8(stat.)$ GeV. We expect that the fit converges to zero number of background events in about 31 % of cases, according pseudo-experiment studies. We see that there is an improvement in mass uncertainty when applying background constrain to the likelihood fit and we expect on average an improvement in statistical uncertainty of 1.5 GeV (according pseudo-experiments studies).

The expected statistical uncertainty based on pseudo-experiments that contain the same number of events as we see in data (30 events on average where 6.4 events out of them are the background events) is shown in Fig. 10.8. The probability for achieving the same or better than observed uncertainty in data is 23 %.

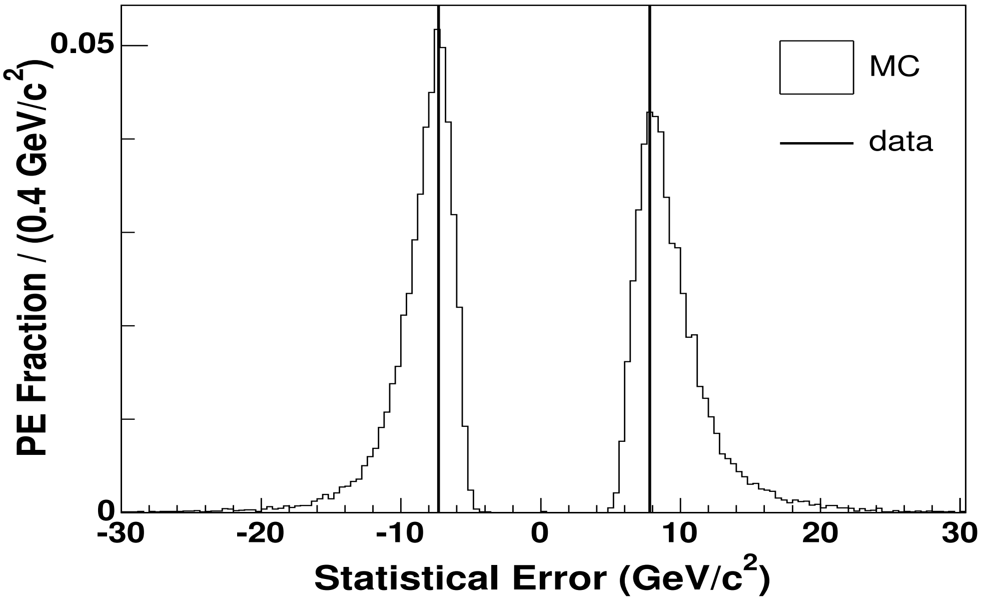


Figure 10.8: Comparison of measured statistical uncertainty (positive and negative) in the data sample (vertical line) with pseudoexperiments generated using the 170 GeV signal template.

Chapter 11

Conclusion

In the work presented in this thesis we measure the top quark mass in dilepton channel using 340 pb^{-1} of data collected by CDF experiment.

We found 33 $t\bar{t}$ candidates passing event selection while the expected number of background was 10.5 ± 1.9 events.

Using our mass reconstruction method, we were able to reconstruct the top quark mass in 30 out of 33 candidates. By comparing reconstructed top quark mass between data and simulated events, we arrived at top quark mass estimate

$$M_{top} = 169.5^{+7.7}_{-7.2}(stat.) \pm 4.0(syst.) \quad \mathrm{GeV}/c^2.$$

The dominant systematic uncertainty of 3.2 GeV is due to uncertainties in jet energy scale.

This measurement is part of the first measurement of top quark mass in dilepton channel at CDF in Run 2 series of data taking [73, 74]. Our measurement is in agreement with the Run 1 measurement from dilepton channel ($M_{top} = 167.4 \pm 10.3(stat.) \pm 4.8(syst.)$ GeV/c^2) [97]. Moreover, the combined result of four measurements in dilepton channel (where one of these four measurements was our measurement) $M_{top} = 167.9 \pm 5.2(stat.) \pm 3.7(syst.)$ GeV/c^2 is not only with agreement with Run 1 result, but it significantly improved (by $\sim 40\%$) the uncertainty on top quark mass. It should be also noted that the combined dilepton result is also in agreement with measurement in other decay channel (lepton+jets channel) [91], so for now, we don't see any hint for physics beyond the Standard Model.

The CDF data taking continues, and so will also this analysis. Using about tree times larger data sample and making improvements to the method (dividing sample into events with identified and non-identified b-jets, using cross-section dependence on top quark mass, etc.) should enable us to decrease the statistical uncertainty below systematic uncertainty for the first time in dilepton channel. At the end of Run 2, we should be able to obtain the top mass statistical uncertainty of about 2 GeV with this method.

Appendix A

Cut on mass reconstruction probability of the event

In our method, we use the cut on reconstruction probability for a given event, in order to pass mass reconstruction procedure. It means that at least certain percentage of the "smeared" events (trials) must be reconstructed for a given physics event, in order to further consider such event in reconstruction procedure.

First, we will describe the optimization of the cut on reconstruction probability. Later, we will show what are the reasons for events not to be reconstructed.

A.1 Optimization of the cut

The cut of 10 % on reconstruction probability for a given event was introduced in Run I [4] in order to have decent statistics in histogram of reconstructed mass for a trial events. The decent number of events in histogram is necessary, in order to perform spline fit to the distribution as we do it, see Fig. 6.11. Using 1000 trials (to allow the computation to finish in reasonable time) then mean at least 100 trial masses in histogram, when requiring 10 % probability. This is reasonable minimum of entries to require.

However, these days the computers are faster, so we can use order of magnitude larger number of trials (10000) and still retain finite amount of computer time. The typical processing time for one MC sample of typical size of about 2000 events is 2 hours at 2 GHz CPU (using 10000 trials for one event), i.e. about 3.5 seconds per event. Using 10000 trials and still requiring at least 100 trial reconstructed masses in histogram, we can go with the cut down to 1 %. However, lowering the cut from 10 % to 1 % cause not only larger number of signal $t\bar{t}$ events are reconstructed, but also background events. Therefore, final decision on what cut to use should be based on

cut	N^{signal}_{expec}	N^{bckg}_{expec}	ΔM_{top}/ GeV
10 %	12.9	6.5	13.52
5 %	14.0	7.2	13.06
2 %	14.8	7.7	12.74
1 %	15.2	8.0	12.52

Table A.1: The dependence of expected number of signal (N^{signal}_{expec}) and background (N^{bckg}_{expec}) events in 340 pb^{-1} and expected uncertainty of the reconstructed top mass from pseudo-experiments (ΔM_{top}, for input top mass 178 GeV) on value of reconstruction probability cut.

testing the top mass resolution on simulated events as a function of the reconstruction probability cut. We performed such study and the results are summarized below.

We try four different cuts on reconstruction probability: 10%, 5%, 2% and 1% cut. For each of these cuts, we performed whole mass reconstruction procedure, i.e. we produced signal and background templates and from them we obtained fitting function describing these templates. Then, using the expected number of signal and background events for each particular reconstruction probability cut (see Tab. A.1), we performed the pseudo-experiments [1]. In Fig. A.1 and Fig. A.2, we show the reconstructed mass and reconstructed error for the reconstruction probability cut 10 % and 1% respectively. It can be seen that the error on top mass is lowered by roughly 1 GeV for 1% cut which is about 7 % descrease in expected uncertainty. We did see this behavior of lowering the error on top mass when we lowered the cut from 10% to 5% to 2% and finally to 1% (see Tab. A.1). Therefore we consider 1% cut as an optimum.

However, results obtained in this work still used 10 % cut on reconstruction probability. Therefore, in next iteration of analysis, we can expect improvement of about 7 % when using 1 % cut on reconstruction probability.

A.2 Reasons for events to not be reconstructed

In Fig. D.2, it can be seen the histogram of reconstruction probability for simulated events. It shows that the reconstruction probability is roughly constant with very high spike for 0 % reconstruction probability. These are the events, where there was none of the 10000 trials reconstructed. We will concentrate on these events and try to understand the reasons why such events are not being reconstructed [2].

---

[1] for description of pseudo-experiments see Chap. 8

[2] of course, because we use numerical method for solving the equations, the reconstruction probability can be increased by releasing conditions of convergence, see Sec. 6.1. This way, we could obtain

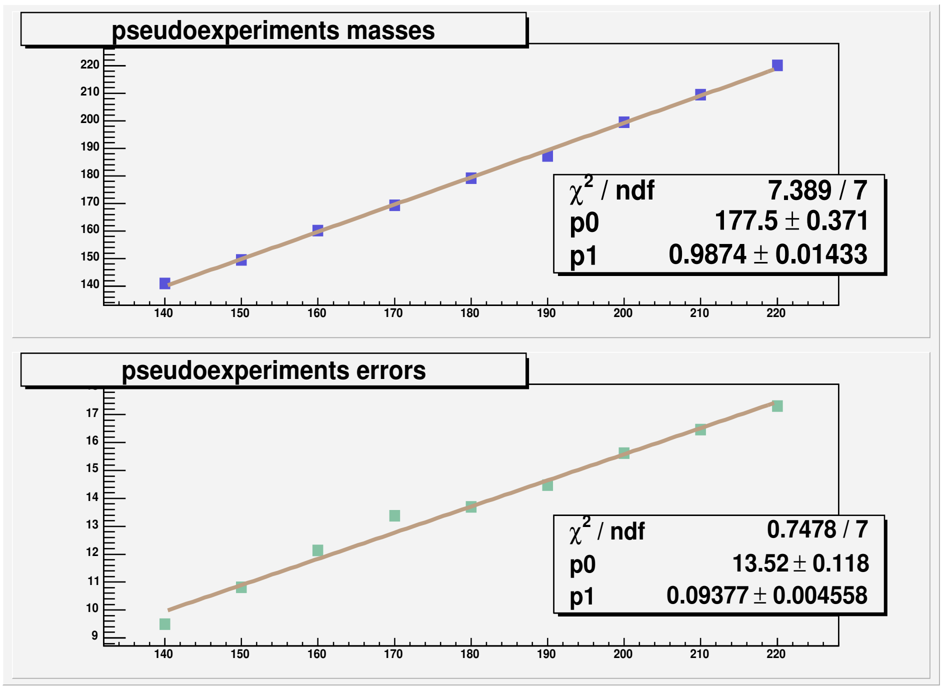


Figure A.1: The dependence of reconstructed mass (above) and reconstructed uncertainty on top mass (down) from pseudo-experiments as a function of input top mass. The plots correspond to the 10 % cut on probability of reconstruction.

First, we will perform the mass reconstruction on parton level, i.e. we will use generated kinematic characteristics of particles rather than reconstructed. In such case, the Tab. A.2 shows that events which have 0 % reconstruction probability (i.e. no trial for a given event was reconstructed) are almost exclusively the events which don't meet our requirements we put on the events in reconstruction procedure ($|M_t - M_{\bar{t}}| < 2$ GeV, $|M_W - 80.4 \text{ GeV}| < 3$ GeV).

