Measurement of the Top Quark Mass in the Dilepton Channel using the Neutrino Weighting Algorithm at CDF II

by

Simon Sabik

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

Abstract

Measurement of the Top Quark Mass in the Dilepton Channel using the Neutrino Weighting Algorithm at CDF II

Simon Sabik

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2006

We measure the top quark mass using approximately 359 pb^{-1} of data from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ GeV at CDF Run II . We select $t\bar{t}$ candidates that are consistent with two W bosons decaying to a charged lepton and a neutrino following $t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow l^+l^-\nu\bar{\nu}b\bar{b}$. Only one of the two charged leptons is required to be identified as an electron or a muon candidate, while the other is simply a well measured track. We use a neutrino weighting algorithm which weighs each possibility of neutrino direction to reconstruct a top quark mass in each event. We compare the resulting distribution to Monte Carlo templates to obtain a top quark mass of $170.8^{+6.9}_{-6.5}\,(stat) \pm 4.6\,(syst)$ GeV/c^2.

Acknowledgements

I would first like to thank my advisor Professor Pierre Savard for his sound direction, for sharing his expertise of physics and for teaching me how to become a better scientist. He was an excellent advisor, and it was a pleasure to do research with him.

I thank David Ambrose, Kostas Kordas and Kevin Sung, who took part in this analysis along with Pierre and myself. I appreciate their efforts and I am very proud of what we accomplished together.

I am also grateful to Professor Bailey, Professor Holdom, Professor Orr, Professor Sinervo and Professor Trischuk of the University of Toronto. Being surrounded by Professors of such a high calibre was inspiring and I always benefited learning from them.

I had a great time working with the students and postdocs of the CDF Toronto group. In particular, Ich möchte mich bei meinen Freunden Bernd, Oliver und Teresa bedanken. تشکرکت Shabnaz خواهم ازدوستم می . Quisiera agradecerle a mi amigo Ian. 我想感謝我的朋友 Stan. Reda و Rachid اربد ان اشکر زملائ. Je voudrais remercier mes amis Hugo, Michael et Dan. I would also like to thank my friend Jean-Francois.

I give my thanks to all my Fermilab collegues for their ideas and advice. Among them are Anwar Bhatti, Florencia Canelli, Douglas Glenzinski, Eva Halkiadakis, Chris Hays, Beate Heinemann, Matt Herndon, Young-Kee Kim, Vaia Papadimitriou, Evelyn Thomson and Igor Volobouev.

Finalement, je tiens à remercier ma mère, Ruwani, Alexandre, Nadia et tous mes amis pour leur soutien.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

As of today, the Standard Model is the best tool we have to predict the outcome of experiments in the subatomic world. The top quark, whose existence was a prediction of the Standard Model, was discovered by CDF [1] and D0 [2] in 1995, using limited statistics. Our current experimental knowledge of it remains quite limited. Its unexpectedly large mass, $\sim$35 times more than that of the next heaviest quark (the b quark), suggests that the top quark might be a special particle. Unlike other quarks, it is massive enough to decay to a real W boson. Its predicted lifetime ($\sim 10^{-25}s$) is expected to be too short for top flavored hadrons to form before decay [3].

The knowledge of the free parameters of the Standard Model, such as the fermion masses, improves the accuracy of the Standard Model predictions. Why is it more interesting to measure the top quark mass than any other Standard Model parameter? The reason is related to another prediction of the Standard Model: the existence of a Higgs boson, which would be responsible for giving mass to every elementary particle. The top quark mass, precisely because it is so large, is related to the prediction of the Higgs boson mass. The understanding of the origin of mass would be a major scientific breakthrough, and the top quark mass plays an important role in this prediction.

Interestingly, the top quark mass is close to the scale of electroweak symmetry break-

ing. This observation led to the development of several alternative models beyond the Standard Model, in which the top quark mass plays a significant role. A more precise measurement will contribute to the experimental testing of these models.

This thesis presents a measurement of the top quark mass using data collected during Run II of the Collider Detector at Fermilab (CDF) experiment. We study the dilepton decay channel (see Section 1.2) using a quantity of data corresponding to an integrated luminosity of 359 pb^{-1} collected during CDF Run II. The event selections are based on the identification of a charged lepton and the observation of a good quality isolated track (see Section 4.2). The top reconstruction method used is the neutrino weighting algorithm (see Section 5.2). In the current Chapter, we introduce the subject of the top quark mass by discussing the motivation for this measurement and previous measurements. In Chapter 2, we will describe the experimental apparatus. In Chapter 3, we will describe the software used to reconstruct particles from raw measurements. In Chapter 4, we will describe the event selections. In Chapter 5, we will explain the analysis methodology used to calculate the top quark mass. Finally, in Chapter 6, we will discuss the systematic uncertainties on the measurement of the top quark mass.

1.1 Motivation for measuring the Top Quark Mass

Even though the Standard Model does not directly predict the quark masses, a precise measurement of the top quark mass combined with other experimental information can constrain important parameters of the Standard Model. For instance, we can constrain the Higgs boson mass through the radiative corrections to the W propagator [4]. In the Standard Model context, these corrections are fermion loops and the emission and absorption of a Higgs boson. The light quark loops are highly suppressed, and, as a result, the $t\bar{b}$ and the Higgs loop (see Figure 1.1) dominate the corrections. Therefore, accurate measurements of the top quark mass and the W boson mass are needed to

predict a Higgs boson mass.

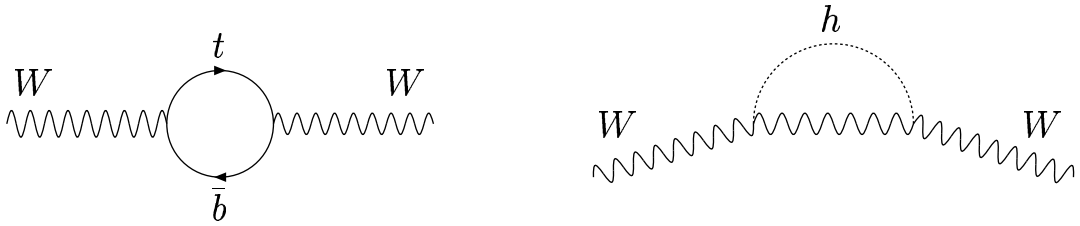

Figure 1.1: Radiative corrections to the W boson propagator. The dominating fermion loop $t\bar{b}$ is shown on the left and the emission and absorption of a Higgs boson is shown on the right.

Figure 1.2 shows how the top quark mass and the W boson mass results from the Tevatron Run I and the LEP2 experiment have constrained the Higgs boson mass. If the Higgs boson is discovered and its mass measured at the LHC, a precise measurement of the top quark mass will allow a test of the self-consistency of the Standard Model. Currently, the uncertainty on the top quark mass limits our prediction on the Higgs mass. For this reason, it is crucial to make a more accurate measurement.

1.2 The dilepton channel

At the Fermilab Tevatron, top quarks are mostly created through quark anti-quark annihilation and the subsequent production of $t\bar{t}$ pairs. In the Standard Model, the weak $t \to Wb$ branching ratio is about 99.8%. The W boson in turn can then decay as a charged lepton-neutrino pair or as a quark anti-quark pair. We call the "dilepton channel" the cases where both W bosons decay into charged leptons with their associated neutrinos. Its signature is therefore two charged leptons (except when we have a τ lepton, where we observe its decay products), two jets, and a missing transverse energy (energy imbalance from the two undetected neutrinos, see Section 3.2). Furthermore, in addition to the expected two jets from the hadronization of the b and the $\bar{b}$, gluons can

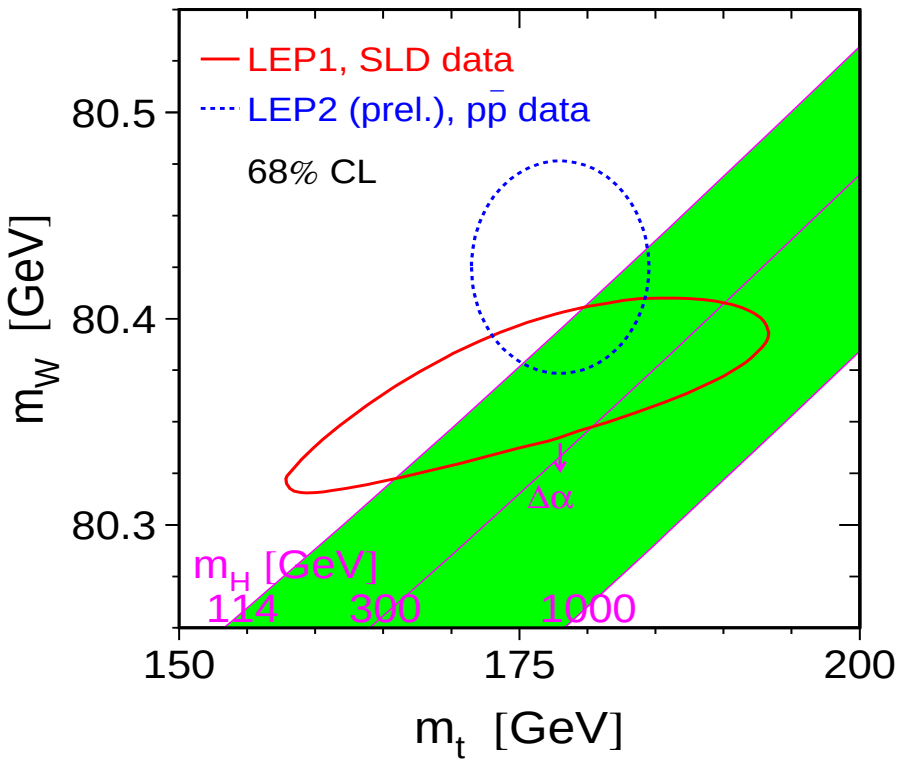

Figure 1.2: Constraints on the Higgs boson mass from the top quark and W boson mass measurements in CDF, D0 Run I and the LEP2 experiment (dashed blue line). Constraints from studies at the Z boson pole (red line) and from a direct search at LEP (green bands) are also shown. For latest Run II results, see Figure 7.2 in conclusion.

be radiated in the initial state or the final state.

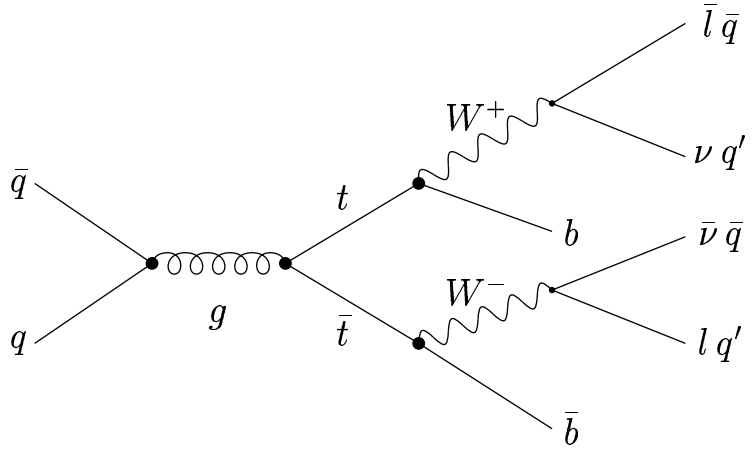

Figure 1.3: Main $t\bar{t}$ production and decay process at the Tevatron

Cases where, from a $t\bar{t}$ pair, one of the two W bosons decays as an electron or a muon and a neutrino and the other W boson decays as a quark anti-quark pair are labeled as the "lepton+jets" channel. Relative to the lepton+jets channel, the dilepton channel has a lower branching ratio (5% of all $t\bar{t}$ decays, excluding τ lepton cases compared to 31% for the lepton+jets channel), but it has the advantage that it is less contaminated by background processes.

1.3 Previous measurements

The combined results from CDF and D0 in Run I give a 178.0 ± 4.3 GeV/c^2 top quark mass (see Figure 1.4). The dilepton result was $167.4 \pm 10.3(stat) \pm 4.8(syst)$ GeV/c^2 for CDF [5]. The lepton+jets channel provides the highest precision measurement, but the dilepton channel provides an important complement. Indeed, the dilepton result contributes to a better combined measurement, and provides a statistically independent cross check.

In the CDF Run II, we have a better detector acceptance and a larger quantity of data. The statistical uncertainty will therefore be greatly reduced, and important gains

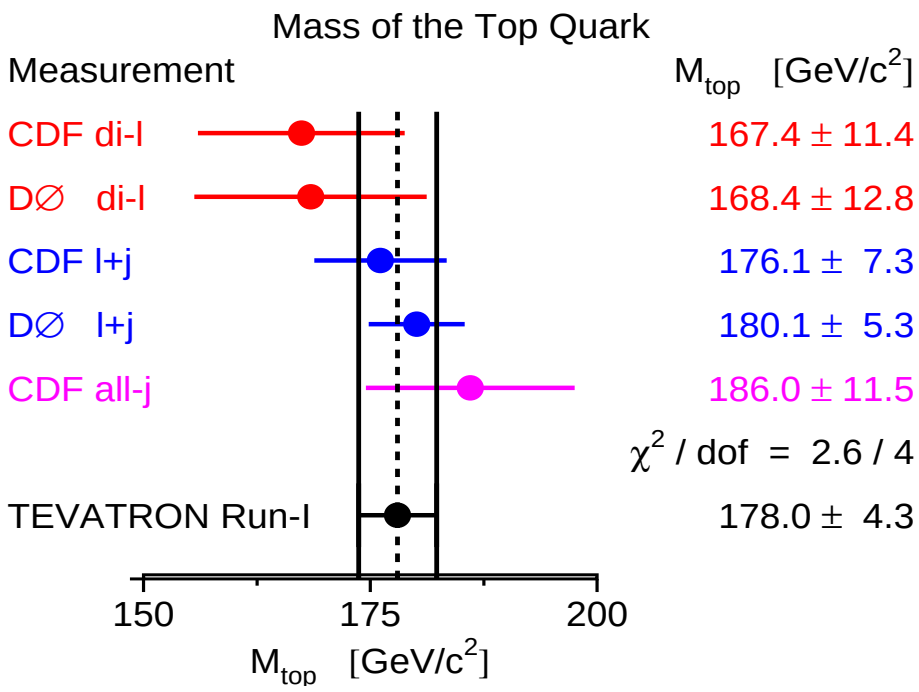

Figure 1.4: CDF and D0 Run I measurements of the top quark mass. For latest Run II results, see Figure 7.1 in conclusion.

can be made by reducing the systematic uncertainty as well. This can be achieved by a better understanding and tuning of the detector simulation.