As a next step, we performed the study on simulated events which pass all the chain of generation, CDF detector simulation and reconstruction. The results are summarized in Tab. A.3. Here, we divided the events into more categories which we thought can be reasons for the event not to be reconstructed. We can see that there are two other reasons comparing to the previous case at parton level. These are the events where the jet is not associated to b-quark (very probably, wrong jet was used in mass reconstruction) or just the jet was really badly reconstructed (factor of 2 difference in energy between quark and associated jet). Once again, the listed categories are

---

even around 100 % reconstruction probability. However, the return value would be to large extent "random" number

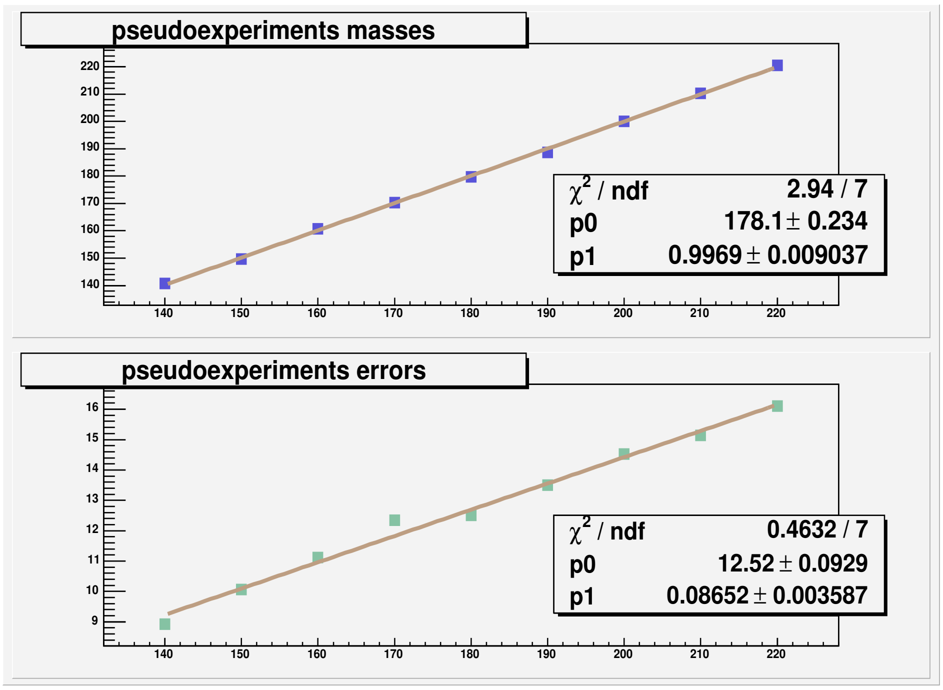


Figure A.2: The dependence of reconstructed mass (above) and reconstructed uncertainty on top mass (down) as a function of input top mass. The plots correspond to the 1 % cut on probability of reconstruction.

event category	N_{ev}	$N_{ev}^{0\%}$	percentage ($N_{ev}^{0\%}$ / N_{ev} * 100)
$\|M_t - M_{\bar{t}}\| < 2$ GeV	4836	87	1.8 %
$\|M_W - 80.4$ GeV$\| < 3$ GeV	1634	26	1.6 %
sum of above	6470	113	1.75%
rest of events	2879	2	0.07 %
total	9349	115	1.23 %

Table A.2: Reconstruction probabilities for different selected categories. The study was performed on parton level. N_{ev} is the total number of events which fail into particular category and $N_{ev}^{0\%}$ is the number of events which have 0 % reconstruction probability. Note, the categories are exclusive, i.e. if event failed into first category, it is not considered in further categories.

exclusive.

event category	N_{ev}	$N_{ev}^{0\%}$	percentage ($N_{ev}^{0\%}$ / N_{ev})
events with τ lepton	1444	52	3.6 %
$\|M_t - M_{\bar{t}}\| < 2$ GeV	4836	337	7 %
$\|M_W - 80.4 \text{ GeV}\| < 3$ GeV	1634	113	6.9 %
no association jet–quark	666	84	12.6 %
$\|E_{jet} - E_{quark}\|/E_{jet} > 2$	147	20	13.60 %
sum of above	8727	606	6.94%
rest of events	2066	50	2.40 %
total	10793	656	6.08 %

Table A.3: Reconstruction probabilities for different selected categories. The study was performed on fully simulated events.

Appendix B

Dependence of template fit parameters

As pointed out in Sec. 7.1, we incorporate the global fit to the signal templates which is combination of the Gauss and Landau distributions with the five parameters (see Eq. (7.1)). We assumed that these parameters depend linearly on top mass. Here, we will show that the dependence is indeed linear.

The global fit to the templates is performed in such a way that all the templates are fitted simultaneously. The five fitting parameters are assumed to be linearly dependent on top mass, so in total we have 10 parameters - slope and offset for each of the 5 linearly dependent parameters.

The proof, that five fitting parameters are linearly dependent on top mass will be performed the following way. We will fit each template separately by the combination of Landau and Gauss function. In doing so, five fitting parameters are free parameters for each of the fitted template, i.e. parameters for each sample are independent from other ones. However, by plotting each parameter as a function of input top mass, we will see whether there is any dependence of these parameters on input top mass and if yes, what is their actual dependence.

Note, the 1st (p_1) and 2nd (p_2) parameter is the mean value and width of the Landau distribution, respectively. The same holds for the p_4 and p_5 parameter, this time for Gauss function. The parameter p_3 controls the relative importance of these functions in total fit, and particularly in this case , it is fraction of Landau function in total fitting function.

First, we will consider the simpler case when the mean values and widths of both Gauss and Landau function are the same, i.e. $p_1 = p_4,\ p_2 = p_5$. In such case, we have just 3 parameters. As described above, we fit each template separately and obtain the parameters estimates.

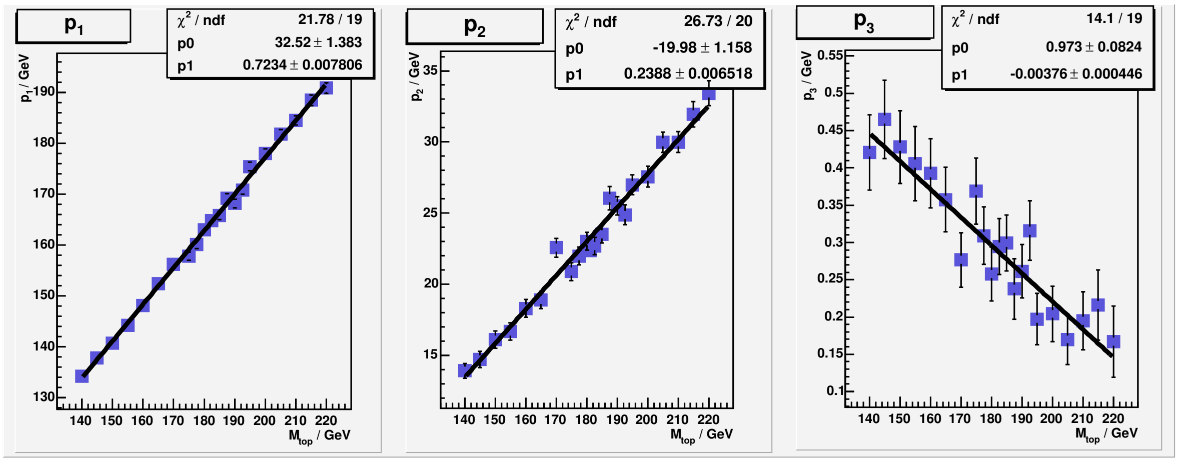


Figure B.1: Top mass dependence of template fit parameters. Here we consider case of only 3 parameters. The parameters in the plot are in the following order (from the left): Gauss and Landau mean value (p_1), Gauss and Landau width (p_2), fraction of Landau function in total fitting function (p_3).

The dependence of these parameters on input top mass together with linear fits can be seen in Fig. B.1. From there, it is clear that the linear fit correctly describes all three parameters as a function of top mass.

Next, we will consider more general case, when the mean values and widths of Landau and Gauss function are different. In such case, there are 5 parameters in the fit. The top mass dependence of these parameters is in Fig. B.2, B.3 and B.4.

We can see once again the linear dependence of these parameters on top mass. However, it's clear, that fit with more parameters (5 comparing to 3) has more freedom

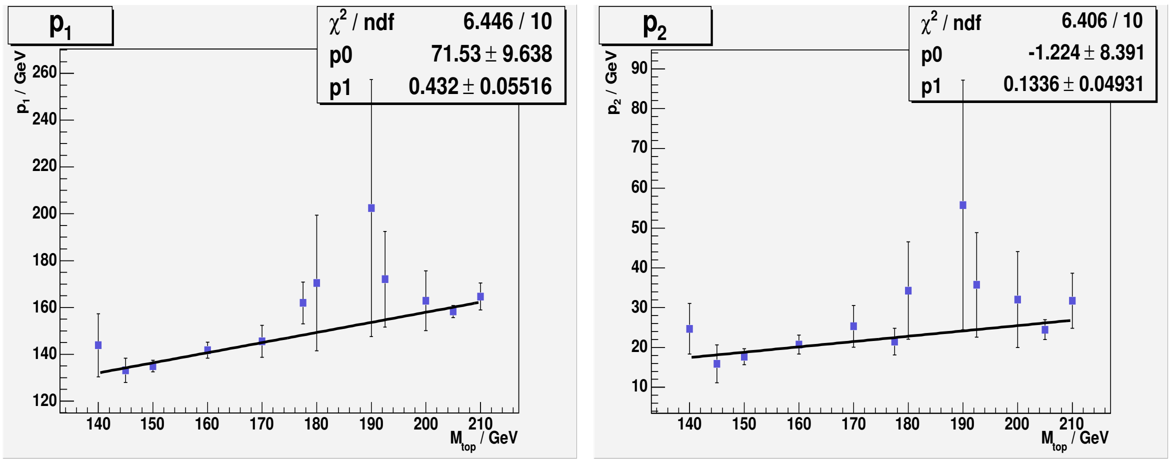


Figure B.2: Top mass dependence of fitting parameter p_1 (left) and p_2 (right).

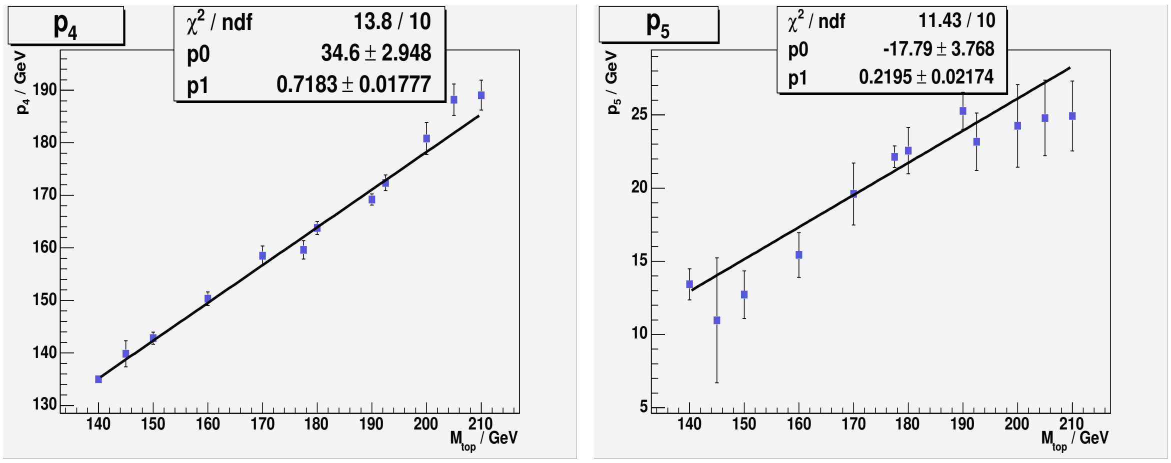


Figure B.3: Top mass dependence of fitting parameter p_4 (left) and p_5 (right).

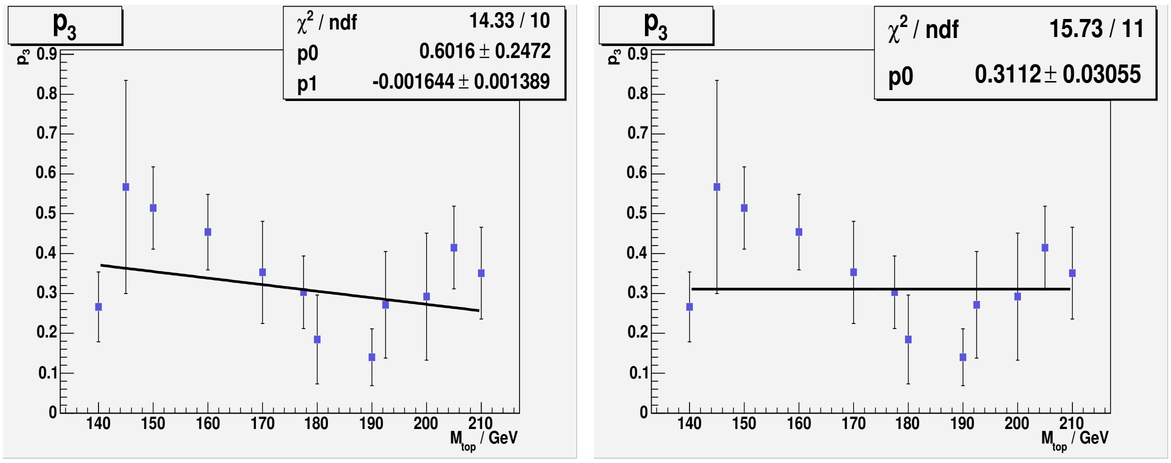


Figure B.4: Top mass dependence of fitting parameter p_3 together with the linear (left) and constant (right) fit.

in fitting procedure and therefore the dependence is not so clear as in case with 3 parameters. Some parameters (mainly, parameters for Landau mean and width) for some fits have large uncertainties in case of 5 parameters fit. However, the overall dependence is still linear.

One specialty is the parameter p_3. There is no strong linear dependence and it seems that it is equally appropriate to use constant top mass dependence for this parameter (see Fig. B.4). Moreover, in practice, we have seen that the fit have been often unstable (the fitting procedure ends without convergence, the error calculation fails, etc.) when we used the linear dependence of parameter p_3. Therefore, we decided to use the linear dependence for all others parameters and to consider parameter p_3 independent on top mass. In such case, when we perform simultaneous fitting of all templates, we have all together 9 parameters: slope and offset for parameters p_1, p_2, p_4, p_5 and just offset for parameter p_3, see Eq. (7.2).

Appendix C

Releasing selection cuts for background events

In order to obtain reasonable (not statistically limited) shape for background templates, we need certain number of events. Ideally, we would like to have similar number of events in background samples as we have in the signal samples. This is roughly 1000 – 2000 events per sample after mass reconstruction, see Sec. 7.1 [1]. Using official CDF MC samples, we gained such number of events only for WW diboson production. As mentioned before, this is because only inside WW events we've got generated all basic objects needed to pass $t\bar{t}$ event selection (2 leptons, 2 jets and missing E_T coming from neutrino). On the other side, Drell-Yan and FAKE events can pass the $t\bar{t}$ event selection only in case there was mismeasurement or bad reconstruction, see Sec. 4.2. They are missing one of the signatures of $t\bar{t}$ events (Drell-Yan events don't have neutrinos in final state, while FAKE events have only one real lepton in final state).

The number of FAKE events passing event selection and mass reconstruction is far too low (~ 10 in total, see Tab. C.1) for a template, using Monte-Carlo samples with production of W boson and associated production of at least 3 jets (we call it "$W+ \geq 3$ jets" sample). The Drell-Yan Monte Carlo sample consists of several sub-samples (see Tab. C.2), which need to be carefully combined into one template. The weights for different sub-samples are within almost two orders of magnitude, causing spikiness in the combined Drell-Yan template.

In order to have more statistics for the FAKE template, and smoother Drell-Yan template, one could generate more Monte-Carlo events. However, as mentioned above, both FAKE and Drell-Yan events pass event selection due to mismeasurements. Such mismeasurements are rare, and thus the number of generated events would need to be

---

[1] Even, when we would be not so strict, roughly 100 events is really the minimum to obtain usable templates.

huge.

The other possibility of increasing statistics in templates is to release event selection cuts. In this appendix, this possibility together with drawbacks, will be discussed.

C.1 FAKEs

As we already mentioned, the main source for FAKE events are $W+\geq 3$ jets events, where one jet fakes a lepton. Approximately an order of magnitude smaller are the contributions from processes like $W+b\bar{b}$, $W+c\bar{c}$, QCD or $W+c$ with associated jets production.

Applying the standard event selection procedure (see Sec. 4.2.2) to the $W+\geq 3$ jets Monte Carlo samples, only 26 events pass the event selection. Furthermore, only 13 events pass the mass reconstruction procedure. The used Monte Carlo samples are generated by Alpgen+Herwig, and the specific characteristics are described in the Tab. C.1. These 13 events is really pure number to be able to create template distribution for FAKE events. In order to get more events, we have released event selection cuts. However, releasing event selection cuts may, and very probably will, change the shape of kinematic distributions and consequently also the shape of the template. Therefore, one needs to make sure that the shape of the template does not change significantly. Because there is not enough statistics to compare templates between original and released event selection, we will compare just the kinematic distributions of the events. Because the kinematics define the shape of template, we will assume that if kinematics agree between different selections then also the templates agree.

We have also included "heavy flavor samples" to the FAKE template. These are special $W+\geq 3$ jets samples where the flavor of two quarks is defined to be either $b\bar{b}$ or $c\bar{c}$ quarks. The events from heavy flavor samples become FAKE events in the same way as all the other $W+\geq 3$ jets events: one of the jets is mis-identified as a lepton. We included these heavy flavor samples in our calculation, to increase the total number of events in MC. We assume, the kinematic characteristics are similar in such events comparing to $W+\geq 3$ jets events and that they get also similar reconstructed mass. We will prove it later in the Appendix. The used heavy flavor samples (see Tab. C.1) are generated by Alpgen+Herwig.

The releasing of the selection cuts must be done very carefully. It's clear, we don't want to release any kinematic cut (e.g. cut on jet E_T), because you would directly change your kinematic distributions and also the template. Rather, one would like to release some quality cuts which presumably have small effect on kinematics. According to comparison of event kinematics, following loosening in event selection preserve the kinematic distributions best, and provide reasonable increase in statistics for the template:

- For Central electrons
  - isolation: 0.1 → 999 (change from 0.1 to 999)
    The cut 999 effectively means, that we don't require electron to be isolated.
  - E_{had}/E_{em}: 0.110+0.00090*E_{ele} → (0.055+0.00045*E_{ele})
    This is the ratio of energy deposited in hadronic and electromagnetic calorimeter. We increased the slope and offset by factor two.
- For Plug electrons
  - isolation: 0.1 → 0.5
    The cut means that we allow for up to 50 % of additional energy in the cone around the electron (comparing to original cut allowing 10 % more energy).
  - E_{had}/E_{em}: 0.5 → 0.10
- For Muons
  - isolation: 0.1 → 999
    effectively, cut removes isolation

In our mass reconstruction procedure, we use leptons,jet four-momenta and the missing transverse energy $\not\!\!E_T$. The above released cut should affect only the leptons P_T and $\not\!\!E_T$. The comparison of these variables between events selected with original and released cuts is presented in Fig. C.1. Events from heavy flavor samples are included in the sample of events which pass released selection cuts while they are not included in the original selected sample. The normalized distributions agree within errors. The number of events passing standard event selection N^{selec}_{orig} and mass reconstruction N^{recon}_{orig} together with the number of events passing released event selection $N^{selec}_{released}$ and mass reconstruction $N^{recon}_{released}$ are shown in Tab. C.1. Listed are all generated MC samples used to create FAKES template.

The similarity of template from heavy flavor samples and the template from other $W+\geq 3$ jets samples is presented in Fig. C.2. We make also the plot of the templates for FAKE events selected with original selection and with released cuts in Fig. C.3. However, as we said, the 13 events in original selection don't allow for meaningful comparison.

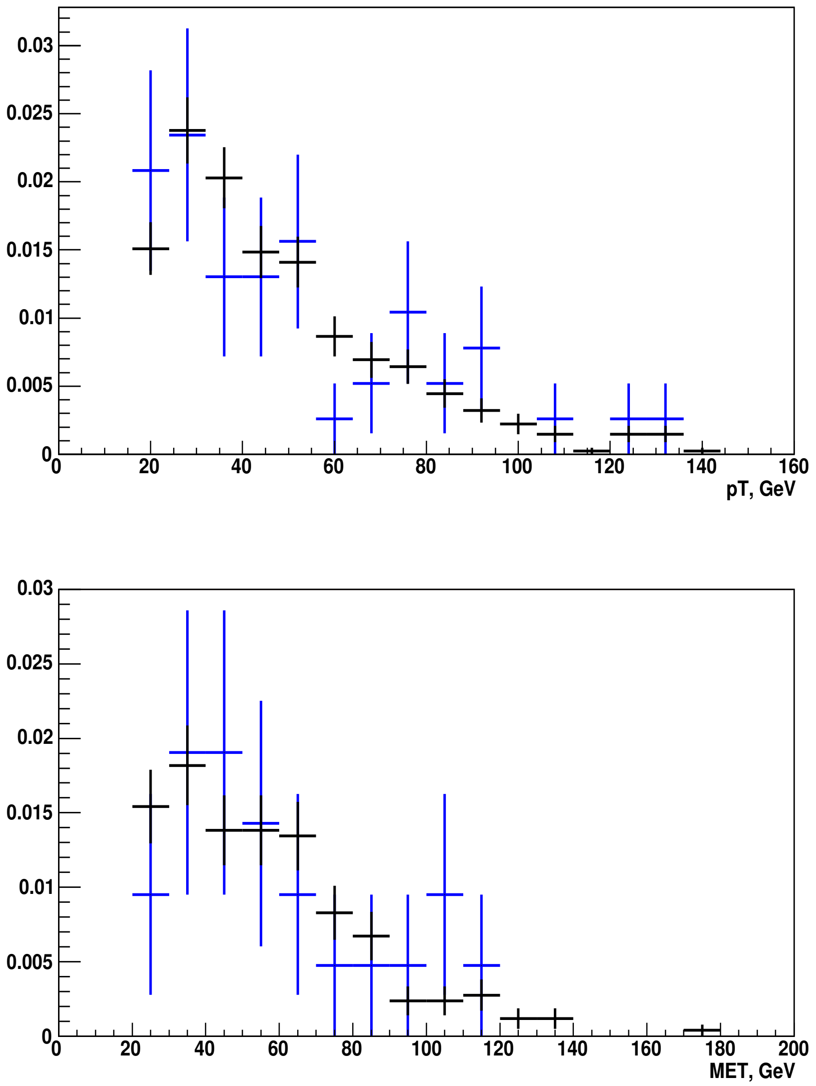


Figure C.1: Comparison of lepton P_T and $\not{E}_T$ between events which pass standard event selection (blue histogram) and released event selection (black histogram). Note, the distributions are normalized.

Up to now, we have shown that our released selection does not distort the kinematics of originally selected MC events. As another check, we will compare the events from MC to data events. We will compare MC FAKE template to a template, which is constructed using fake-able data events (for definition of fake-able data event see Appendix C.1.1).

There are totally 80 fake-able data events, and 62 out of them pass the mass recon-

generated MC sample	N_{tot}	N^{selec}_{orig}	N^{recon}_{orig}	$N^{selec}_{released}$	$N^{recon}_{released}$
W→ e + 2 partons	180k	1	1	2	1
W→ e + 3 partons	265k	8	4	15	8
W→ e + 4 partons	243k	8	6	18	13
W→ μ + 2 partons	252k	1	0	2	0
W→ μ + 3 partons	291k	3	1	10	3
W→ μ + 4 partons	287k	5	1	27	7
W→ $e + c\bar{c}$ + 2 partons	194k			18	9
W→ $e + b\bar{b}$ + 2 partons	235k			84	54
W→ $\mu + c\bar{c}$ + 2 partons	255k			23	13
W→ $\mu + b\bar{b}$ + 1 parton	202k			20	10
W→ $\mu + b\bar{b}$ + 2 partons	239k			38	27
total	2.64M (1.52M)	26	13	257	145

Table C.1: Number of FAKE events passing standard and released event selection and mass reconstruction. The total number of events generated in each sample (N_{tot}) is also shown ('k' means thousands, 'M' means millions).