Chapter 2

The CDF II experiment

In this chapter, we describe the apparatus used in this measurement in two parts:

- The production and collision of protons and anti-protons by the accelerator chain at Fermilab
- The CDF detector (described in details in [6]), with an emphasis on the aspects most relevant to this analysis

2.1 The accelerator

In a series of steps, the Fermilab accelerator chain creates protons and anti-protons, and accelerates them to provide $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV (see Figure 2.1).

2.1.1 Creating 8 GeV protons

The beam creation starts with a negative hydrogen ion gas. First, the Cockcroft-Walton pre-accelerator uses a DC electrical field to accelerate the H^- ions to 750 KeV. Then, a linear accelerator, using an AC electrical field, brings the H^- ions to 400 MeV. The H^- ions are stripped of their two electrons to become bare protons. The protons then enter

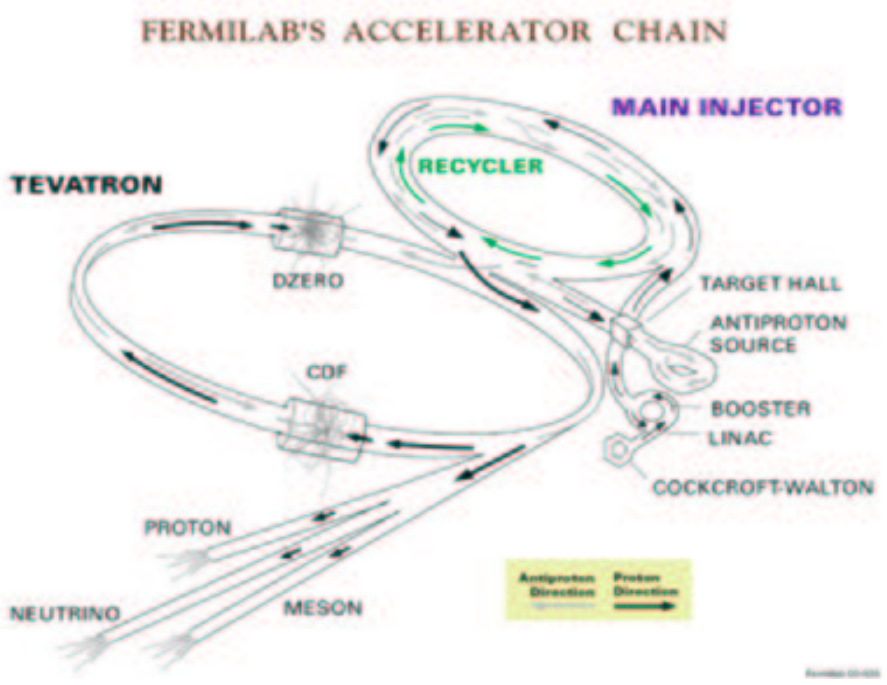

Figure 2.1: The Fermilab accelerator complex

the Booster synchrotron accelerator, which brings them to 8 GeV using radio frequency cavities.

2.1.2 The Main Injector

The protons are led from the Booster to another synchrotron, the Main Injector, where they can be accelerated up to 150 GeV. To create anti-protons, the Main Injector leads 120 GeV protons to a nickel target. Several types of particles are produced from this interaction, including anti-protons which are selected using a bending magnetic field. The anti-protons are transferred to the Debuncher, a synchrotron where they are cooled to 8 GeV to create a more uniform beam. They are then stored in the Accumulator, a storage ring. When enough anti-protons have been stored, they are sent back to the Main Injector and accelerated to 150 GeV.

2.1.3 The Tevatron

The protons and anti-protons are then passed to the Tevatron, where they are accelerated using an AC electric field and kept on a circular trajectory using superconducting dipole magnets. The Tevatron holds 36 bunches of protons and as many anti-protons. Protons and anti-protons travel in opposite directions. A series of quadrupole magnets focus the beam at the interaction points, where a bunch crossing occurs every 396 ns. The center of mass energy of the collisions is 1.96 TeV.

2.2 The CDF Detector

We reconstruct events in the dilepton channel $t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}l^+l^-\nu\bar{\nu}$. The visible signature is two jets, two charged leptons (in the case of a τ lepton, only its decay products can be observed) and a large missing transverse energy, due to the two undetected neutrinos (see Section 3.2). In order to measure the top quark mass, the CDF detector must precisely measure the kinematic quantities and provide the identity of these particles. The selections are described in Chapter 4.

The coordinate system is defined as follow. x points horizontally outward from the accelerator, y points straight up and z points in the proton direction. ϕ and r are in the x-y plane. θ is the angle from the proton direction and $\eta \equiv -ln(tan(\theta/2))$ is the pseudo-rapidity. Transverse quantities are projections in the x-y plane. For instance, $E_t \equiv E\sin(\theta)$ denotes the transverse energy and $p_t \equiv p\sin(\theta)$ denotes the transverse momentum. A slice of the CDF detector is seen in Figure 2.2.

2.2.1 The tracking system

The CDF tracking system is located nearest the interaction point and comprises (from outside to inside), the Central Outer Tracker (COT), the Intermediate Silicon Layers (ISL), and the Silicon Vertex detector (SVX) (see Figure 2.3). The tracking system lies

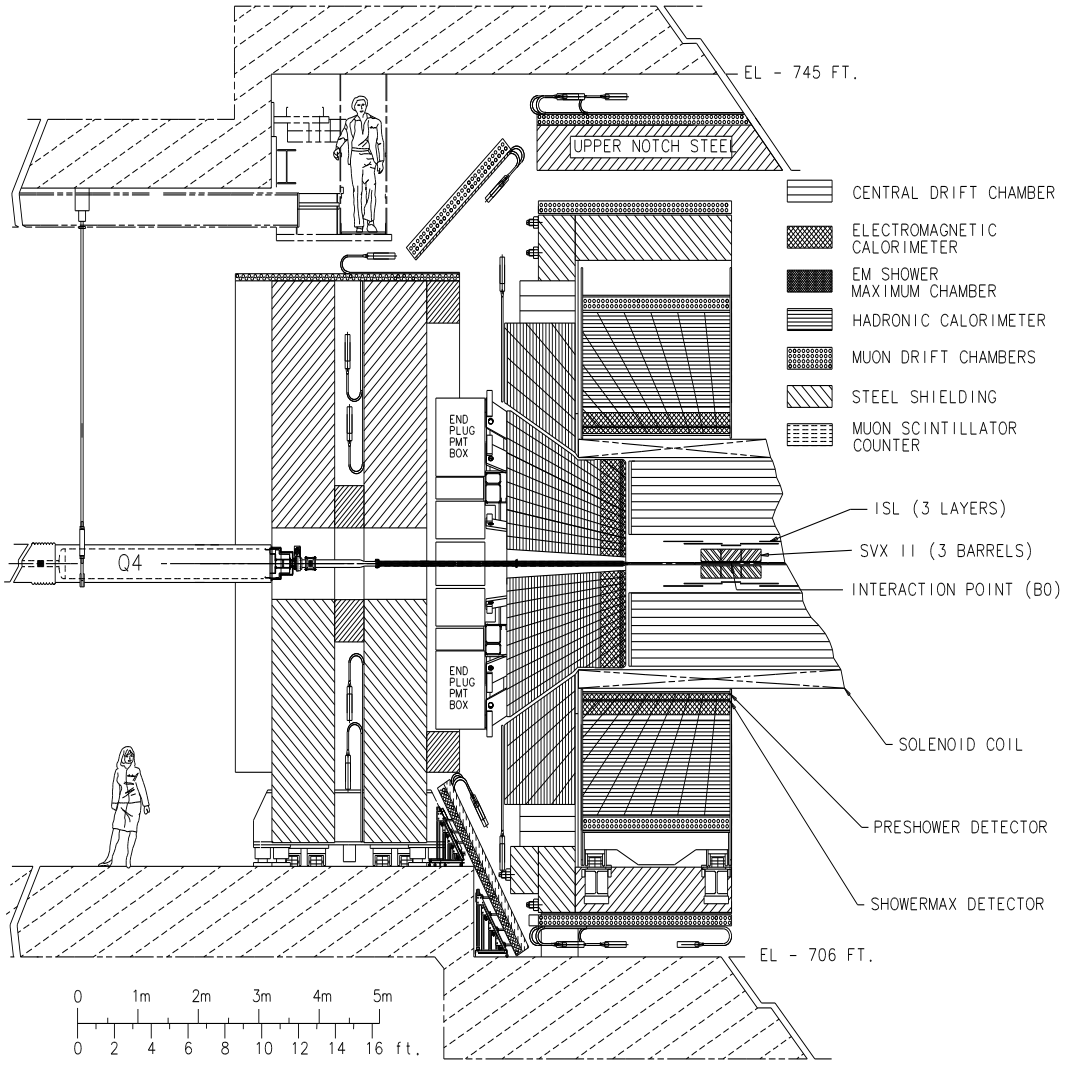

Figure 2.2: Schematic view of half the CDF II detector

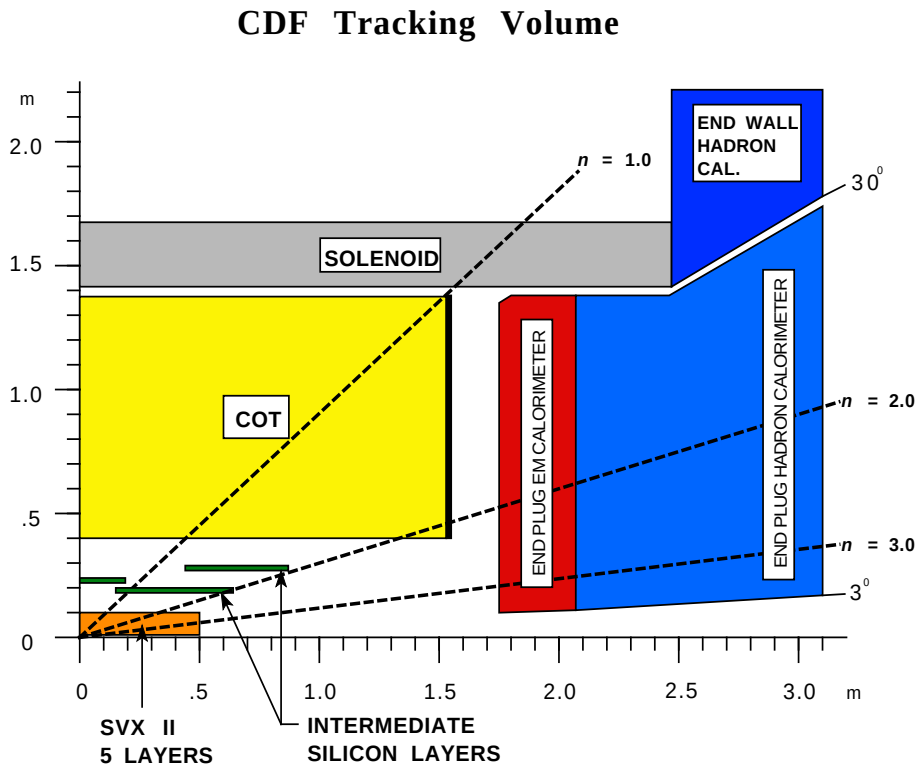

Figure 2.3: Schematic view of a quadrant of the CDF tracking system

inside a superconducting solenoid magnet, with length of 2.8 m and radius of 1.5 m. The coil produces a 1.4 T magnetic field parallel to the beam, curving the charged particle trajectories. The main purpose of the tracking system is to precisely measure the trajectory and momentum of charged particles. It also plays a role in particle identification (see Section 4.2.1). Tracking is essential to the reconstruction of the $t\bar{t}$ events signature in the dilepton channel, as it provides information about the two charged leptons. It is also useful for the study of charged particles in jets, as we will see in Section 6.1.3.

COT

The COT is the main tracking device of the CDF detector. It is a cylindrical drift chamber consisting of 96 concentric layers subdivided in 8 superlayers of 12 layers each. The COT is 310 cm long with an inner radius of 44 cm and an outer radius of 132 cm. It is centred at $z = 0$ and its angular coverage is $|\eta| \leq 1$.

In the COT cells, an electric field is established using longitudinal sense wires (anodes) and field planes (cathodes) strung between two end-plates. The chamber is filled with an Argon-Ethane 50% : 50% gas mixture.

When charged particles pass through the gas, it creates electron-ion pairs. The electric field causes the free electrons to drift toward the sense wires. By this motion, the electrons can gain enough energy to cause secondary ionizations. The resulting electron avalanche is detected by the sense wire as a pulse. The hit position is determined using the known drift speed and the measured drift time. Curved by the magnetic field, the charged particle leaves a helical track of hits.

The drift speed is approximately 50 μm/ns. Given that the maximum drift distance is 0.88 cm, the maximum drift time is below 200 ns. Since a bunch crossings occur every 396 ns, bunch to bunch ambiguity is avoided.

Even numbered superlayers (superlayer 1 being at the inner radius) are axial to the beam line for measurement of the hit position in the $r-\phi$ plane. Odd numbered superlayers are tilted by an alternating stereo angle of $\pm 2°$ for measurement of the hit position in the $r-z$ plane. The hit resolution is approximately 140 μm. The momentum resolution of the COT is $\frac{\delta p_t}{p_t} \simeq 1.5 \times 10^{-3}$ GeV$^{-1} \times p_t$. The tracking algorithm is described in Section 3.1.1.

Silicon detectors

A system of silicon micro-strip detectors is located inside the COT volume. When a charged particle passes through the semiconductor, it creates electron-hole pairs, analogous to the electron-ion pairs in the COT. The electric field causes the electron and holes to separate and be detected at the electrodes, giving a signal proportional to the energy loss of the incident charged particle.

The SVX consists of 5 concentric double sided layers. Each layer is separated in 12 wedges in ϕ. It is 89 cm long with the innermost layer at radius of 2.44 cm and the

outermost layer at radius of 10.6 cm. It is centered at $z = 0$ and its angular coverage is $|\eta| \leq 2$. For each layer, one side of the double sided chip is for measurement of the hit position in the $r - \phi$ plane. On the other side, 3 of the layers have 90° stereo alignment and the remaining 2 have a 1.2° stereo alignment for measurement of the hit position in the $r - z$.

The ISL are additional silicon layers located around the SVX. One layer covers up to $|\eta| \leq 1$ to provide additional tracking information. Two layers are extended from $1 \leq |\eta| \leq 2$ to provide information for the tracks beyond the COT acceptance.

Being located closest to the interaction point, the silicon detector system's main goal is to measure the position of the decay vertices. The combined results from the inner tracking system yields an impact parameter d_0 * resolution of 40 μm and a z_0 † resolution of 70 μm.

2.2.2 Calorimetry

The CDF calorimeter system is located around the tracking system (see Figure 2.2). The inner part is designed for electromagnetic energy measurements and comprises the Central ElectroMagnetic calorimeter (CEM) and the Plug ‡ ElectroMagnetic calorimeter (PEM). The outer part is designed for hadronic energy measurements and comprises the Central HAdronic calorimeter (CHA), the Plug HAdronic calorimeter (PHA), and the Wall HAdronic calorimeter (WHA), which fills a geometric gap between the CHA and PHA.

The CDF calorimeter system plays an important role in the measurement of the top quark mass, as it measures the energy of the jets and helps particle identification.

*Distance of closest approach of the track to the primary vertex in the x-y plane
†z position of the point of closest approach
‡"Central" designates the region with $|\eta| \leq 1$ and "Plug" designates the forward region $1 < |\eta| < 3.6$

Electromagnetic Calorimeter

The electromagnetic calorimeter consists of 23 (22 in the plug) alternating layers of lead absorber (4.5 mm) and scintillator (4 mm). It is separated into towers pointing toward the origin.

The electromagnetic calorimeter's main goal is to precisely measure the energy of electrons and photons. When an incoming electron crosses the absorber, it radiates photons (mostly through the Bremsstrahlung process). As they interact with the absorber, photons convert to electron-positron pairs. The resulting shower is detected in scintillators and converted to a light pulse proportional to the energy of the particles. Wavelength shifting fibers conduct the light from the scintillators to photomultiplier tubes.

Embedded at the point where electromagnetic showers are maximally developed is a multi-wire proportional chamber (CES) in the central region and a scintillator strip detector (PES) in the plug. The CES and PES precisely measures the $\phi - \eta$ position of the shower. They also give information about the shower profile, which is used to identify electrons and photons.

The energy resolution of the electromagnetic calorimeter is given by $\frac{\sigma_E}{E} = \frac{13.5\%}{\sqrt{Esin(\theta)}}$ (where E is in GeV) for the central component and $\frac{\sigma_E}{E} = \frac{14.4\%}{\sqrt{E}} \oplus 0.7\%$ (where E is in GeV) for the plug component.

Hadronic Calorimeter

The hadronic calorimeter consists of 23 (22 in the plug) alternating layers of an iron absorber (5 cm) and scintillator (6 mm). It is separated into towers pointing toward the origin.

Its main goal is to precisely measure the energy of hadrons. As the incoming hadron crosses the absorber, it collides with a nucleus, creating secondary hadrons. Electromagnetic processes also contribute to the shower. The energy is measured with the same system of scintillator, wavelength shifting fibers, and photomultiplier tubes as the

electromagnetic calorimeter.

The energy resolution of the hadronic calorimeter is given by $\frac{\sigma_E}{E} = \frac{75\%}{\sqrt{E}}$ (where E is in GeV) for the central component and $\frac{\sigma_E}{E} = \frac{74\%}{\sqrt{E}} \oplus 3.8\%$ (where E is in GeV) for the plug component.

2.2.3 The Muon identification system

Muons are the only charged particles that interact little with the calorimeter. Muon identification is based on the fact that any charged particle observed beyond the calorimeter is most likely a muon. Around the calorimeter, and limited to the $|\eta| < 0.6$ region, is the Central MUon detector (CMU), which consists of 4 layers of single-cell drift chambers. A 60 cm thick steel absorber surrounds the CMU to further prevent non-muon particles from reaching the Central Muon uPgrade (CMP). The CMP is a second set of 4 drift chambers covering the $|\eta| < 0.6$ region. The Central Muon eXtension (CMX) also consists of single-cell drift chambers. They are arranged in the shape of a frustum concentric around the beam in the $0.6 < |\eta| < 1.0$ region.

2.2.4 The trigger system

Bunch crossings occur at a rate of 2.5 MHz (396 ns). The writing of data on tape occurs at a rate of 75 Hz. Therefore, a fast and efficient trigger system is needed to quickly discard uninteresting events. For this analysis, the trigger decision is based on the identification of a high p_t charged lepton. The CDF trigger system consists of 3 levels. The purpose for each level is to reduce the rate enough to avoid dead time in the next level.

Level 1 is a hardware trigger operating at the bunch crossing rate and with an acceptance rate of 20 kHz. It quickly collects primary information from the detectors. The eXtra Fast Trigger (XFT) collects hit information from the 4 axial superlayers of the COT and attempts to reconstruct $r - \phi$ tracks and to measure their momentum. The XFT information is relayed to the calorimeter and muon trigger. Energy measurements

are made in electromagnetic and hadronic trigger towers [1]. The central electron trigger requires a track that points toward an electromagnetic cluster. The central muon trigger requires a track pointing toward a set of hits in the muon chambers. In the case of the plug electron trigger, only the calorimeter information is used at level 1.

Level 2 is a combined hardware and software trigger with acceptance rate of 350 Hz. The electron trigger merges the calorimeter trigger towers into clusters and additional cuts can be imposed. The level 2 trigger imposes no additional requirement for muon triggering.

Level 3 is a purely software trigger with acceptance rate of 75 Hz. Electrons and muons can be fully reconstructed using the offline software (see Section 4.2).

The trigger selections for our analysis are described in Section 4.1.

2.3 Monte Carlo simulations

Monte Carlo simulations are a crucial part of the top quark mass measurement. They are used to model the signal and background processes, which are compared with data to extract the top quark mass (see Chapter 5).

Monte Carlo production is done in three stages: event generation, detector simulation, and particle reconstruction.

The event generation randomly creates the production, decay process chain, and kinematics of the particles. The generation of $t\bar{t}$ events is made using the HERWIG software [7]. The generation of the background processes is made using the PYTHIA [8], ALPGEN [9], and HERWIG softwares.

The detector simulation reproduces the interaction of particles with the detector. The simulation must account for all detector efficiencies so that the Monte Carlo be as similar to the data as possible. Most detector simulations are made using the GEANT

[1] Collection of two towers defined by an η-ϕ region

[10] software package. The calorimeter showering simulation is made using the GFLASH [11] software.

The particle reconstruction software is the same in data and Monte Carlo, and its features most relevant to this analysis are described in Chapter 3.

Chapter 3

Particle reconstruction

In this chapter, we discuss how the raw data collected by the CDF detector described in the previous chapter is transformed into the quantities useful to reconstruct $t\bar{t}$ dilepton events.

3.1 Charged lepton reconstruction

The dilepton signature requires two charged leptons (electrons, muons, or its decay products in the case of τ leptons). In this section, we discuss how muons and electrons are reconstructed. The selections of charged leptons are detailed in Chapter 4.

3.1.1 Central tracking

We reconstruct tracks from hits left by charged particles in the tracking system.

Central tracks are first reconstructed in the COT, and then extrapolated to the inner tracking system. The first step is to find $r - \phi$ segments in axial superlayers. The segments are linked between each of the four axial superlayers to form a full $r - \phi$ track. Independently, segments are formed in the stereo superlayers. Stereo superlayer segments are matched to the axial segments to obtain a complete 3 dimensional COT

track. The COT tracks are then projected to the inner tracking system to find hits that are consistent with the extrapolated trajectory. From the curvature of the tracks, caused by the magnetic field, we can extract the momentum of the track.

3.1.2 Central electron reconstruction

Electrons deposit most of their energy in the electromagnetic part of the calorimeter. Clusters are formed from two or three CEM towers. We have an electron candidate when an EM cluster matches an extrapolated COT track trajectory. After identification, we extract the electron energy from its electromagnetic cluster. The direction of the electron is taken from its track.

3.1.3 Muon reconstruction

For this analysis, we only use central region muons ($|\eta| \leq 1$). In the muon drift chambers, we reconstruct stubs, which are sets of 3 or 4 hits. To obtain muon candidates, stubs are extrapolated to be matched with COT tracks. After identification, we extract the muon's momentum from its track only.

3.1.4 Plug Electron reconstruction

Tracks in the $|\eta| > 1$ region don't pass through all 8 superlayers of the COT . For this reason, we start with an electromagnetic cluster in the PEM. The energy of the cluster gives us information about the magnitude of the curvature of the particle. A precise measurement of the cluster position is made in the PES. Supposing that the electron originates from the 3-D primary vertex, we have all the information needed to define a helix, but with a sign ambiguity. For each charge possibility, the helix is extrapolated to the silicon detectors for matching with silicon hits. If no match is found, we have no plug electron candidate. If one helix is matched, we have a plug electron candidate. If both

helices are matched, the one with the best χ^2 fit is kept as a plug electron candidate.

3.2 Missing transverse energy

Since the momentum of the $p\bar{p}$ system is in the z direction, the sum of its collision products should have no perpendicular momentum component. However, when this is not the case, it means that the detector did not accurately measure some of the energy. It can be due to a particle passing outside the detector acceptance, detector inefficiencies or inaccuracies, or the presence of neutrinos, which don't interact with the detector. The missing transverse energy is measured from the energy and η position of each calorimeter tower. It is defined as $\not\!\!E_T$= - $\sum_i^{n\,towers} \vec{E}_i \sin\theta_i$. It measures the energy that was missed by the detector, given a balanced collision. In the dilepton channel case, the missing transverse energy ideally gives us information about the energy of the two neutrinos which were not detected.

3.3 Jet reconstruction and energy scaling

Strongly interacting particles produce jets of hadrons and other particles through fragmentation. The calorimeter system measures the energy of the jets.

3.3.1 Jet reconstruction

Energy is measured in the calorimeter towers. The jet reconstruction algorithm JETCLU [12] starts by selecting "seed towers" defined as calorimeter towers with $E_t > 1$ GeV. The sum of the electromagnetic and hadronic components of that tower is used to calculate its energy. A cluster is defined as the group of towers within a cone of $\Delta R \equiv \sqrt{(\eta^{centre} - \eta^{tower})^2 + (\phi^{centre} - \phi^{tower})^2} < 0.4$ around the geometrical centre of the tower. The jet transverse energy is the sum of the transverse energy in all the towers

in the cluster. The centroid of the cluster represents the direction in η and ϕ of the cluster. It is calculated as the weighted average of the energy over all towers in the cluster in ϕ and η. From this centroid is created a new cluster and this process is iterated until the centroid is stable.

3.3.2 The jet energy corrections

Due to several effects, the measured jet energy in the calorimeter doesn't reflect the true energy of the jets. In this section, we describe each of these effects, and the correction applied to obtain the true energy [13]. Apart from the relative correction, each correction is applied to the data and Monte Carlo separately. The systematic difference between the simulation of each of these effects will be discussed in Section 6.1.

η relative correction

The calorimeter response is not uniform in η. This is mainly due to the gaps between sections of the central calorimeter at $\eta = 0$ and where the central and plug calorimeters join at $\eta \simeq 1.1$. The relative correction aims at calibrating the jet energy to that of the central region, but apart from the central gap, in a region defined by: $0.2 < |\eta| < 0.6$. We select events where we have two jets back-to-back in ϕ, at least one of the two being in the $0.2 < |\eta| < 0.6$ region ("trigger jet"). When both jets are in the central region, the trigger jet is selected randomly. The second jet ("probe jet") is studied as a function of η. The correction factor is the ratio of the P_t of the probe jet and the trigger jet.

This process is repeated separately for data and Monte Carlo. Figure 3.1 shows the correction as a function of η.

Multiple interactions correction

With the high luminosity at the Tevatron, it often occurs that we have more than one $p\bar{p}$ interaction in a bunch crossing. The particles produced in these secondary interactions

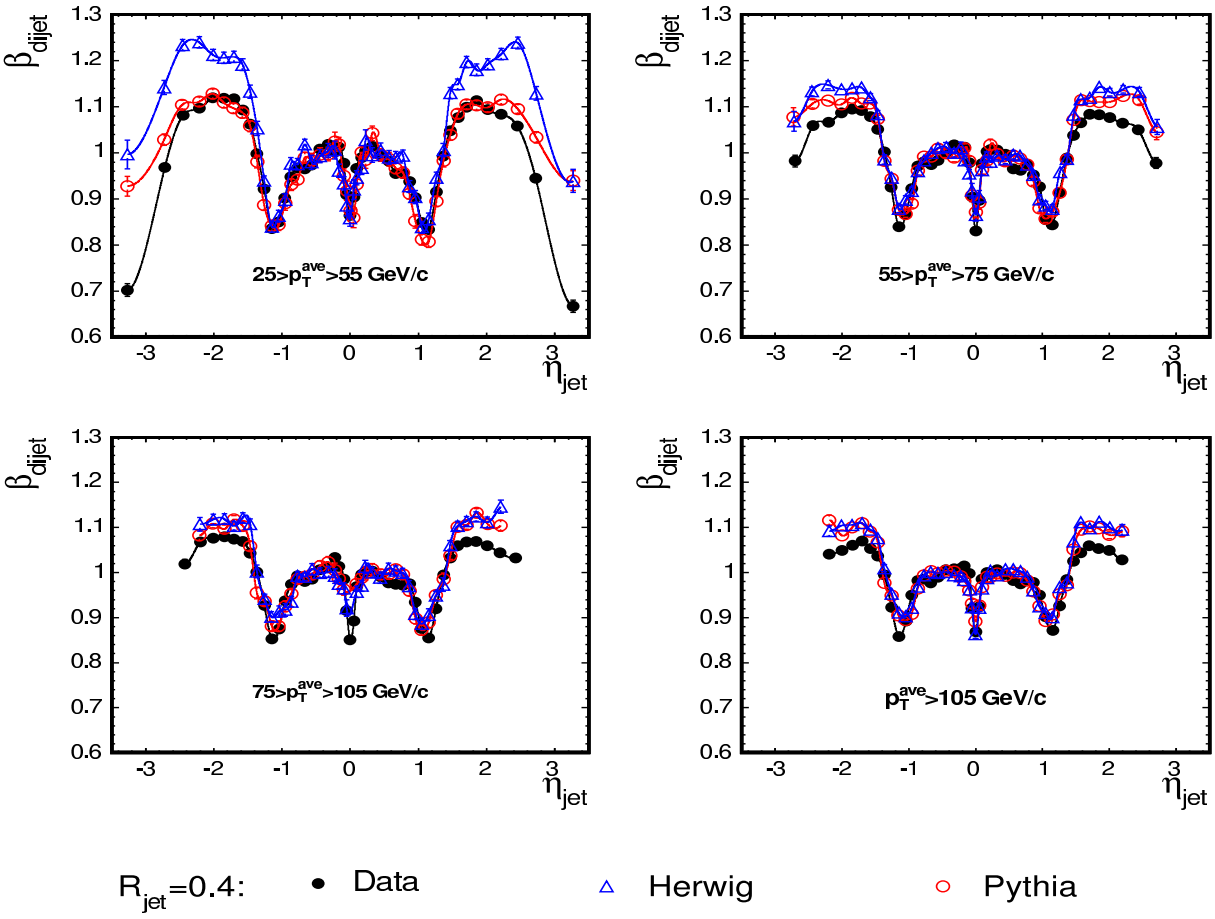

Figure 3.1: The relative jet energy correction in Monte Carlo and data for several average jet P_t ranges. The $\beta_{jet} \equiv P_t^{probe}/P_t^{trigger}$ factor is used to correct the jet energy response to that of the central region. The lines show the interpolation between the individual measurements used to correct the jet energy.

can fall in our jet reconstruction cone and contaminate our energy measurement.