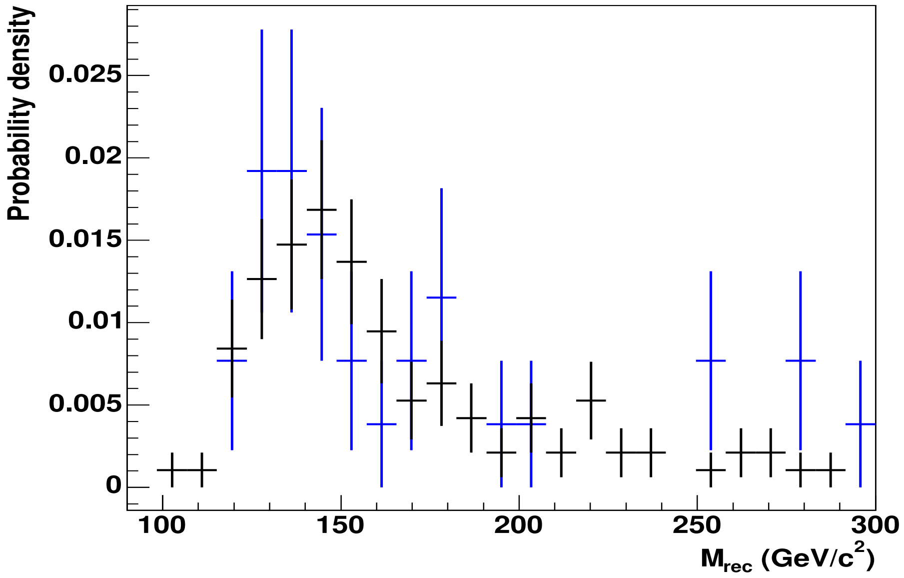


Figure C.2: Comparison of FAKE templates from original samples (blue) and from heavy flavor samples (black). The released event selection has been used for both templates.

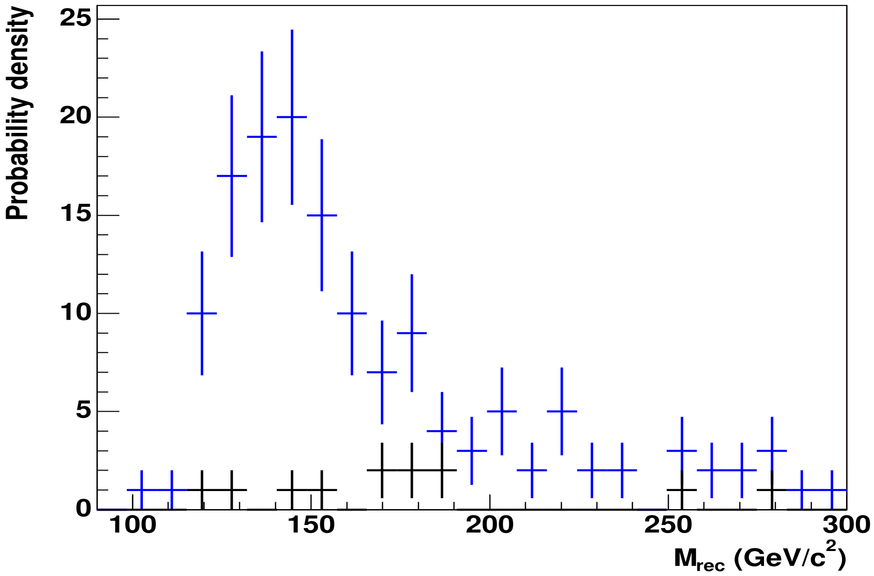


Figure C.3: The FAKE template. The blue histogram depicts the template with released cuts and from both original and heavy flavor samples. The black histogram shows the raw top masses for the 13 events passing original event selection.

struction. The comparison between Monte-Carlo and data FAKE templates is shown in Fig. C.4. The Monte-Carlo and data templates agree within statistics. This agreement also proves that $W+\geq 3$ jets events are the main source of FAKE events or at least that if there are other significant sources of FAKE events, they don't change the template shape.

C.1.1 Definition of fake-able data event

The fake-able data events have a good lepton plus a fake-able object. Fake-able events must then pass the same $Z-veto$, MET and $L\ cut$ as for the standard selection (see Sec. 4.2.2) using the fake-able object as the second lepton in the event.

The fake-able object could be a jet, a track, or a lepton which has passed some minimal cuts.

A jet faking an electron must contain a highly energetic object, leaving most of its energy in the electromagnetic calorimeter, like a π^0, η and another charged object with a track (π^+, π^-). The other jet fragments do not contribute much to the hadronic energy.

A jet might fake a muon for many reasons: a punch-through (a hadron which

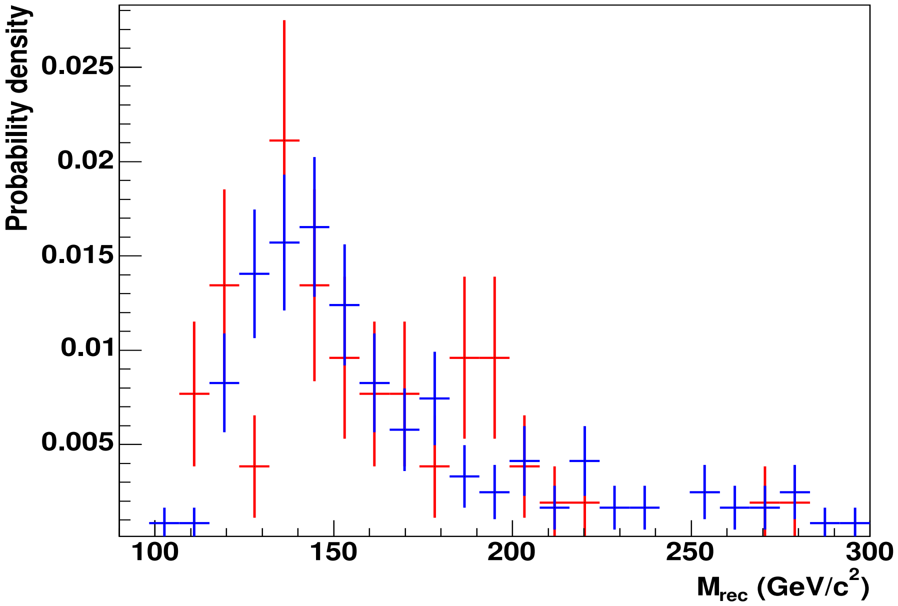


Figure C.4: Comparison between Monte-Carlo and data FAKE templates. The blue histogram depicts Monte-Carlo and the red histogram depicts data. The released cuts has been used for the Monte-Carlo template.

reaches the muon chambers), a decay-in-flight (low P_T kaon decays and a high-P_T track is mis-reconstructed) or muons from semileptonic decays of quarks.

Fake-able object used in this analysis is object which is identified as electron passing minimal set of cuts ($E_T > 20$ GeV, the ratio $E_{HAD}/E_{EM} \leq 0.125$ and for appropriate categories (CEM, PHX) also $Isol < 0.1$) or is identified as muon which pass some minimal set of cuts (cuts on P_T,$|z_0|$,$|d_0|$, COT track quality cuts, Iso_P_T from Tab. 4.3 and $(E_{HAD} + E_{EM})/P < 1$).

C.2 Drell-Yan

We used several Drell-Yan MC sub-samples in order to create Drell-Yan template. These samples were generated by Alpgen+Herwig and are listed in Tab. C.2. The acceptance is very low for the low mass samples ($M_{\ell\ell} \in (10, 75)$), higher for medium mass samples ($M_{\ell\ell} \in (50, 130)$) and highest for the high mass samples ($M_{\ell}\ell \in (105, 800)$). The overall Drell-Yan template, where subsamples are combined according corresponding cross-sections is dominated by the medium mass samples and have been obtained the following way.

background	$M_{\ell\ell}$ range	n_{tot}	N^{orig}_{selec}	N^{orig}_{recon}	N^{rel}_{selec}	N^{rel}_{recon}	σ	w
$Z/\gamma^* \to ee$+2p	(15,75)	265k	2	1	5	3	50.5	93.4
	(50,130)	494k	27	14	39	20	25.0	25.0
	(105,800)	292k	179	55	196	57	0.63	1.07
$Z/\gamma^* \to \mu\mu$+2p	(15,75)	268k	6	4	8	4	50.5	92.4
	(50,130)	293k	43	27	43	27	25.0	42.3
	(105,800)	118k	169	48	187	57	0.63	2.64
$Z/\gamma^* \to \tau\tau$+2p	(15,75)	294k	0	0	0	0	50.5	84.2
	(50,130)	288k	44	36	45	36	25.0	43.0
	(105,800)	111k	62	46	66	49	0.63	2.81
total		2.42M	532	231	731	325		

Table C.2: Number of Drell-Yan events passing original and released event selection and mass reconstruction for each particular MC sample ('k' means thousands, 'M' means millions). There are listed the cross-sections σ of the corresponding processes and also the weights w we use to combine the samples together.

The Z $\to$ ee+2p and Z $\to$ $\mu\mu$+2p samples were combined into one template by normalizing the number of events by cross section, acceptance and mass reconstruction probability. Practically, it was done by creating the template for each subsample and weighting them by a weights w shown in Tab. C.2. The weights w were calculated with the following formula

$$w = \frac{n_{tot}}{494k}\,\sigma, \tag{C.1}$$

where n_{tot} is the size of data set (number of generated events) and σ is the cross section of the corresponding process.

Z $\to$ $\tau\tau$+2p samples were combined exactly the same way as above to obtain one Z $\to$ $\tau\tau$ template.

The overall Drell-Yan template was obtained by normalizing Z $\to$ $ee/\mu\mu$ and Z $\to$ $\tau\tau$ templates and weighting each of them by the expected number of the events passing both event selection and mass reconstruction.

Due to low statistics and the weights differing almost two orders of magnitude, some particular events have high contribution to overall template and causing the spikiness of the distribution. In order to smoothen this spikiness, we released the selection cuts similarly as for FAKE events. This time, we released just isolation cut on leptons from 0.1 to 0.5. The number of events passing standard event selection N^{orig}_{selec} and mass reconstruction N^{orig}_{recon} together with the number of events passing released event selection N^{rel}_{selec} and mass reconstruction N^{rel}_{recon} are shown in Tab. C.2. The comparison