The average transverse energy that falls in the jet reconstruction cone, as a function of the number of vertices, is calculated using minimum bias data events. A reconstruction cone is randomly selected centered in the central region $0.2 < |\eta| < 0.6$. Figure 3.2 shows the average transverse energy in the cone as a function of the number of vertices. The multiple interactions correction consists of subtracting this energy from the jet energy in each of the reconstructed jets.

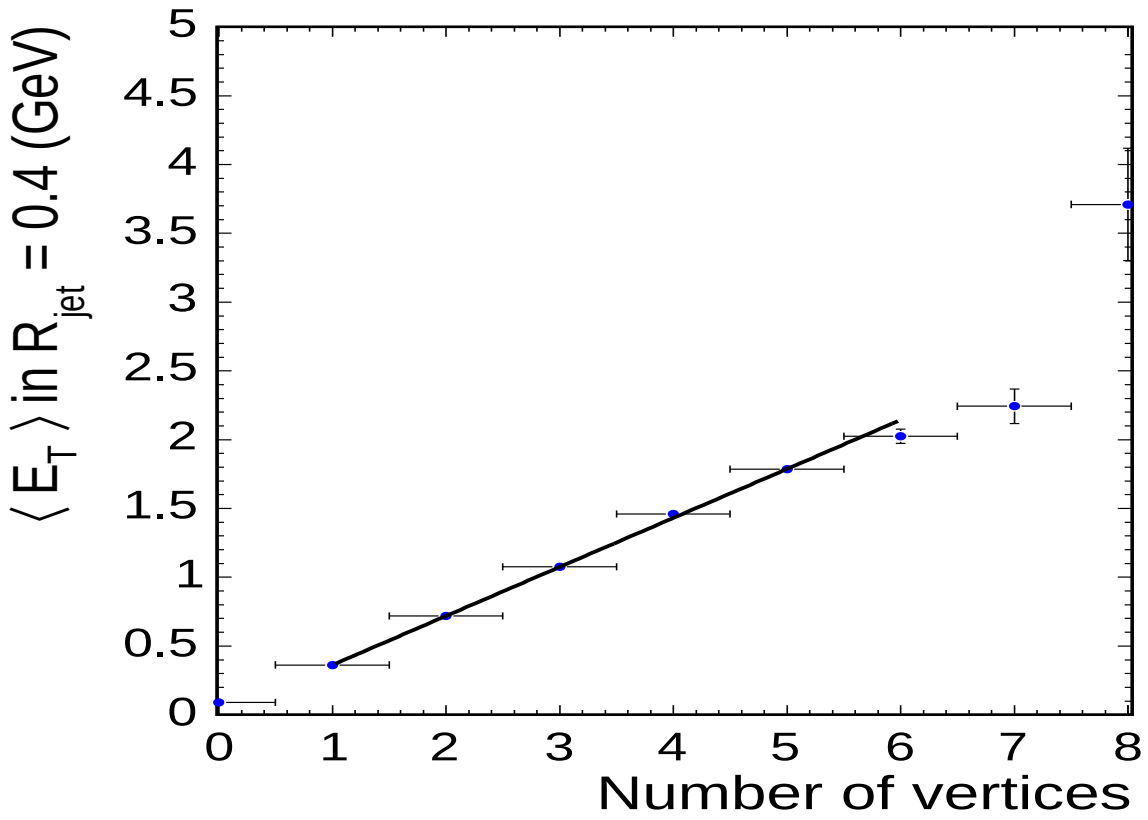

Figure 3.2: Average transverse energy that falls in a cone of $R_{jet} = 0.4$ as a function of the number of vertices in minimum bias data. A linear fit is shown. Jet energies are corrected by substracting this value as a function of the number of vertices.

Absolute jet energy scale correction

When a hadron interacts with the calorimeter, the nuclear collision is inelastic, and some energy is lost. This occurs at every step of the ensuing hadronic shower. As a result, the energy in the calorimeter is lower than the true energy of the incoming hadron. For single particles, this discrepancy is not linear with the energy. For this reason, the fragmentation of jets to multiple particles will affect the calorimeter response to jets.

The correction factor is calculated using Monte Carlo PYTHIA simulations, by comparing the sum of the transverse energy of the generated particles that fall in the jet cone p_T^{HAD} with the simulated calorimeter measurement p_T^{JET}. The correction factor, $C_{ABS} = p_T^{HAD}/p_T^{JET}$, is parameterized as a function of p_T^{JET}. This way, we obtain, on average, the true energy in the jet reconstruction cone.

Figure 3.3 shows the absolute jet energy correction factor as a function of the measured jet p_t.

Underlying event correction

Within a $p\bar{p}$ collision where a hard interaction of interest occurs, spectator quarks can interact with each other. The products of these interactions can end up in the jet reconstruction cone, increasing the measured jet energy artificially. To measure this effect, we use minimum bias data, in the same way as the multiple interactions correction, but for cases where we detect exactly one decay vertex. The underlying event correction is measured to be 0.6 GeV (this correction is applied after the absolute jet energy correction, hence the discrepancy with Figure 3.2).

Out-of-cone correction

A fraction of the jet energy can be moved away transversely from the jet center and fall outside the jet cone. Particles can exit the cone in the case of final state radiation

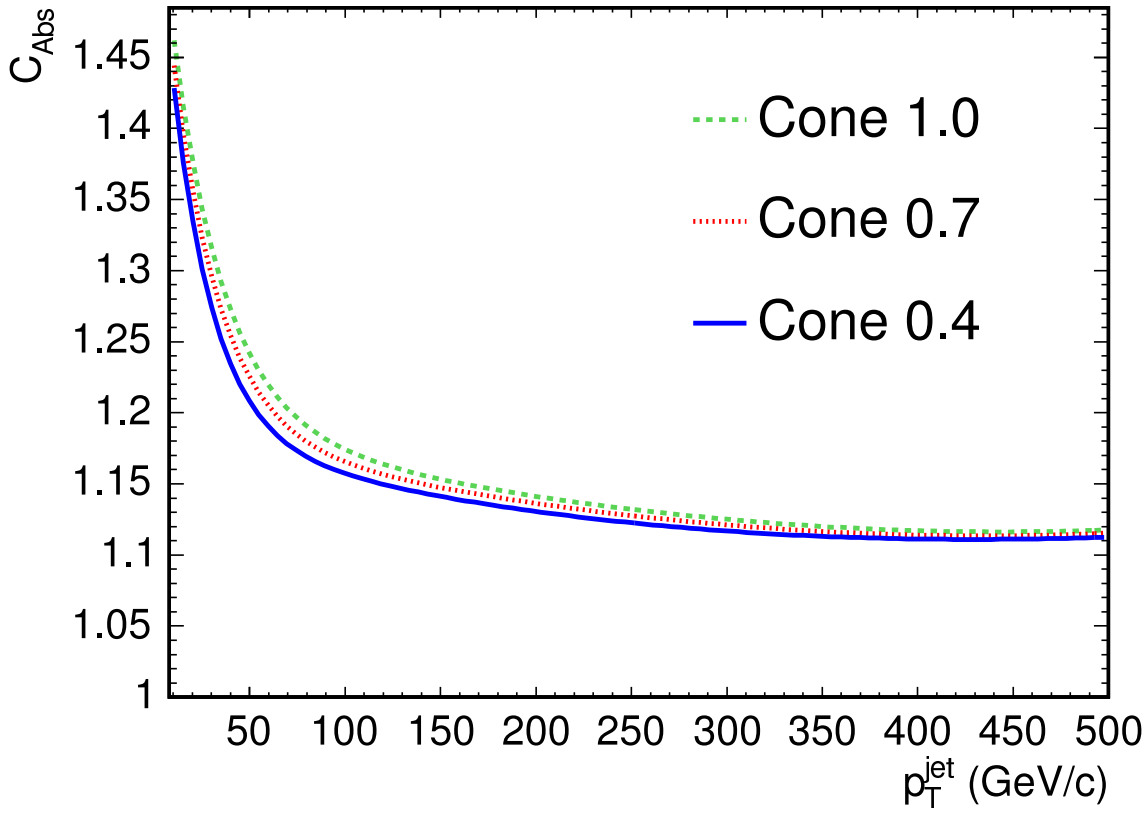

Figure 3.3: The absolute jet energy scale C_{Abs} as a function of jet P_t. This factor is used to correct the measured jet energy that falls in the indicated cone size to that of the original incoming particles.

or during the fragmentation process. In addition, low p_t charged particles can be bent away from the jet cone. The out-of-cone correction aims at accounting for the energy lost outside the jet reconstruction cone.

The correction factor is calculated using Monte Carlo PYTHIA simulations, by comparing the measured jet energy (with all the previous corrections applied) with that of the generated quark. Figure 3.4 shows the out-of-cone correction as a function of jet P_t.

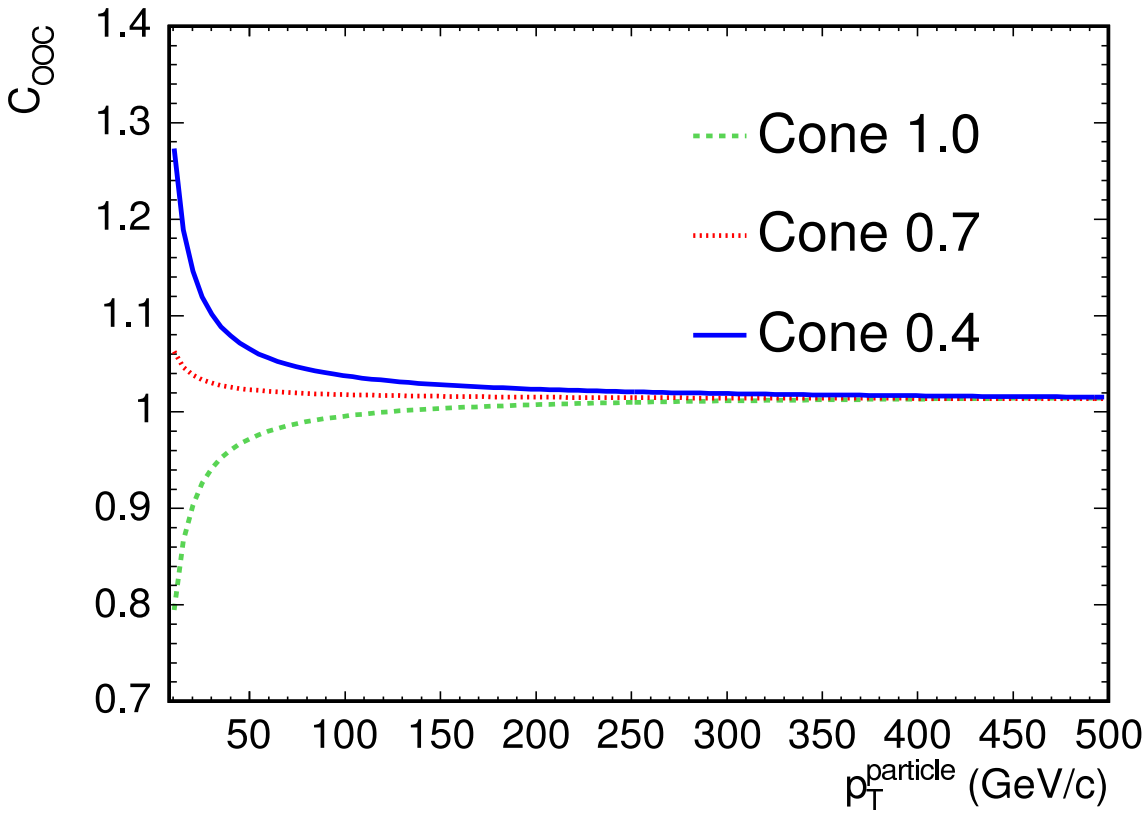

Figure 3.4: The out-of-cone jet energy scale C_{OOC} as a function of jet P_t. This correction factor is used to correct the jet energy that falls in the indicated cone size to the total energy of the jet.

Chapter 4

Event Selections

In this chapter, we discuss the $t\bar{t}$ candidates selection and the expected signal and background yields. We first describe the trigger selections, which are applied during data taking, and then the offline selections, which follow the full reconstruction described in Chapter 3.

4.1 The trigger selections

Events of interest are first selected at the trigger level. The trigger requirement for this analysis is the identification of a high p_t electron or muon. The electron can be in the central region or in the plug region. We only accept muons in the central region.

4.1.1 High p_t central electron trigger

At level 1, the high p_t central electron trigger requires a calorimeter trigger tower with $E_t > 8$ GeV. The ratio of the energy deposited in the hadronic and electromagnetic portions of the calorimeter is selected at $E_{had}/E_{em} < 0.125$. One XFT track with $p_t > 8$ GeV/c must be pointing at the tower. At level 2, we require a cluster with $E_t > 16\,\mathrm{GeV}$ and the XFT tracks with $p_t > 8$ GeV/c must be matched to the seed

tower of the cluster. At level 3, we require a cluster with $E_t > 18$ GeV and a fully reconstructed 3-D track with $p_t > 9$ GeV/c pointing at the cluster.

The central electron trigger efficiency is dominated by the XFT efficiency. We measure the XFT efficiency for electrons using a trigger that selects events consistent with a W boson decaying to an electron and a neutrino, that is with large missing E_t and an electromagnetic calorimeter cluster. However, the electromagnetic cluster is not necessarily matched with a track at trigger level. The observation of a matching track indicates that the XFT successfully reconstructed the track. Using the electron E_t distribution from PYTHIA simulated $t\bar{t}$ events, we obtain a weighted average central electron trigger efficiency of 98.6%.

4.1.2 High p_t central muon trigger

At level 1, the high p_t central muon trigger requires hits with arrival time within 124 ns in the CMUP (CMU+CMP) or CMX. These hits must be matched with a XFT track with $p_t > 8$ GeV/c. At level 3, we require a fully reconstructed 3-D track with $p_t > 18$ GeV/c matched to the muon detector stub.

The muon trigger efficiency is measured using $Z \to \mu^+\mu^-$ events, where we trigger on the first muon and check if we triggered on the second muon as well. The first muon must be fully identified at trigger level, while the second one might just be an oppositely charged stubless track. The invariant mass of the two must be consistent with the Z mass hypothesis. By comparing the number of cases where we have two identified muons at trigger level to the number of cases where we have only one, we can extract the muon trigger efficiency. We measure 90.8% in the CMUP and 96.5% in the CMX.

4.1.3 High p_t plug electron trigger

At level 1, the high p_t plug electron trigger requires a calorimeter trigger tower with $E_t > 8$ GeV and $E_{had}/E_{em} < 0.125$. To cut the background rate, we require $\not\!\!E_T > 15$

GeV. At level 2, we require a cluster with $E_t > 20$ GeV in the $1.1 < |\eta| < 3.6$ region. At level 3, we require $\not\!\!E_T > 20$ GeV.

The main plug electron trigger inefficiency comes from the level 2 $E_t > 20$ GeV requirement. We compare the events obtained with a trigger path that applies the same trigger requirement, except for the level 2 calorimeter cluster. The efficiency of the level 1 $\not\!\!E_T > 15$ GeV cut was also studied, using Monte Carlo $t\bar{t}$ dilepton events to ensure we account for the effect that the second lepton (especially if it is a muon) will have on the $\not\!\!E_T$. We select plug electron candidates passing all cuts and count how many have uncorrected $\not\!\!E_T > 15$ GeV. The overall efficiency of the plug electron trigger is 94.5%.

4.2 Offline selections

In the $t\bar{t}$ dilepton channel, the signature consists of two high E_t jets, two charged leptons (electron or muon) and large $\not\!\!E_T$corresponding to the undetected neutrinos. In cases where we have τ lepton, we can use a charged particle originating from its decay. The strategy consists of identifying one of the leptons and, for the second one, requiring only a well measured and isolated high p_t track. This increases the chance of accepting τ lepton events, since their decay can be hadronic. Finally, a series of cuts reduce the possibility that the presence of $\not\!\!E_T$was due to a mis-measurement.

4.2.1 Charged Lepton identification

The event selection starts with the identification of a high p_t charged lepton. We first look for central electrons, then for muons in the CMUP, then for muons in the CMX, and finally for plug electrons. After a charged lepton is selected, some criteria are applied regarding its direction relative to the $\not\!\!E_T$(see Section 4.2.4).

Central electron selection

After full reconstruction, the following criteria are used to select a central electron candidate:

- Cluster $E_t > 20$ GeV
- Track $p_t > 10$ GeV/c
- Calorimeter energy in cone of 0.4 around the lepton must be less than 10% of the energy of the lepton
- $\gamma \to e^+e^-$ are reconstructed and rejected
- Cluster energy over track momentum $E/p < 2.0$ (applied for $E_t < 50$ GeV only)
- Track $|z_0| < 60$ cm
- $E_{had}/E_{em} < 0.055 + 0.00045 \times E$ where E is in GeV
- Lateral shower sharing $0.14\frac{\sum_{i towers}(M_i - P_i)}{\sqrt{(0.14\sqrt{E_{EM}})^2 + \sum_{i towers}(\Delta P_i)^2}} < 0.2$ based on electromagnetic shower profiles measured in test beams. Where:

  Sums are over towers in the EM cluster adjacent to the seed tower.

  M_i is the measured energy (GeV).

  P_i is the predicted energy (GeV).

  E_{EM} is the total energy (GeV) in the cluster.

- The position of the electromagnetic cluster measured in the CES must match the track extrapolation within -3.0 cm $< track\,charge \times \Delta x < 1.5$ cm where x is defined in the $r-\phi$ plane. The asymmetry of this cut is to account for bremsstrahlung.
- $|\Delta z(track, cluster)| < 3$ cm.

- COT track quality: >= 3 axial segments and >= 2 stereo segments with more than 5 hits each.

Using Z events, the central electron reconstruction efficiency is calculated to be 81.2±0.4% in data and 81.5 ± 0.1 % in Monte Carlo [14].

Central muon selection

After full reconstruction, the following criteria are used to select a central muon candidate:

- Track $p_t > 20$ GeV/c
- Track $|z_0| < 60$ cm
- Calorimeter transverse energy in cone of 0.4 around the candidate must be less than 10% of its p_t
- $E_{had} < 6 + max(0, 0.028 \times (p - 100))$ GeV, where p is in GeV/c
- $E_{em} < 2 + max(0, 0.0115 \times (p - 100))$ GeV, where p is in GeV/c
- Track $|d_0| < 0.02$ cm if track has silicon hits. Track $|d_0| < 0.2$ cm if track has no silicon hits.
- The position of the stub in the muon detection system must match the track extrapolation in $r - \phi$ within $|\Delta x| < 3$ cm for CMU stubs, within $|\Delta x| < 5$ cm for CMP stubs, and within $|\Delta x| < 6$ cm in the case of a CMX stub.
- Cot track quality: >= 3 axial segments and >= 2 stereo segments with more than 5 hits each.
- Track χ^2 probability (efficiency of $\sim 93\%$)

Using Z events, the CMUP reconstruction efficiency (excluding the track χ^2 probability cut) is calculated to be $87.1 \pm 0.6\,\%$ in data and $92.1 \pm 0.1\,\%$ in Monte Carlo. The CMX reconstruction efficiency (excluding the track χ^2 probability cut) is calculated to be $93.2 \pm 0.6\,\%$ in data and $92.3 \pm 0.1\,\%$ in Monte Carlo [15].

Plug electron selection

After full reconstruction, the following criteria are used to select a plug electron candidate:

- Cluster $E_t > 20$ GeV
- Calorimeter energy in cone of 0.4 around the lepton must be less than 10% of the energy of the lepton
- Silicon track $|z_0| < 60$ cm
- At least 3 silicon hits.
- $E_{had}/E_{em} < 0.05$.
- Electromagnetic shower profile corresponds to shape observed in plug test beam data.
- $1.2 < |\eta| < 2.0$.
- Ratio of the sum of the energy in the 5 central PES strips of the cluster divided by the 9 strips of the cluster > 0.65.
- $\Delta R(PES, PEM) < 3.0$.

Using Z events, the central electron reconstruction efficiency is calculated to be $72.8 \pm 0.7\%$ in data and $76.7 \pm 0.1\,\%$ in Monte Carlo [14].

4.2.2 Isolated track selection

After full reconstruction, the following criteria are used to select the isolated track candidate:

- $\geq$ 24 axial COT hits
- $\geq$ 20 stereo COT hits
- track $p_t > 20$ GeV/c
- Isolation: $\frac{p_T}{p_T+\sum p_T} < 0.9$ where $\sum p_T$ is over tracks within a cone of 0.4, excluding the candidate

4.2.3 Jet selection

For the event selection, jets are corrected with the relative corrections, the multiple interactions, and the absolute jet energy scale, described in Section 3.3.2. We require two jets with $|\eta| < 2.0$ and $E_t > 20$ GeV. We also require the lepton and the isolated track to be outside the jet cones (0.4).

After an event is accepted, we re-correct the jets using all the corrections described in Section 3.3.2.

4.2.4 $\not\!\!E_T$ selection

The $\not\!\!E_T$ is a calorimeter measurement of the sum of the transverse energies of the two undetected neutrinos. It is important that we use all the information available to minimize the effect of any $\not\!\!E_T$ mis-measurement. For instance, if the selected lepton is a muon, $\not\!\!E_T$ is corrected using its track p_t. The missing E_t is also corrected using all the other tracks in the event, accounting for the difference between the track p_t and a corresponding 3×3 block of calorimeter towers. The $\not\!\!E_T$ is also recalculated to account for the corrected jet energies.

We select events with $\not\!\!E_T >$ 25 GeV. When the invariant mass of the lepton candidate and the isolated track is within 15 GeV of the Z mass, we require $\not\!\!E_T >$ 40 GeV. This is to further cut the Drell-Yan background events (see Section 4.5), which feature instrumentally induced $\not\!\!E_T$.

Measures are taken to reject cases of "fake" $\not\!\!E_T$ due to the mis-measurement of one of the leptons. For instance, when a lepton points within 5° of the $\not\!\!E_T$, we take the next one and start the algorithm from the beginning. When an isolated track is parallel or anti-parallel within 5° with the $\not\!\!E_T$, we also start from the beginning with the next lepton hypothesis. We iterate on all identified leptons until all criteria are satisfied or the event is rejected.

After an event was accepted, we re-correct the $\not\!\!E_T$ to account for all jet corrections described in Section 3.3.2.

4.3 Selection due to reconstruction method

An additional selection is related to the top quark mass reconstruction method, the Neutrino Weighting Algorithm (NWA), which is described in Section 5.1. Events where the Neutrino Weighting Algorithm fails to solve for the two neutrinos' 4-momenta are rejected.

4.4 Signal estimates

To estimate the signal yield, we assume a $t\bar{t}$ production cross section of 6.1 pb, which is the theoretical value [16] corresponding to the top quark mass world average of 178 GeV/c^2 from Run I. Using PYTHIA $t\bar{t}$ events, the efficiency times acceptance is estimated to be 0.89 ± 0.06 %. Though the event selection is optimized to accept $t\bar{t}$ events where both W bosons decayed to an electron or a muon with their corresponding neutrinos, dilepton events where one or both of the W bosons decayed to a τ lepton also contribute

to the signal. A priori, we obtain an expected mean number of signal of 19.4 ± 1.4 events [17]. This number is used to study the top quark mass reconstruction method as an input value of pseudo-experiments (see Section 5.5). It is not used to extract the top quark mass from the data.

4.5 Background estimates

Background in the $t\bar{t}$ dilepton channel is expected from three main sources [18]:

- **Drell-Yan decays:** $Z/\gamma \to \ell^+\ell^-$ ($\ell = e, \mu$, or τ). In the Drell-Yan background, no neutrino is produced. Therefore, very few of these should pass the $\not\!\!E_T$ cut. However, we can have false $\not\!\!E_T$ due to the mis-measurement of one or more objects in the event. To estimate the number of Drell-Yan events, we use data events for which the invariant mass of an identified lepton and an isolated track falls in the Z mass window 76 GeV/c^2 $< M <$ 106 GeV/c^2. We take events with $\not\!\!E_T >$ 40 GeV to estimate the number of Drell-Yan events inside the Z mass window, and we use events with $\not\!\!E_T >$ 25 GeV to estimate the number Drell-Yan events outside the Z mass window. The latter is corrected with the ratio of events outside to inside the Z mass window, which is calculated using Monte Carlo. To increase the statistics, the selection is loosened and we accept events with any number of jets, for which we correct at the end. We estimate and subtract the number of non Drell-Yan events in this sample using Monte Carlo and the $t\bar{t}$ theoretical cross section. To minimize any bias caused by using the $t\bar{t}$ cross section in its own measurement, we recalculate the Drell-Yan yield using the final measured value, and iterate until we have a convergence.

- **Diboson production:** In these cases, we have a production of WW, WZ and ZZ, which then decay mimicking the $t\bar{t}$ dilepton signature. Using PYTHIA events, the acceptance times efficiency is estimated to be 0.028 ± 0.006 %.

Background	Expected number	Combined uncertainty
$Z/\gamma \to e^+e^-$	5.1	
$Z/\gamma \to \mu^+\mu^-$	1.1	3.3
$Z/\gamma \to \tau^+\tau^-$	2.1	
Fakes	3.9	1.2
WW	1.4	
WZ	0.4	0.4
ZZ	0.1	
Total	14.1	3.5

Table 4.1: Expected event yields for 359 pb^{-1} from background sources after event selection and application of the NWA method.

- **"Fakes"**: "Fakes" are events that are selected due to a jet being misidentified as a lepton (usually the isolated track). They usually arise from $W + jets$ events. The first step in evaluating the frequency of this happening is to estimate the "fake rate", which is the probability that a jet looks like an isolated track. The denominator is the list of all jets and the numerator is the number of jets which match with the extrapolation of an isolated track in η and ϕ. The fake rate varies from $\sim 0.1\%$ at jet $E_t = 10$ GeV to $\sim 0.5\%$ at jet $E_t = 100$ GeV. To estimate the number of fake events, we look at $W + jets$ data. We don't apply the isolated track selection, but instead require 3 jets. Using the fake rate as a function of jet E_t and ϕ, we estimate the number of fake events.

The expected event yields for these background modes after selection and NWA processing are shown in Table 4.1. A more detailed description of how the background

estimates were obtained can be found in [17], which however exclude the NWA selection (see Section 5.1).

Chapter 5

Top Quark Mass Measurement

In this chapter, we discuss how the top quark mass is reconstructed for each selected event. We also discuss how the likelihood function extracts the most probable top quark mass from a set of events by comparing it to signal and background templates. Before extracting the top quark mass from the data, we describe the series of pseudo-experiment tests that were performed to ensure that the method is unbiased.

5.1 $t\bar{t}$ dilepton reconstruction

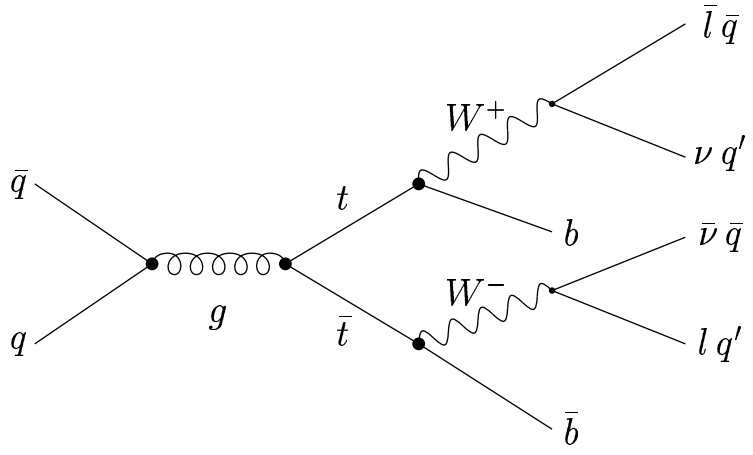

Figure 5.1: Main $t\bar{t}$ production and decay process at the Tevatron

To directly reconstruct the top quark mass in the $t\bar{t}$ channel, we need the values of

the 4-momenta of all the six decay products, that is 24 values. We measure 14 values in the detector:

- 3-momenta of two jets (6 values)
- 3-momenta of two charged leptons (6 values)
- Missing transverse energy in x and y (2 values)

We can also make the following assumptions:

- If we make the hypothesis that we are indeed observing the $t\bar{t}$ dilepton signature, we can deduce the nature, and therefore the masses of the 6 particles (6 values).
- For the two W bosons, we can constrain the invariant mass of its decay products to equal the known W boson mass (2 values).
- Similarly, we can constrain the invariant mass of the top and anti-top decay products to equal each other (1 value).