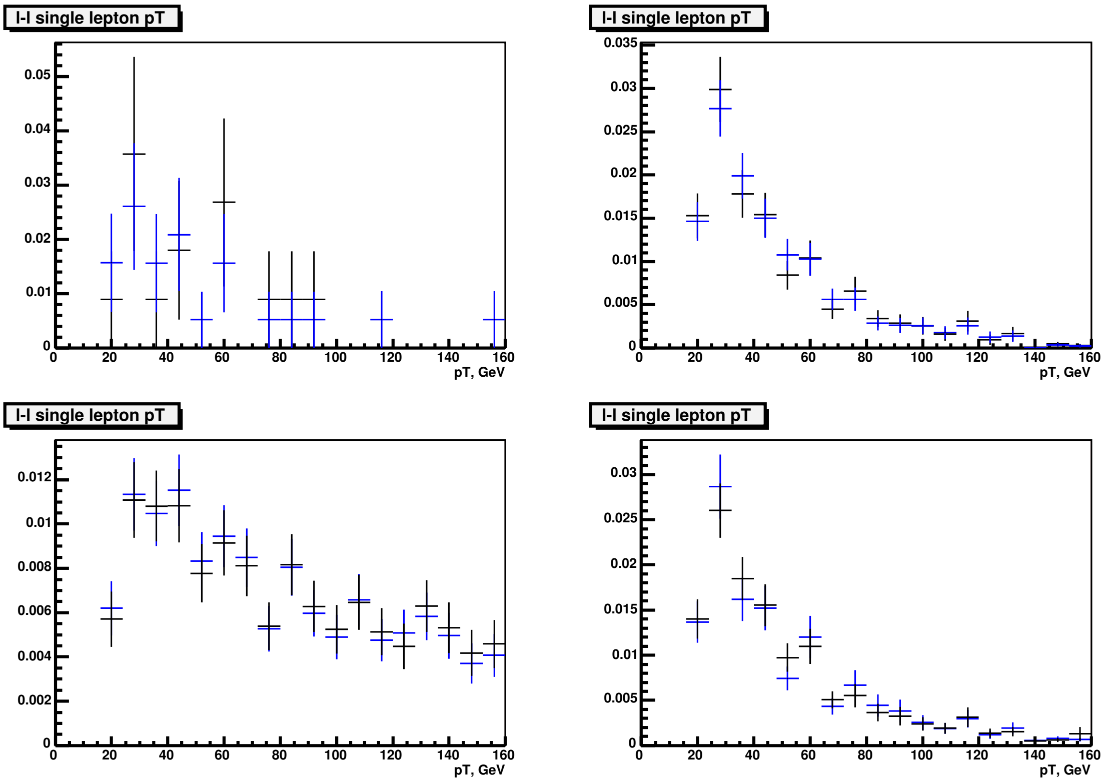


Figure C.5: Lepton p_T comparison for events passing original (black) and released cuts (blue). The upper left plot is for low mass sample, the upper right plot is for medium mass sample and the lower left plot is for high mass sample. The lower right plot is a distribution combined from weighted subsamples.

of lepton p_T, missing E_T and jet E_T between events passing the original and the released cuts is shown in Fig. C.5, Fig. C.6, and Fig. C.7, respectively. The jet E_T distribution is there only as a consistency check, because we released the cuts only for leptons, so we don't expect any significant change in kinematics of jets. The distributions agree between these two different selections. The contribution to Drell-Yan template from low mass samples is still limited by statistics. There are only 13 events from all low mass samples passing the released event selection. By definition, Drell-Yan events have two leptons, so one cannot expect significant increase in statistics by releasing lepton identification cuts.

The statistics for low mass samples was increased in the following way, called inverted missing E_T cut selection [98] hereafter:

1 Invert missing E_T cut, i.e. require missing $E_T < 25$ GeV, and remove H_T cut.

2 Smear the jet E_T (corrected as in standard event selection, see Sec. 4.2.2) for the

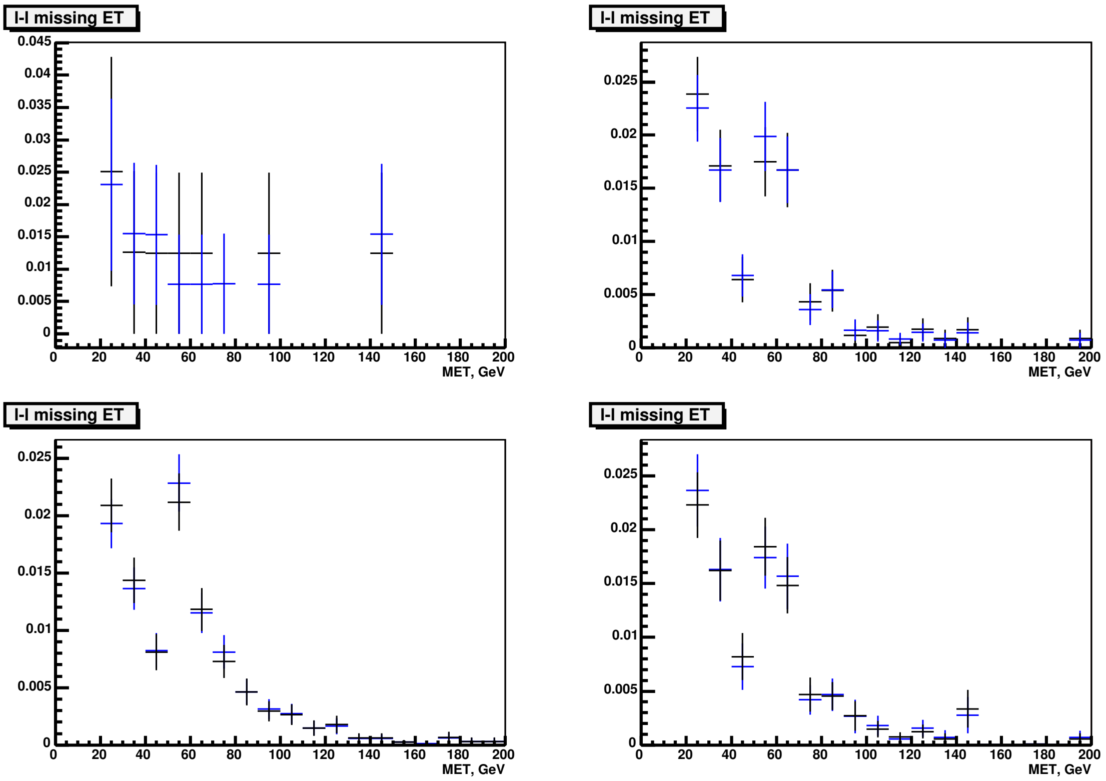


Figure C.6: Missing E_T comparison for events passing original cuts (black) and released cuts (blue). The upper left plot is for low mass sample, the upper right plot is for medium mass sample and the lower left plot is for high mass sample. The lower right plot is a distribution combined from weighted subsamples.

two highest E_T jets and recalculate transverse missing energy $\not\!\!E_T$. This step is repeated until $\not\!\!E_T$ cut is passed $\not\!\!E_T \geq 25$ GeV. The maximum number of iterations is 500. The event is cut out, if it did not achieve $\not\!\!E_T \geq 25$ GeV in 500 iterations.

3 Apply H_T cut.

4 After the event selection, the smeared jet E_T is corrected as in standard event selection for remaining jet corrections, i.e. underlying event correction and out of cone correction.

The jet E_T, missing E_T and lepton P_T distributions are compared between events passing released selection and above described inverted $\not\!\!E_T$ selection in Fig. C.8. The missing E_T distribution for events from the inverted missing E_T cut event selection favors lower values than the events from normal selection. However, we haven't come up with a better idea how to increase statistics for low mass sample.

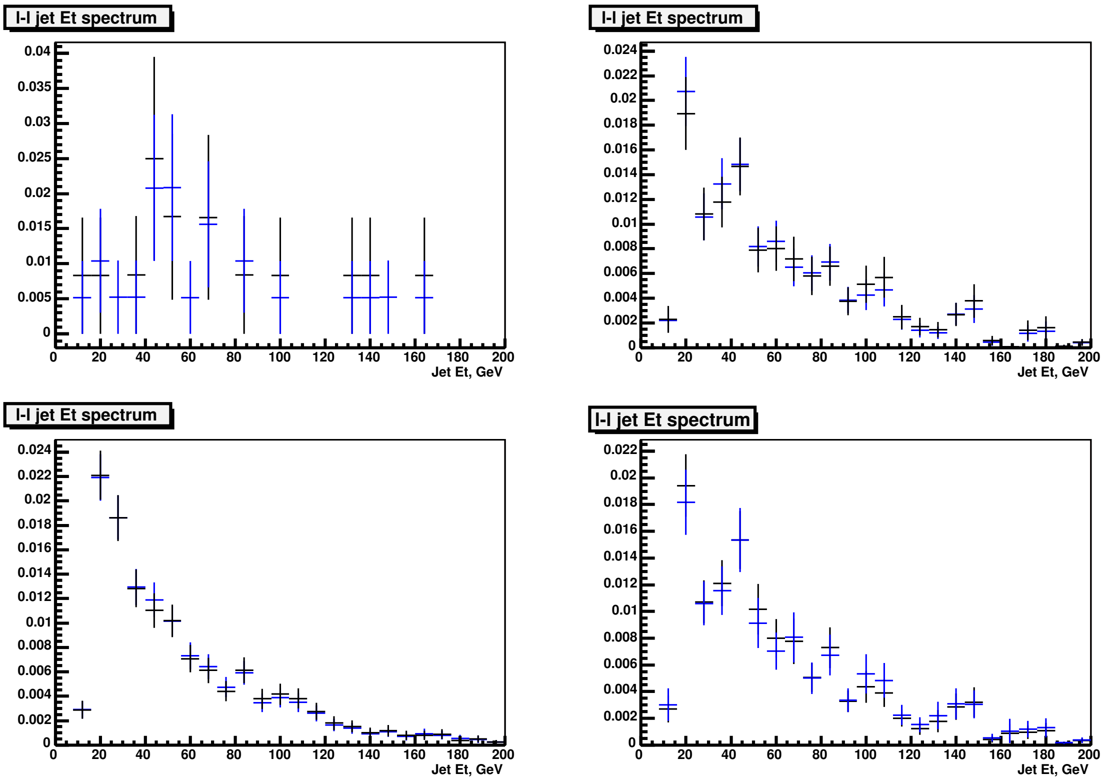


Figure C.7: Jet E_T comparison for events passing original cuts (black) and released cuts (blue). The upper left plot is for low mass sample, the upper right plot is for medium mass sample and the lower left plot is for high mass sample. The lower right plot is a distribution combined from weighted subsamples.

After applying the inverted missing E_T cut event selection, the normal procedure to reconstruct top mass was performed. The template from this procedure was normalized to correspond to the numbers of events passing the released event selection.

There is one more complication in using more sub-samples, because of overlapping in $Z_{\ell\ell}$ range between sub-samples. In the ideal case, the invariant mass of the two measured leptons corresponds to the mass of Z-boson or virtual γ^*. Thus, the double counting of events was minimized by applying cuts on invariant lepton mass distributions [2]. The cuts were only applied for events originating from Z → ee+2p and Z → $\mu\mu$+2p samples, and only for identified ee and $\mu\mu$ events [3]. For low mass sample, the lepton invariant mass was limited to be $M_{l\bar{l}} < 70$ GeV, 70 GeV $\leq M_{l\bar{l}} \leq 120$ GeV

---

[2] We applied the cut on reconstructed invariant lepton mass. Applying the cut on HEPG level would be more correct, but at our level of precision, we do not expect to see significant difference.

[3] For Z → $\tau\tau$ samples, one cannot make a cut on reconstructed dilepton invariant mass because the measured leptons come from τ decays. At HEPG level, the cut on $m_{\tau\tau}$ could have been applied though.