Despite the assumptions made, because we don't detect the two neutrinos, we still only have 23 of the 24 values needed to directly reconstruct the top quark mass.

The strategy used to solve this under-constrained problem and to obtain the top quark mass with the information available is the Neutrino Weighting Algorithm (NWA) [19].

5.2 Neutrino Weighting Algorithm

The first step in the NWA method consists in assigning a top quark mass hypothesis from 80 GeV/c^2 to 380 GeV/c^2 with increments of 1 GeV/c^2. We also make a hypothesis of the η direction of the two neutrinos, scanning from $\eta = -3$ to $\eta = 3$ with increments of 0.1. Meanwhile, we set the $\not\!\!E_T$ measurement in x and y aside and reserve it to assign a weight

to each top quark mass and neutrino η configuration hypothesis. With these additional assumptions, we can solve the event kinematics for the 4-vectors of the two neutrinos and assign a weight to this top quark mass hypothesis depending on how consistent it is with the measured $\not\!\!E_T$.

In detail, the algorithm works as follows. Assuming we know the correct lepton-jet pairing ($\ell * jet$), m_W, m_t and using $\boldsymbol{b}$, $\boldsymbol{\ell}$ and $\boldsymbol{\nu}$ for the 4-vectors of the b-quark, the lepton and the neutrino, respectively, so that:

$$(\boldsymbol{\ell}+\boldsymbol{\nu})^2 = m_W^2 \quad , \quad (\boldsymbol{b}+\boldsymbol{\ell}+\boldsymbol{\nu})^2 = m_t^2$$

we can write for the neutrino side [20][21]:

$$B \equiv 2\boldsymbol{b\nu} = m_t^2 - m_W^2 - m_b^2 - 2\boldsymbol{b\ell} \tag{5.1}$$

$$\begin{aligned} L \equiv 2\boldsymbol{\ell\nu} &= m_W^2 - m_\ell^2 - m_\nu^2 \\ &= m_W^2 - m_\ell^2 \end{aligned} \tag{5.2}$$

where we have used $m_\nu = 0$. We thus have two equations with three unknowns (P_x, P_y, P_z). We then assume we know the neutrino η and write its 4-vector as:

$$\boldsymbol{\nu} \equiv (P_x, P_y, P_z, E) = (P_T cos\phi, P_T sin\phi, P_T sinh\eta, P_T cosh\eta) \tag{5.3}$$

This allows us to solve Equations 5.1 and 5.2 for B and L simultaneously and get ϕ from a quadratic equation; sometimes we get two and sometimes we get no solutions. We then use L to get the P_T for each ϕ solution and we end up with two or no possible neutrino 4-vectors. Using for the 4-vectors of the b and the lepton:

$$\begin{aligned} \boldsymbol{b} &\equiv (P_x^b, P_y^b, P_z^b, E^b) \\ \boldsymbol{\ell} &\equiv (P_x^\ell, P_y^\ell, P_z^\ell, E^\ell) \end{aligned}$$

we get:

$$B = 2\boldsymbol{b\nu} = 2P_T(E^b cosh\eta - P_x^b cos\phi - P_y^b sin\phi - P_z^b sinh\eta) \tag{5.4}$$

$$L = 2\boldsymbol{\ell\nu} = 2P_T(E^\ell cosh\eta - P_x^\ell cos\phi - P_y^\ell sin\phi - P_z^\ell sinh\eta) \tag{5.5}$$

which can be solved simultaneously for ϕ:

$$\beta = cos(\phi - \alpha) \Rightarrow \begin{cases} \phi = \alpha \pm arccos\beta & (|\beta| \leq 1) \\ \text{“no } \phi \text{ solutions”} & (|\beta| > 1) \end{cases} \tag{5.6}$$

where:

$$tan\alpha = \frac{LP_y^b - BP_y^\ell}{LP_x^b - BP_x^\ell} \tag{5.7}$$

$$\beta = \frac{(LE^b - BE^\ell)cosh\eta - (LP_z^b - BP_z^\ell)sinh\eta}{\sqrt{(LP_x^b - BP_x^b)^2 + (LP_y^b - BP_y^b)^2}} \tag{5.8}$$

We repeat the above steps for the antineutrino side and we thus have either no simultaneous solutions or up to four pairs of neutrino-antineutrino 4-vectors ($\boldsymbol{\nu}$, $\bar{\boldsymbol{\nu}}$). Each solution gets a different weight (w_i for the i^{th} solution) based on how consistent it is with the measured $\not\!\!E_T$:

$$w_i = exp(-\frac{(\not\!\!E_x - P_x^\nu - P_x^{\bar{\nu}})^2}{2\sigma_x^2}) \cdot exp(-\frac{(\not\!\!E_y - P_y^\nu - P_y^{\bar{\nu}})^2}{2\sigma_y^2}) \tag{5.9}$$

where σ_x and σ_y are the resolutions on the $\not\!\!E_x$ and $\not\!\!E_y$ measurements, respectively. The value of the σ_x and σ_y has been determined by comparing the simulated $\not\!\!E_T$ to the generated neutrinos in $t\bar{t}$ Monte Carlo. Figure 5.2 shows the distribution of the difference between the $\not\!\!E_T$ measurement to the sum of the true neutrino E_t. We take a constant resolution of 15 GeV. To ensure that this choice causes no bias, we doubled it and halved it, and saw a negligible difference in the reconstructed top quark mass (see Section 6.5).

Given the assumed top quark mass and the neutrino η values, any of the four solution pairs ($\boldsymbol{\nu}$, $\bar{\boldsymbol{\nu}}$) could have occurred in nature. We therefore add up the four weights:

$$w(m_t, \eta_\nu, \eta_{\bar{\nu}}, \ell * jet) = \sum_{i=1}^{4} w_i \tag{5.10}$$

Since we have not measured the neutrino η's, we loop over all possible $(\eta_\nu, \eta_{\bar{\nu}})$ pairs. Based on the neutrino η distributions from Monte Carlo generation (see Figure 5.3),

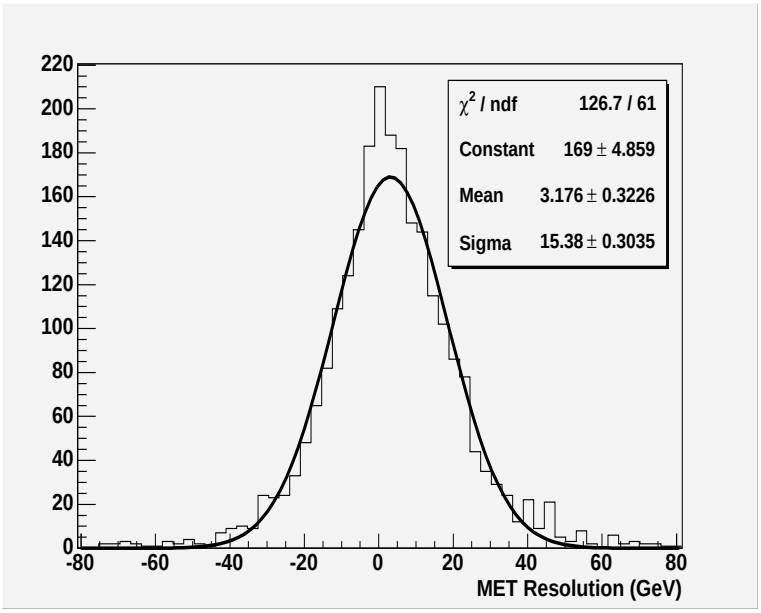

Figure 5.2: Difference between the simulated $\not\!\!E_T$ and the added E_t of the two generated neutrinos. A fitted Gaussian is shown.

each pair is assigned a probability of occurrence $P(\eta_\nu, \eta_{\bar{\nu}})$, which is taken into account when summing the weights of all the possible $(\eta_\nu, \eta_{\bar{\nu}})$ pairs:

$$w(m_t, \ell * jet) = \sum_{\eta_\nu, \eta_{\bar{\nu}}} P(\eta_\nu, \eta_{\bar{\nu}}) \cdot w(m_t, \eta_\nu, \eta_{\bar{\nu}}, \ell * jet) \tag{5.11}$$

The last source of ambiguity arises from the fact that we do not distinguish b jets from $\bar{b}$ jets and thus the correct pairing of leptons and jets is unknown. So, the problem is solved with both possible lepton-jet pairings and the two resulting weights are added up. Therefore, the final weight is only a function of the top quark mass, with all other unknowns integrated out:

$$W(m_t) = \sum_{\ell * jet} w(m_t, \ell * jet) \tag{5.12}$$

We use an input W boson mass of 80.5 GeV/c^2 [22]. We studied the effect of the choice of W boson mass value and found that it has a negligible effect on the top quark mass determination.

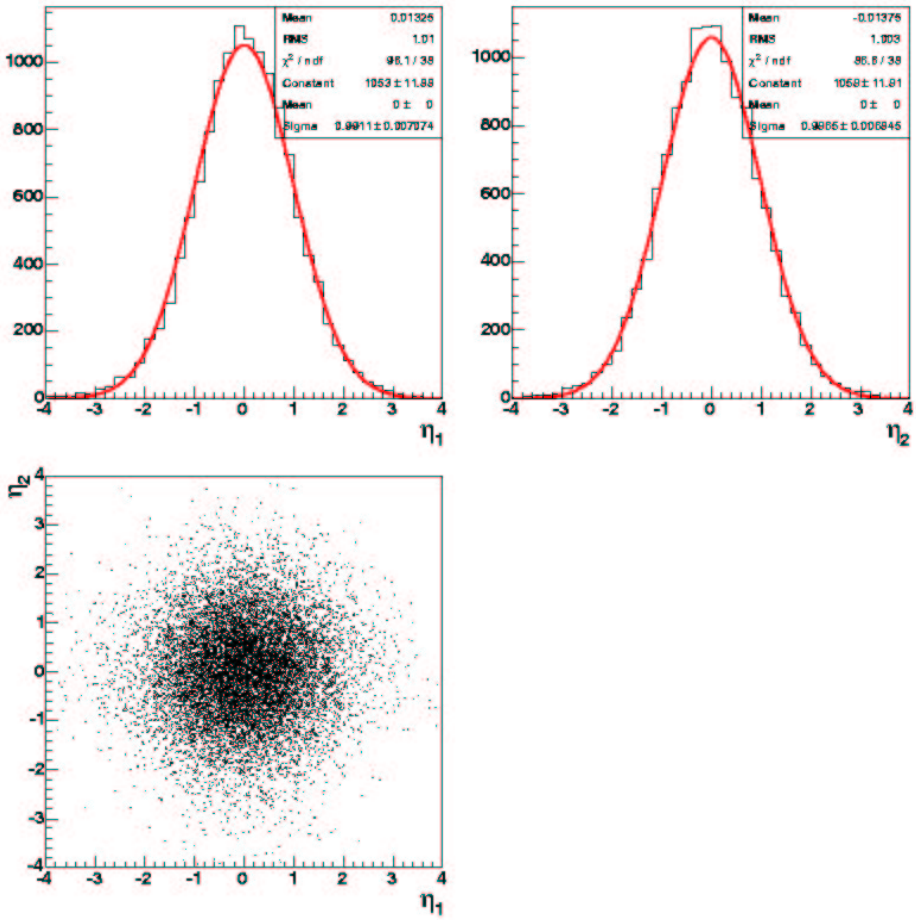

Figure 5.3: η distribution of the two neutrinos from $t\bar{t}$ dilepton Monte Carlo events (top left and right). The correlation between the two is also shown (bottom left).

From the weight distribution, we tried three methods to extract a top quark mass most representative of the distribution:

- The most probable top quark mass
- The median of the distribution
- We select the region defined by the half maximum of the most probable top quark mass peak and calculate the mean within that region.

The three methods gave the same precision on the final top quark mass measurement. For simplicity, we used the most probable top quark mass of the distribution.

5.3 Mass template construction

We apply the above NWA procedure to Monte Carlo dilepton signal and background samples to form template distributions for each sample. These templates are parameterized to build probability density functions which describe the probability of reconstructing a certain top quark mass given a signal or background sample.

5.3.1 Signal templates

To model the signal, we use 31 HERWIG $t\bar{t}$ samples with generated top quark masses ranging from 130 to 230 GeV/c^2. After event selection, we build signal templates using the most probable top quark mass of each event. The templates are fitted with a 6-parameter function. Each of the 6 parameters α_i is constrained to be linearly dependent on the input top quark mass such that we perform a 12 parameter p_i fit on all templates at the same time: *i.e.*, $\alpha_i = p_i + p_{i+6}(M - 175\ \text{GeV/c}^2)$. The fitting function, which represents the probability of observing a reconstructed top quark mass m if the true top quark mass is M, is the addition of a Gaussian and the integrand of the Gamma function [23]:

$$\begin{aligned} P_{signal}(m;M) &= \alpha_5 \frac{\alpha_2^{1+\alpha_1}}{\Gamma(1+\alpha_1)}(m-\alpha_0)^{\alpha_1}\exp\left(-\alpha_2(m-\alpha_0)\right) \\ &+ (1-\alpha_5)\frac{1}{\alpha_4\sqrt{2\pi}}\exp\left(-\frac{(m-\alpha_3)^2}{2\alpha_4^2}\right) \end{aligned} \tag{5.13}$$

This probability distribution function, though arbitrary, is a smooth parameterization of our signal templates as a function of the true top quark mass. Figure 5.4, 5.5, and 5.6 show the signal templates for several generated top quark masses with the overlaid parametrization.

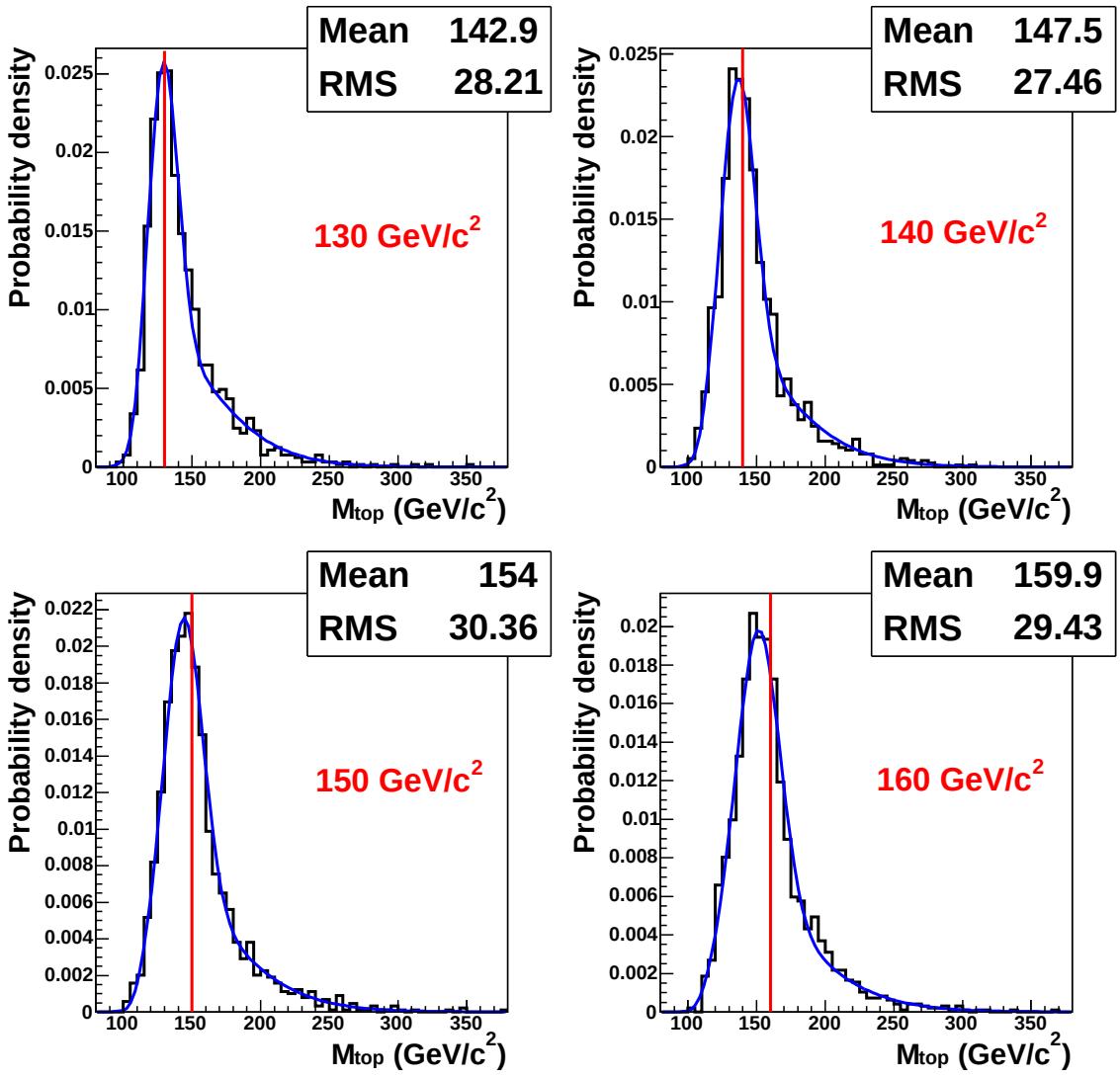

Figure 5.4: Example signal templates constructed from the HERWIG Monte Carlo $t\bar{t}$ samples with masses from 130 to 160 GeV/c^2. The parameterization from Equation 5.13 is shown in blue, while the generated sample mass is indicated by the vertical red line.

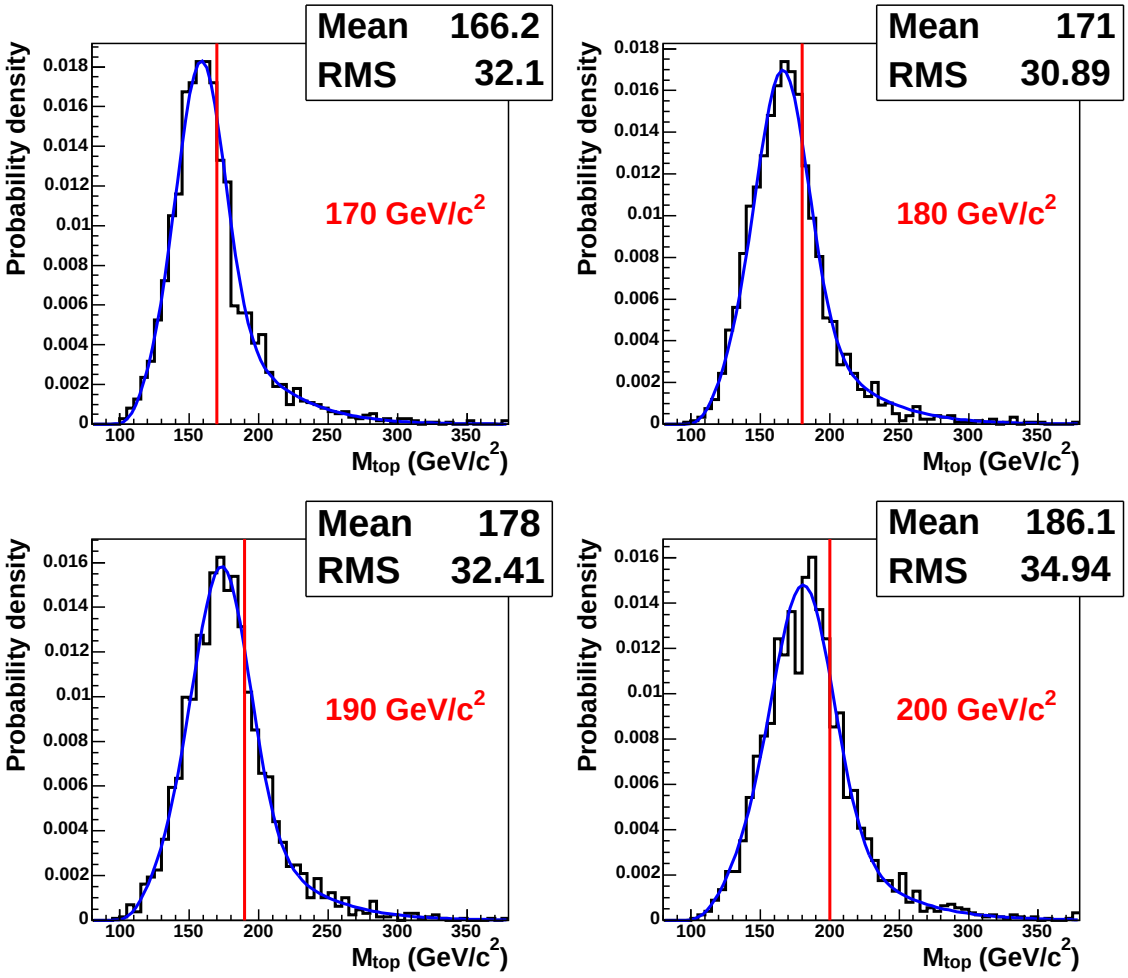

Figure 5.5: Example signal templates constructed from the HERWIG Monte Carlo $t\bar{t}$ samples with masses from 170 to 200 GeV/c^2. The parameterization from Equation 5.13 is shown in blue, while the generated sample mass is indicated by the vertical red line.

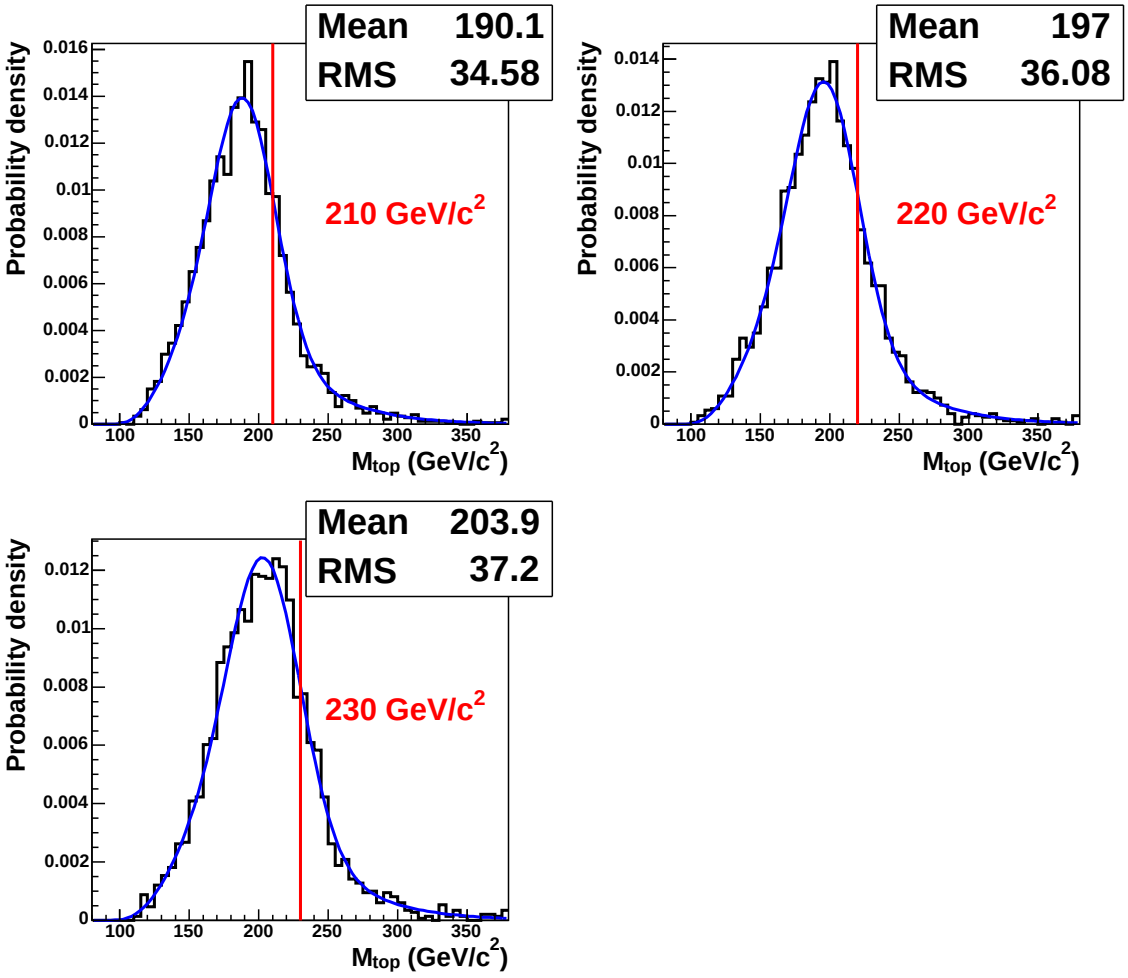

Figure 5.6: Example signal templates constructed from the HERWIG Monte Carlo $t\bar{t}$ samples with masses from 210 to 230 GeV/c^2. The parameterization from Equation 5.13 is shown in blue, while the generated sample mass is indicated by the vertical red line.

5.3.2 Background template

We must also model the background processes (described in Section 4.5) in order to disentangle them from the signal. To model the Drell-Yan background, we use PYTHIA generated events. The "fake" background is modeled with ALPGEN+HERWIG events generated with three jets and one W boson decaying leptonically. Therefore, to pass the event selections, one of the jets must fake a lepton. The diboson samples are also generated with ALPGEN+HERWIG.

Background templates from each source are built in the same way as signal templates. We create one combined background template by adding the individual templates from each background source according to their expected yields (see Table 4.1). We obtain a background probability distribution function by fitting the combined background shape to the following 6 parameter (p_i) function:

$$\begin{aligned} P_{background}(m) &= p_5 \frac{p_2^{1+p_1}}{\Gamma(1+p_1)} (m-p_0)^{p_1} \exp\left(-p_2(m-p_0)\right) \\ &+ (1-p_5) \frac{1}{p_4\sqrt{2\pi}} \exp\left(-\frac{(m-p_3)^2}{2p_4^2}\right) \end{aligned} \tag{5.14}$$

which is similar to that for the signal, but independent of true top quark mass M. Figure 5.7 shows the main background templates and 5.8 shows the combined background template with the overlaid parametrization.

5.4 The likelihood function

The last step in the determination of a final top quark mass from the data sample is to perform a likelihood minimization. The role of the likelihood function described here is to find the probability that our data are described by a certain admixture of background events and dilepton $t\bar{t}$ decays with a certain top quark mass. By minimizing the likelihood function, we find the most likely top quark mass given the set of measurements made.

At the individual event level, we assign a probability that each event i is signal and

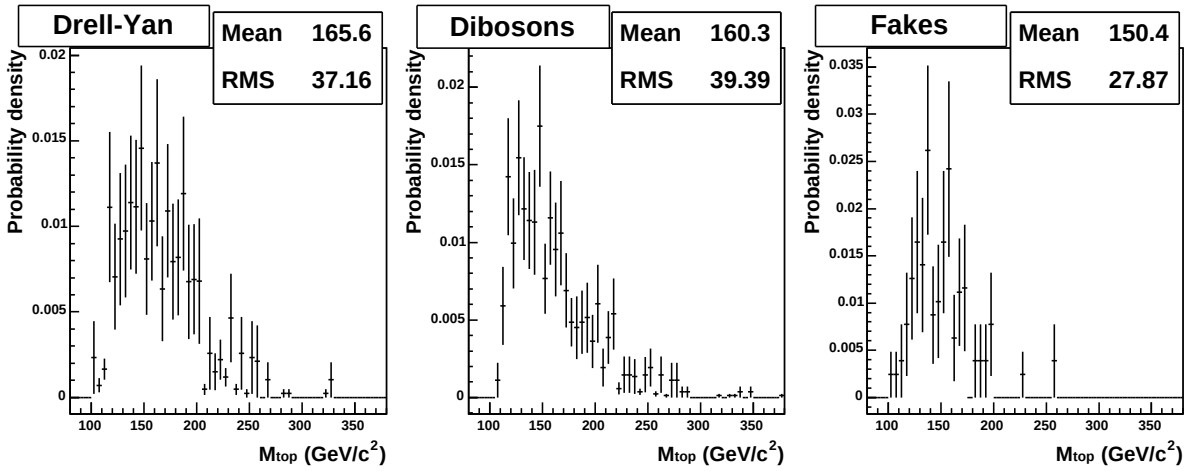

Figure 5.7: Monte Carlo mass templates for the three sources of background.