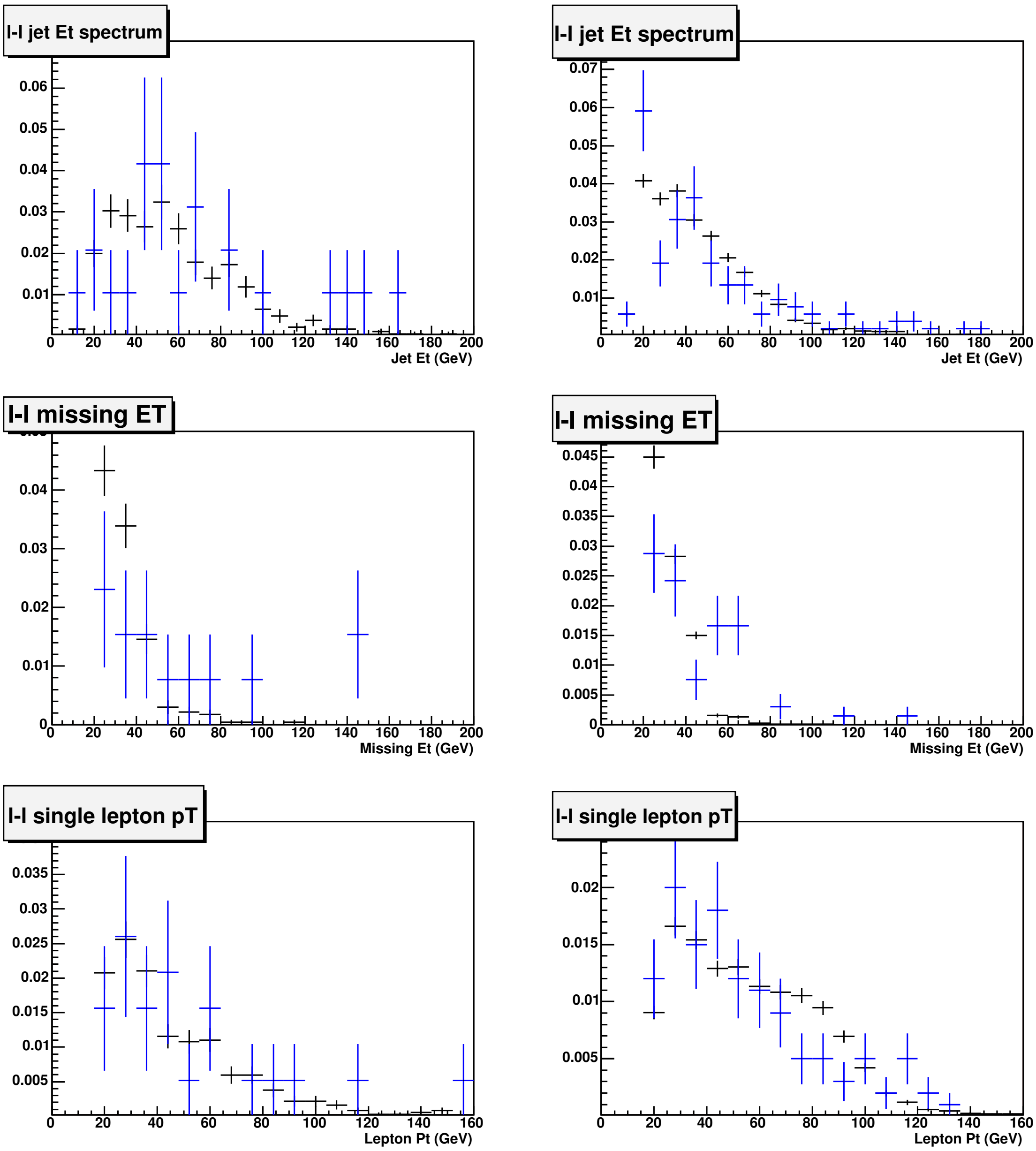


Figure C.8: Jet E_T, missing E_T and lepton P_T normalized distributions for Drell-Yan events. The blue histograms correspond to normal event selection with released isolation cut and the black histograms correspond to selection with inverted missing E_T cut. The distributions on left side include low mass ee and $\mu\mu$ events and the distributions on right side included medium mass ee and $\mu\mu$ events.

for medium mass sample and $M_{l\bar{l}} > 120$ GeV for high mass sample.

The Z $\rightarrow ee/\mu\mu$ and Z $\rightarrow \tau\tau$ templates are shown in Fig C.9. The combined Drell-Yan template is presented in Fig. C.10. It is just smooth enough for top mass analysis.

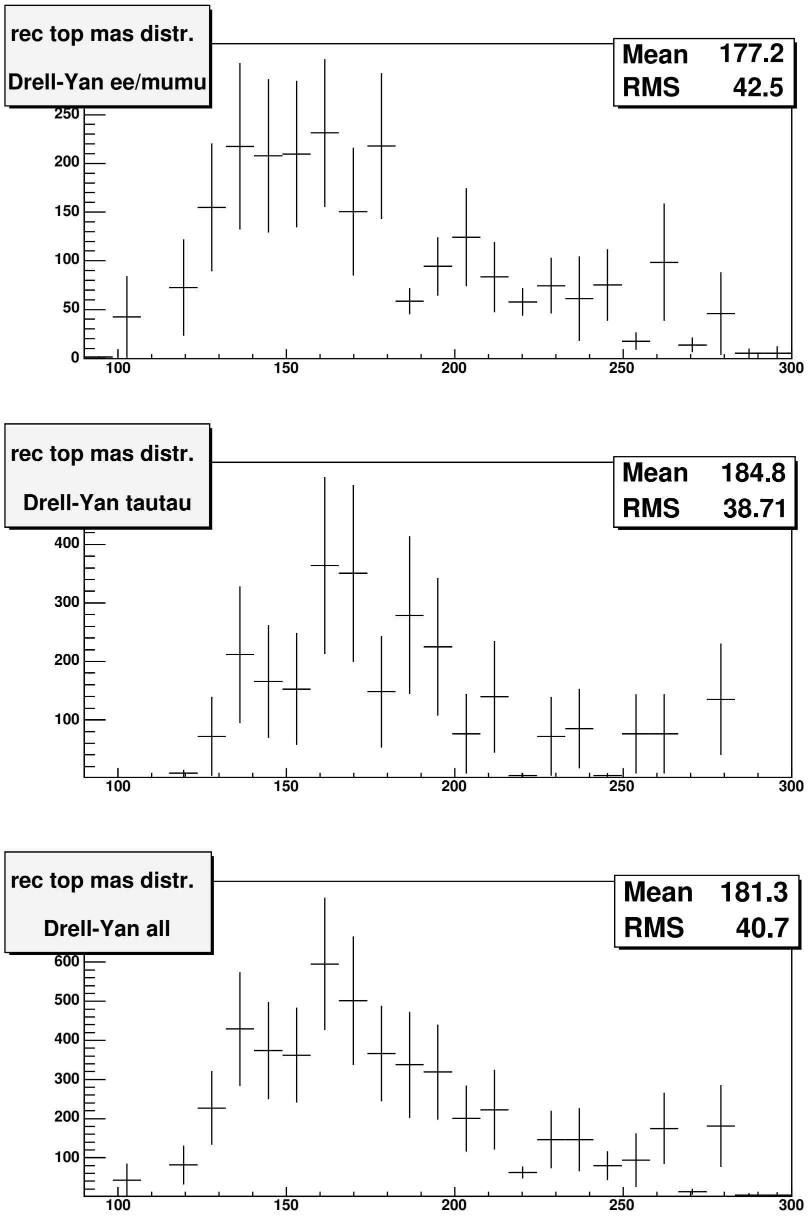


Figure C.9: Drell-Yan templates. The upper plot is the Z $\rightarrow ee/\mu\mu$ template, the middle plot is the Z $\rightarrow \tau\tau$ template and the lower plot is the combined Drell-Yan template.

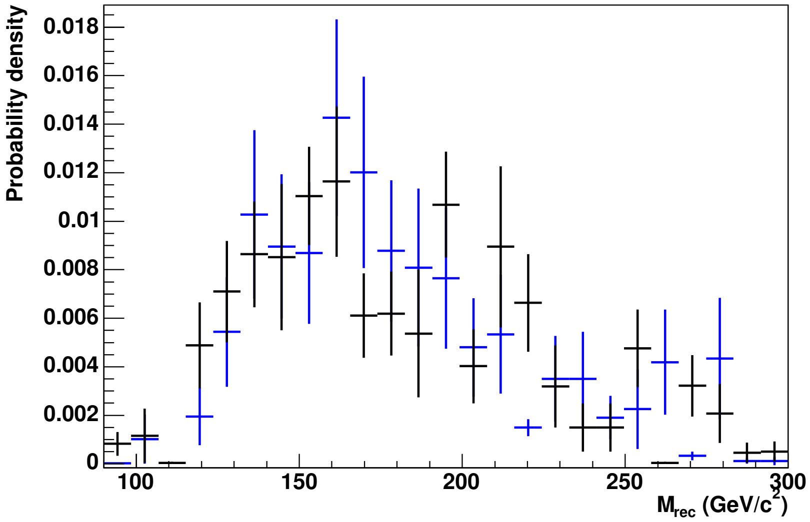


Figure C.10: Drell-Yan template with original event selection and with released event selection. The black histogram is with original cuts, and the blue histogram is with released cuts.

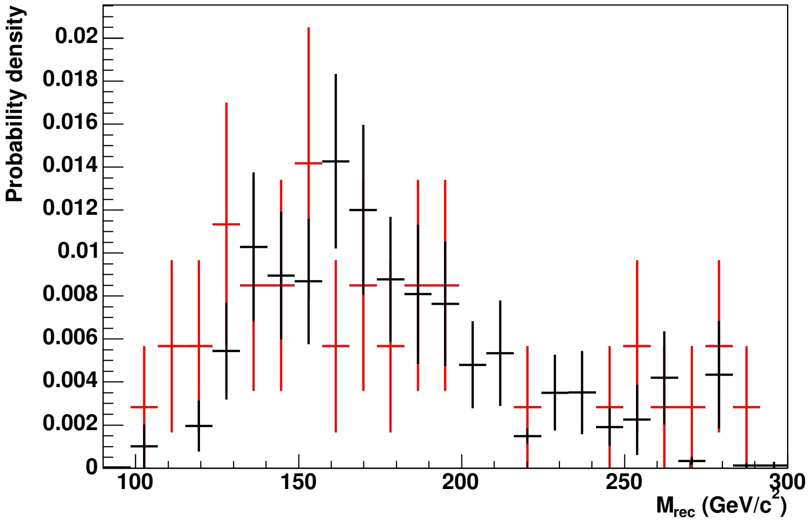


Figure C.11: Drell-Yan data (red) and Monte-Carlo (black) templates.

The comparison between data and Monte-Carlo was performed also for Drell-Yan events. We applied inverted missing E_T selection to data. The selection passed 66 data events. The shape of templates (Fig. C.11) as well as lepton p_T, missing E_T and jet E_T (Fig. C.12) are similar for Monte-Carlo and data events.

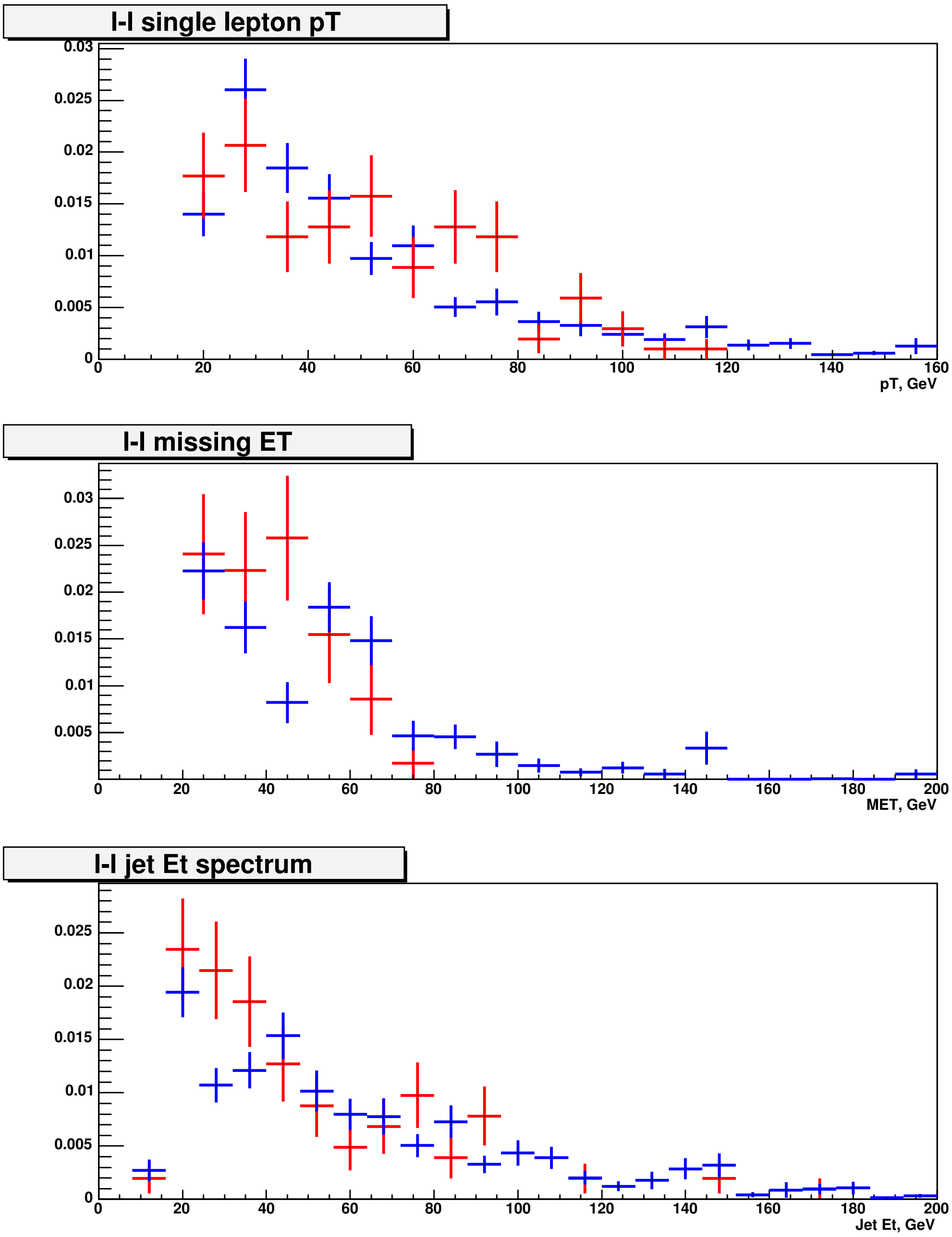


Figure C.12: Lepton P_T, missing E_T and jet E_T distributions for Drell-Yan events. The blue histograms correspond to Monte-Carlo and the red histograms correspond to data.

In summary, it was shown that the events passing released event selection have similar lepton p_T, missing E_T and jet E_T distributions and consequently also the templates as the ones passing original event selection. The Monte-Carlo templates were compared to templates from data, and they agree within statistics.

The statistics gained is just good enough to describe the shape of Drell-Yan and FAKE templates. In the future rounds of analysis, larger Monte Carlo samples will be needed.

Appendix D

Mass reconstruction probability for data and MC

In this appendix, we will address the question of virtual inconsistency between data and MC reconstruction probability.

As was already pointed out in Sec. 10.1, the probability that event pass mass reconstruction procedure is unexpectedly high for data events comparing to MC simulated events.

In simulated events, the reconstruction probability is 75.0 ± 0.4 % for signal $t\bar{t}$ events (see Fig. 7.2) and 61 ± 12 % for background events (see Tab. 7.4). Thus, the combined reconstruction probability for expected number of signal ($n^{exp}_{t\bar{t}}(m_{top} = 175$ GeV$) = 17.2$) and background events ($n^{exp}_{bckg} = 10.5$) is 70 ± 5 % (see Tab. 7.6). In CDF data sample, we reconstruct the top mass in 30 out of 33 events, which gives 91 ± 5 % reconstruction probability. However, the 33 $t\bar{t}$ candidates we see in data is a little bit more than expected (27.7 events, see Tab. 7.6). Therefore, to perform fair comparison, it's appropriate to calculate combined MC reconstruction probability for number of events seen in data. In such case, assuming the number of background events is estimated correctly ($n_{bckg} = 10.5$), there are 33 - 10.5 = 22.5 $t\bar{t}$ events [1]. The combined reconstruction probability for MC is than $71 \pm 5\%$. The difference between data and MC reconstruction probability is then 20 ± 7 %, which is difference almost 3 standard deviations.

As we saw in Sec. 6.2.3, the kinematic characteristics between data and MC events agree within the current data statistics. Moreover, here we compare the kinematic characteristic of events with non-zero reconstruction probability. These are the events, where we were able to reconstruct top mass in at least one trial (out of 10000). The comparison can be found in Fig. D.1. It shows the data is consistent with MC.

---

[1] of course, this is meant in average. In a given experiment, there can not be half of $t\bar{t}$ event.

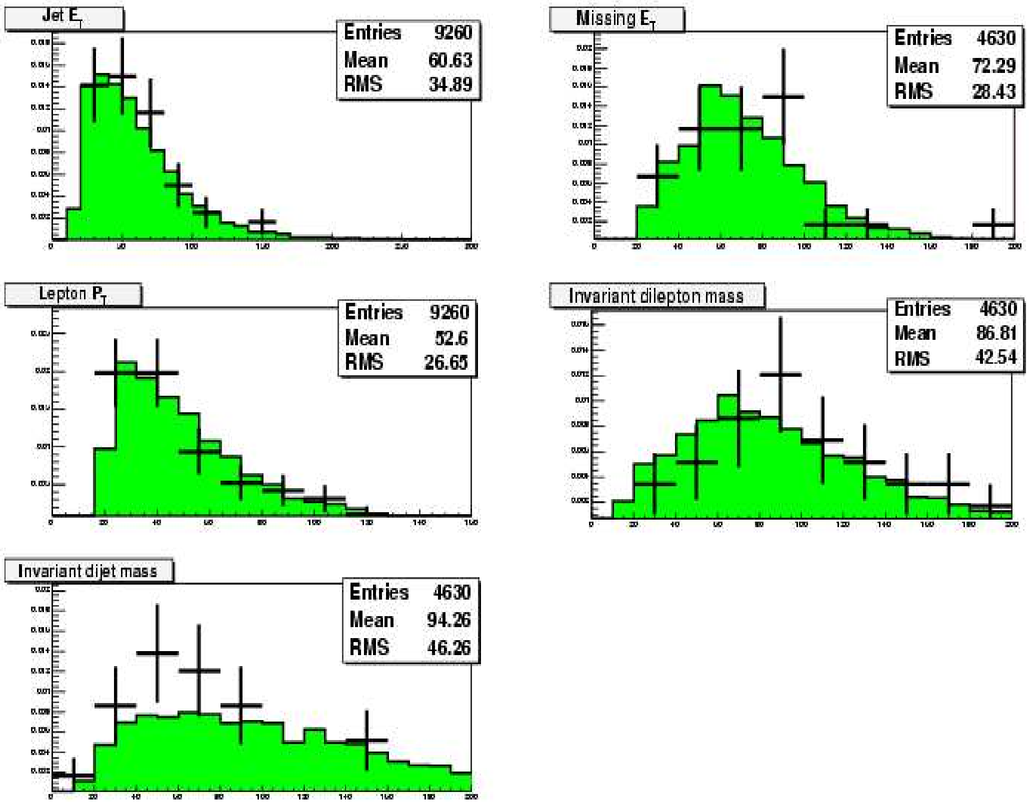


Figure D.1: Comparison of different kinematic characteristics between data and simulated events. Here are compared just the events with non-zero reconstruction probability.

As a next check, we created the MC combined (signal+background) plot for reconstruction probability of trails for a given event and the same plot we created also for data events, see Fig. D.2. The reconstruction probability in this case means what was the probability for mass reconstruction for trials in a given event. Note, event must pass 10 % probability cut to be considered as reconstructed. We can see, the number of events in zero bin are roughly similar, about 10%. However, there is gap in data reconstruction probability distribution in range 0.1 - 0.11 (2nd – 10th bin). This is

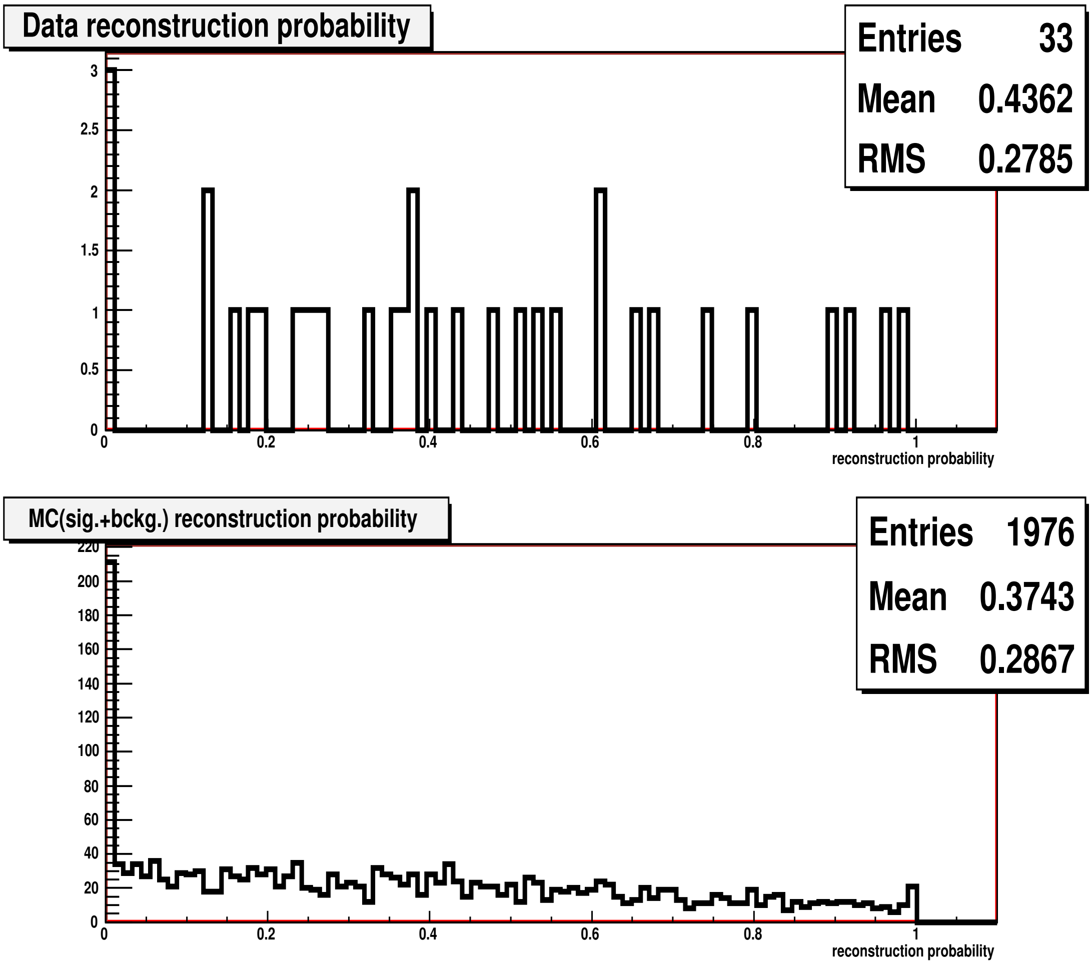


Figure D.2: The histogram of reconstruction probability for "smeared events" of the particular physics events. Shown is a plot for combined MC events (signal and background) (lower plot) and for data events (upper plot).

due to low statistics, however a such gap can be anywhere. By picking randomly the 33 events (total number of events we see in data) from MC reconstruction probability distribution, we see that there is very small probability of $1.21 \pm 0.03\%$ to have such gap between 0.01 and 0.11 (excluding the zero bin), i.e. a gap which we see in data.

Therefore, as we see that data kinematic characteristics are in agreement with MC, we conclude that there is nothing strange with kinematics of data events and the unexpected high reconstruction probability is just consequence of very rare shape of data reconstruction probability. With more data acquainted, we should confirm (or reject) this statement.

Bibliography

[1] F. Abe *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **74** (1995) 2626

[2] S. Abachi *et al.* [D0 collaboration], *Phys. Rev. Lett.* **74** (1995) 2632

[3] S. Herb *et al.*, *Phys. Rev. Lett.* **39** (1977) 252; W. R. Innes *et al.*, *Phys. Rev. Lett.* **39** (1977) 1240

[4] J. Antoš, *Acta Phys. Slovaca* **49** (1999) 127

[5] S. L. Glashow, Nucl. Phys. **22** (1961) 579

[6] S. Weinberg, *Phys. Rev. Lett.* **19** (1967) 1264

[7] A. Salam, *Originally printed in *Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm 1968, 367-377*

[8] S. Willenbrock, *Rev. Mod. Phys.* **72** (2000) 1141 [arXiv:`hep-ph/0008189`]

[9] S. Dawson, arXiv:`hep-ph/0303191`

[10] D. Chakraborty, J. Konigsberg and D. L. Rainwater, *Ann. Rev. Nucl. Part. Sci.* **53** (2003) 301 [arXiv:`hep-ph/0303092`]

[11] W. Wagner, Rept. Prog. Phys. **68** (2005) 2409 [arXiv:`hep-ph/0507207`]

[12] S. Willenbrock, arXiv:`hep-ph/0211067`

[13] LEP Electroweak Working Group arXiv:`hep-ex/0511027`

[14] E. Brubaker *et al.* [Tevatron Electroweak Working Group], arXiv:`hep-ex/0603039`

[15] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, *J. High Energy Phys.* **04** (2004) 068 [arXiv:`hep-ph/0303085`]

[16] N. Kidonakis and R. Vogt, *Phys. Rev.* **D 68** (2003) 114014 [arXiv:`hep-ph/0308222`]

[17] CDF conference note 8148

[18] S. R. Rappoccio, FERMILAB-THESIS-2005-72

[19] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan and S. Weinzierl, *Phys. Rev.* **D 66** (2002) 054024 [arXiv:`hep-ph/0207055`]; Z. Sullivan, *Phys. Rev.* **D 70** (2004) 114012 [arXiv:`hep-ph/0408049`]

[20] V. M. Abazov *et al.* [D0 Collaboration], *Phys. Rev. Lett.* **98** (2007) 181802 [arXiv:`hep-ex/0612052`]

[21] S. Eidelman *et al.* [Particle Data Group], *Phys. Lett.* **B 592** (2004) 1

[22] M. Beneke *et al.*, arXiv:`hep-ph/0003033`, and references therein

[23] M. C. Smith and S. S. Willenbrock, *Phys. Rev. Lett.* **79** (1997) 3825 [arXiv:`hep-ph/9612329`]

[24] CDF Run II top quark group public web page: http://www-cdf.fnal.gov/physics/new/top/top.html, January 2006

[25] http://www-cdf.fnal.gov

[26] http://www.fnal.gov

[27] K. Kodama *et al.* [DONUT Collaboration], *Phys. Lett.* **B 504** (2001) 218 [arXiv:`hep-ex/0012035`]

[28] E. Brubaker *et al.* [Tevatron Electroweak Working Group], arXiv:`hep-ex/0608032`

[29] CDF conference note 8665

[30] D. Acosta *et al.* [CDF Collaboration], *Phys. Rev.* **D 71** (2005) 112001 [Erratum-ibid. D **71** (2005) 119901] [arXiv:`hep-ex/0410076`]

[31] CDF conference note 7835

[32] CDF conference note 7459

[33] CDF conference note 8590

[34] A. Abulencia *et al.* [CDF Run II Collaboration], *Phys. Rev.* **D 74** (2006) 071103 [arXiv:`hep-ex/0512020`]

[35] A. Abulencia *et al.* [CDF II Collaboration], *Phys. Rev. Lett.* **96** (2006) 122001 [arXiv:`hep-ex/0512062`]

[36] A. Abulencia *et al.* [CDF - Run II Collaboration], *Phys. Rev. Lett.* **97** (2006) 062003 [arXiv:`hep-ex/0606027`]

[37] L. Babukhadia *et al.* [CDF and D0 Working Group Members], FERMILAB-PUB-03/320-E (2003)

[38] M. Carena *et al.* [Higgs Working Group Collaboration], arXiv:`hep-ph/0010338`

[39] G. Aubrecht *et al.*, *A Teacher's Guide to the Nuclear Science Wall Chart - Chapter 11*, Contemporary Physics Education Project (2003), http://www.lbl.gov/abc/wallchart/teachersguide/pdf/Chap11.pdf

[40] Fermilab Accelerator Division public web page, http://www-bd.fnal.gov/public/multiturn.html

[41] C. W. Schmidt, FERMILAB-CONF-93-111

[42] E. L. Hubbard, FERMILAB-TM-0405

[43] V. Lebedev, A. Burov, W. Pellico and X. Yang, FERMILAB-CONF-06-205-AD

[44] Fermilab Beam Division, *Run II Handbook*; http://www-bd.fnal.gov/runII/index.html

[45] R. Blair *et al.* [CDF-II Collaboration], FERMILAB-PUB-96/390-E (1996)

[46] A. Sill [CDF Collaboration], *Nucl. Instrum. Meth.* **A447** (2000) 1

[47] A. A. Affolder *et al.* [CDF Collaboration], *Nucl. Instrum. Meth.* **A526** (2004) 249

[48] L. Balka *et al.* [CDF Collaboration], *Nucl. Instrum. Meth.* **A267** (1988) 272

[49] S. Bertolucci *et al.* [CDF Collaboration], *Nucl. Instrum. Meth.* **A267** (1988) 301

[50] M. G. Albrow *et al.* [CDF Collaboration], *Nucl. Instrum. Meth.* **A480** (2002) 524

[51] R. Brun and F. Rademakers, *Nucl. Instrum. Meth.* **A389** (1997) 81, http://root.cern.ch

[52] The SAMGrid project, http://projects.fnal.gov/samgrid

[53] CVS - Concurrent Versions System, http://www.nongnu.org/cvs

[54] J. Antos *et al.* [CDF - Run II Collaboration], IEEE Trans. Nucl. Sci. **53** (2006) 2897 [arXiv:`physics/0606042`]

[55] J. Antos *et al.*, arXiv:`hep-ex/0603008`

[56] Farms and Clustered Systems Group, Fermi National Accelerator Laboratory, "FBSNG – Next Generation of FBS", http://www-isd.fnal.gov/fbsng

[57] Condor project, University of Wisconsin-Madison, http://www.cs.wisc.edu/condor

[58] CDF Central Analysis Farm (CAF), http://cdfcaf.fnal.gov

[59] M. Norman *et al.* "OSG-CAF - A single point of submission for CDF to the Open Science Grid:, in "CHEP 2006, Computing in high energy and nuclear physics", Mumbai, India, 2006; OSG Document 371

[60] J. E. Huth *et al.* in *Proceedings of Research Directions For the Decade: Snowmass 1990*, July 1990, edited by E. L. Berger (World Scientific, Singapore, 1992) p. 134

[61] G. C. Blazey *et al.*, arXiv:`hep-ex/0005012`

[62] A. Bhatti *et al.*, *Nucl. Instrum. Meth.* **A566** (2006) 375 [arXiv:`hep-ex/0510047`]

[63] D. Acosta *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **93** (2004) 142001 [arXiv:`hep-ex/0404036`]

[64] F. Abe *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **80** (1998) 2779 [arXiv:`hep-ex/9802017`]

[65] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, *Comput. Phys. Commun.* **135** (2001) 238 [arXiv:`hep-ph/0010017`]

[66] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *J. High Energy Phys.* **07** (2003) 001 [arXiv:`hep-ph/0206293`]

[67] G. Corcella *et al.*, *J. High Energy Phys.* **01** (2001) 010 [arXiv:`hep-ph/0011363`]

[68] CDF conference note 8375

[69] A. Abulencia *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **96** (2006) 022004 [arXiv:`hep-ex/0510049`]

[70] A. Abulencia *et al.* [CDF - Run II Collaboration], arXiv:`hep-ex/0612061`

[71] A. Abulencia *et al.* [CDF Collaboration], *Phys. Rev.* **D 73** (2006) 092002 [arXiv:`hep-ex/0512009`]

[72] CDF conference note 7102

[73] A. Abulencia *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **96** (2006) 152002 [arXiv:`hep-ex/0512070`]

[74] A. Abulencia *et al.* [CDF Collaboration], *Phys. Rev.* **D 73** (2006) 112006 [arXiv:`hep-ex/0602008`]

[75] CDF conference note 8233

[76] CDF conference note 8420

[77] CDF conference note 8573

[78] E. Accomando *et al.* [ECFA/DESY LC Physics Working Group], *Phys. Rept.* **299** (1998) 1 [arXiv:`hep-ph/9705442`]

[79] I. Borjanovic *et al.*, *Eur. Phys. J.* **C 39S2** (2005) 63 [arXiv:`hep-ex/0403021`]

[80] A. I. Etienvre, PoS **TOP2006** (2006) 023

[81] J. Antos and G. P. Yeh, FERMILAB-CONF-99-260

[82] D. Acosta *et al.* [CDF Collaboration], *Phys. Rev.* **D 71** (2005) 052003 [arXiv:`hep-ex/0410041`]

[83] W. Press *et al.*, Numerical recipes in C: the art of scientific computing, Cambridge University press, 2nd edition, available on-line at *http://www.numerical-recipes.com/nronline_switcher.html*

[84] *Jet energy resolution in dijet events*, talk presented by K. Hatakeyama at the CDF Joint physics group meeting, Fermilab, June 22, 2005

[85] *Min-Bias and the Underlying Event at Tevatron and the LHC*, talk presented by R. Field at the Fermilab ME/MC Tuning Workshop, Fermialb, October 4, 2002; *Toward an Understanding of Hadron Collisions: From Feynman-Field until Now*, talk presented by R. Field at the Fermilab Joint Theoretical Experimental "Wine & Cheese" Seminar, Fermilab, October 4, 2002

[86] J. Antoš, CDF internal note 4165

[87] http://www-cdf.fnal.gov/˜tecchio/internal/topntuple/DIL_plot.html#NEW_PLOT, results are updated in [89]

[88] F. James and M. Roos, *Comput. Phys. Commun.* **10** (1975) 343

[89] CDF conference note 8103

[90] R. Barlow: SLUO lectures on statistics and numerical methods in HEP, Lecture 6: Resampling and the bootstrap. Available online at http://www.hep.man.ac.uk/u/roger, May 2006

[91] A. Abulencia *et al.* [CDF Collaboration], *Phys. Rev.* **D 73** (2006) 032003 [arXiv:`hep-ex/0510048`]

[92] V. N. Gribov and L. N. Lipatov, *Sov. J. Nucl. Phys.* **15** (1972) 438 [Yad. Fiz. **15** (1972) 781], V. N. Gribov and L. N. Lipatov, *Sov. J. Nucl. Phys.* **15** (1972) 675 [Yad. Fiz. **15** (1972) 1218], L. N. Lipatov, *Sov. J. Nucl. Phys.* **20** (1975) 94 [Yad. Fiz. **20** (1974) 181], Y. L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 641 [Zh. Eksp. Teor. Fiz. **73** (1977) 1216], G. Altarelli and G. Parisi, *Nucl. Phys.* **B 126** (1977) 298

[93] S. Frixione, P. Nason and B. R. Webber, *J. High Energy Phys.* **08** (2003) 007 [arXiv:`hep-ph/0305252`]

[94] H. L. Lai *et al.* [CTEQ Collaboration], *Eur. Phys. J.* **C 12** (2000) 375 [arXiv:`hep-ph/9903282`]

[95] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, *Eur. Phys. J.* **C 4** (1998) 463 [arXiv:`hep-ph/9803445`]

[96] CDF internal web page: http://www-cdf.fnal.gov/internal/physics/joint_physics/pdfs/pdf_acceptance.html

[97] F. Abe *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **82** (1999) 271 [Erratum-ibid. **82** (1999) 2808] [arXiv:`hep-ex/9810029`]

[98] Private communication of one of my colleagues (T. Mäki) with Un-Ki Yang