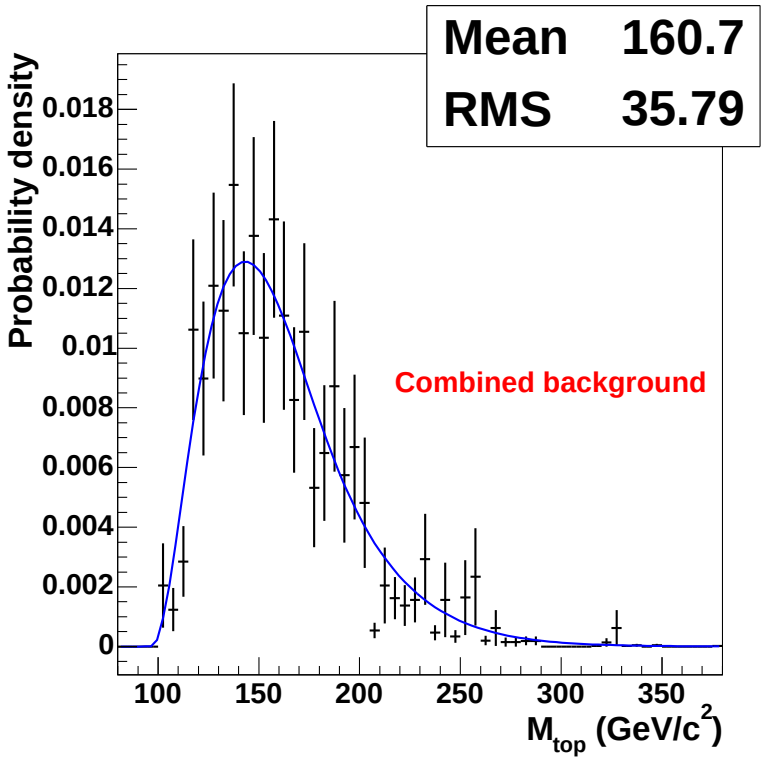

Figure 5.8: Combined NWA background template with parameterization.

a probability that it is background. The signal probability is assigned by comparing the top quark mass values obtained from the data to the template signal distributions, each made with a different generated top quark mass (Figures 5.4 and 5.5). The background probability is derived from a similar comparison with the single template of the combined background shape (Figure 5.8). The number of background events is constrained to the *a priori* background estimate with a Gaussian, while the total number of events is constrained to the observed number of events with a Poisson term. The likelihood becomes:

$$\mathcal{L}(M_{top}) = \mathcal{L}_{shape}(M_{top}) \times \mathcal{L}_{n_b} \tag{5.15}$$

where,

$$\mathcal{L}_{shape}(M_{top}) = \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \prod_{i=1}^{N} \frac{n_s P_{sig}(m_i; M_{top}) + n_b P_b(m_i)}{n_s + n_b} \tag{5.16}$$

and

$$-\ln \mathcal{L}_{n_b} = \frac{(n_b - n_b^{exp})^2}{2\sigma_{n_b}^2} \tag{5.17}$$

Here:

- M_{top} is the unknown true top quark mass
- n_s is the likelihood-estimated number of signal events
- n_b is the likelihood-estimated number of background events
- n_b^{exp} is the *a priori* estimated background, with uncertainty σ_{n_b}
- N is the total number of events in our sample
- m_i is the reconstructed top quark mass for the i^{th} event
- $P_{sig}(m_i; M_{top})$ is the probability that the measurement m_i is signal with top quark mass M_{top} (from Equation 5.13)

- $P_b(m_i)$ is the probability that the measurement m_i is background (from Equation 5.14)

The top quark mass which minimizes $-ln(\mathcal{L})$ is retained. At this stage, the best estimate of the statistical uncertainty is the difference between this mass and the mass at $-ln(\mathcal{L}_{max}) + 0.5$.

5.5 Method check with pseudo-experiment tests

5.5.1 Linearity check

For each $t\bar{t}$ sample with generated top quark mass between 130 GeV/c^2 to 230 GeV/c^2, we perform 10 000 pseudo-experiments, randomly generating events according to the template distributions (see Figures 5.4 and 5.5). The expected number of signal is fluctuated following a Poisson distribution around the central value of 19.4, to simulate statistical fluctuations.

The signal is mixed with background, randomly drawing events from the combined background template (see Figure 5.8). The expected number of background events, 14.1 $\pm$ 3.5 is first fluctuated following a Gaussian distribution with sigma of 3.5 to account for the uncertainty in the evaluation of the background estimation. Then, this Gaussian fluctuated number is fluctuated following a Poisson distribution to simulate statistical fluctuations.

The likelihood machinery, described in the previous section, gives a reconstructed top quark mass from each pseudo-experiment. As seen in Figure 5.9, the output (NWA reconstructed) top quark mass traces the input (generated) top quark mass linearly within uncertainty. The difference with a $y = x$ line can be better observed in Figure 5.10, where we see the offset between the reconstructed mass and the generated mass. The discrepancies, though small compared to our statistical uncertainties, which can be

seen in Figure 5.11 as a function of the generated top quark mass, are corrected for once we obtain a data result (see Section 5.6).

5.5.2 "Pull" distributions

We define the "pull" of a variable as the difference between the reconstructed value and the generated value of the variable, normalized by dividing by the estimated uncertainty of the reconstructed value. The top quark mass "pull" distributions are made by taking

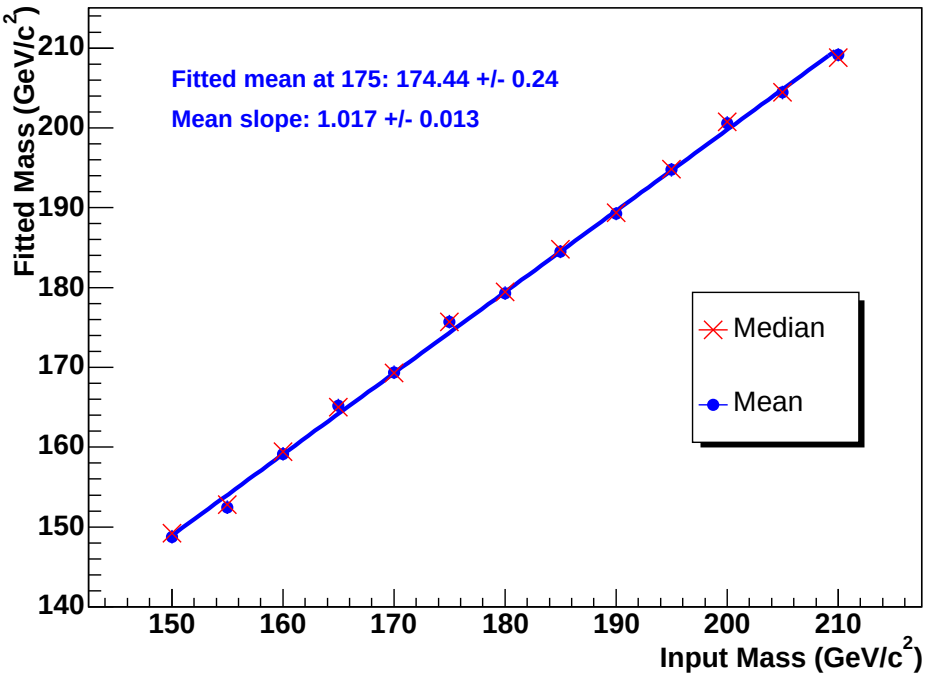

Figure 5.9: The median and mean of the output (reconstructed) top quark mass as a function of the input (generated) top quark mass. For each input top quark mass, 10 000 pseudo-experiments with 19.4 ± 1.4 signal and 14.1 ± 3.5 background events are constructed. The likelihood machinery gives a reconstructed top quark mass for each pseudo-experiment.

the difference between the reconstructed mass and the generated mass in each pseudo-experiment divided by the average of the superior and inferior statistical uncertainty taken from the top quark mass at $-ln(\mathcal{L}_{max}) + 0.5$. Example of the "pull" distributions can be seen in Figures 5.12 and 5.13, and a summary of the mean and sigma of the "pulls" can be seen in Figure 5.14. The "pull" mean is generally consistent with 0 except for a fluctuation at $M_{top} = 155$ GeV/c^2. The "pull" sigma is consistent with 1 above $M_{top} = 175$ GeV/c^2 but consistently higher than 1 below $M_{top} = 175$ GeV/c^2.

The cause of these "pull" sigmas being larger than 1 has been studied. We found that it is due to a discrepancy between the templates and their parametrization. We found that, if we draw events directly from the parametrized probability density functions

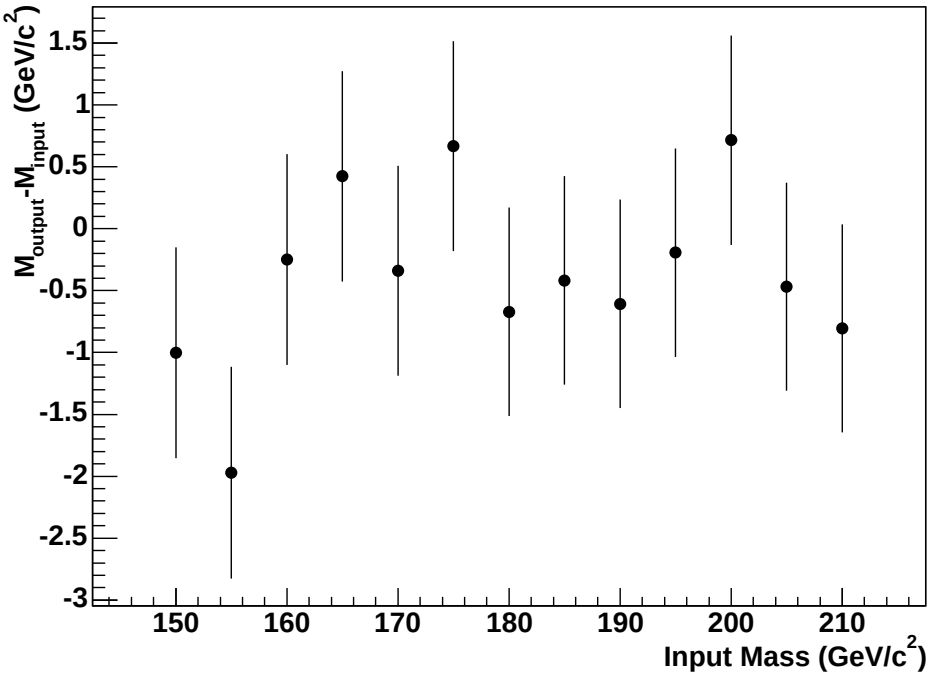

Figure 5.10: Average mass offset (reconstructed mass - generated mass) as a function of the generated top quark mass. The uncertainties shown are correlated, because of the common uncertainty due to the limited background template statistics.

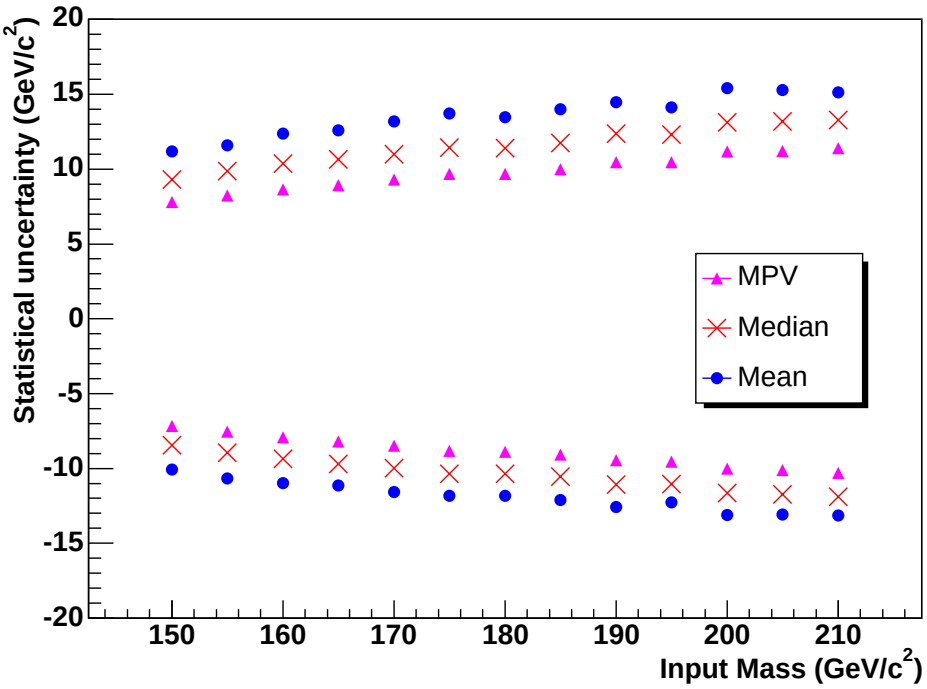

Figure 5.11: The average, median, and most probable statistical uncertainty on the reconstructed top quark mass as a function of the generated top quark mass. The scale factor to account for the observed "pull" width (see Section 5.6) has not been applied here.

(blue curves in Figures 5.4, 5.5, and 5.8), we obtain "pulls" consistent with sigma = 1. Nevertheless, this shows an underestimation of the statistical uncertainty, which will be addressed based on the obtained top quark mass value in data (see Section 5.6).

Now that we studied the consistency and robustness of the method using Monte Carlo of known generated masses, we will, in the next section, verify this consistency further with blinded samples.

5.5.3 Blinded sample tests

Before looking at the data, we analysed five blinded Monte Carlo samples. These contained $t\bar{t}$ signal with unknown randomly generated top quark input masses mixed with background events. They were generated by the members of the CDF top mass group who are not directly involved in this analysis to determine if any bias is present in the reconstruction method. We ran pseudo-experiments on those samples to extract the top quark mass. Table 5.1 shows the difference between the obtained mass and the true generated mass. Four points are within 1 σ and one is at 1.1 σ. Satisfied with this result, we proceeded to analyse the data sample.

Generator	$(mass_{fitted} - mass_{true})/(\mathrm{GeV/c^2})$
HERWIG sample 1	$+0.8 \pm 0.8$
HERWIG sample 2	$+0.2 \pm 0.8$
HERWIG sample 3	$+0.4 \pm 0.8$
PYTHIA sample 1	-0.5 ± 0.8
PYTHIA sample 2	-0.9 ± 0.8

Table 5.1: Results from tests on blinded samples

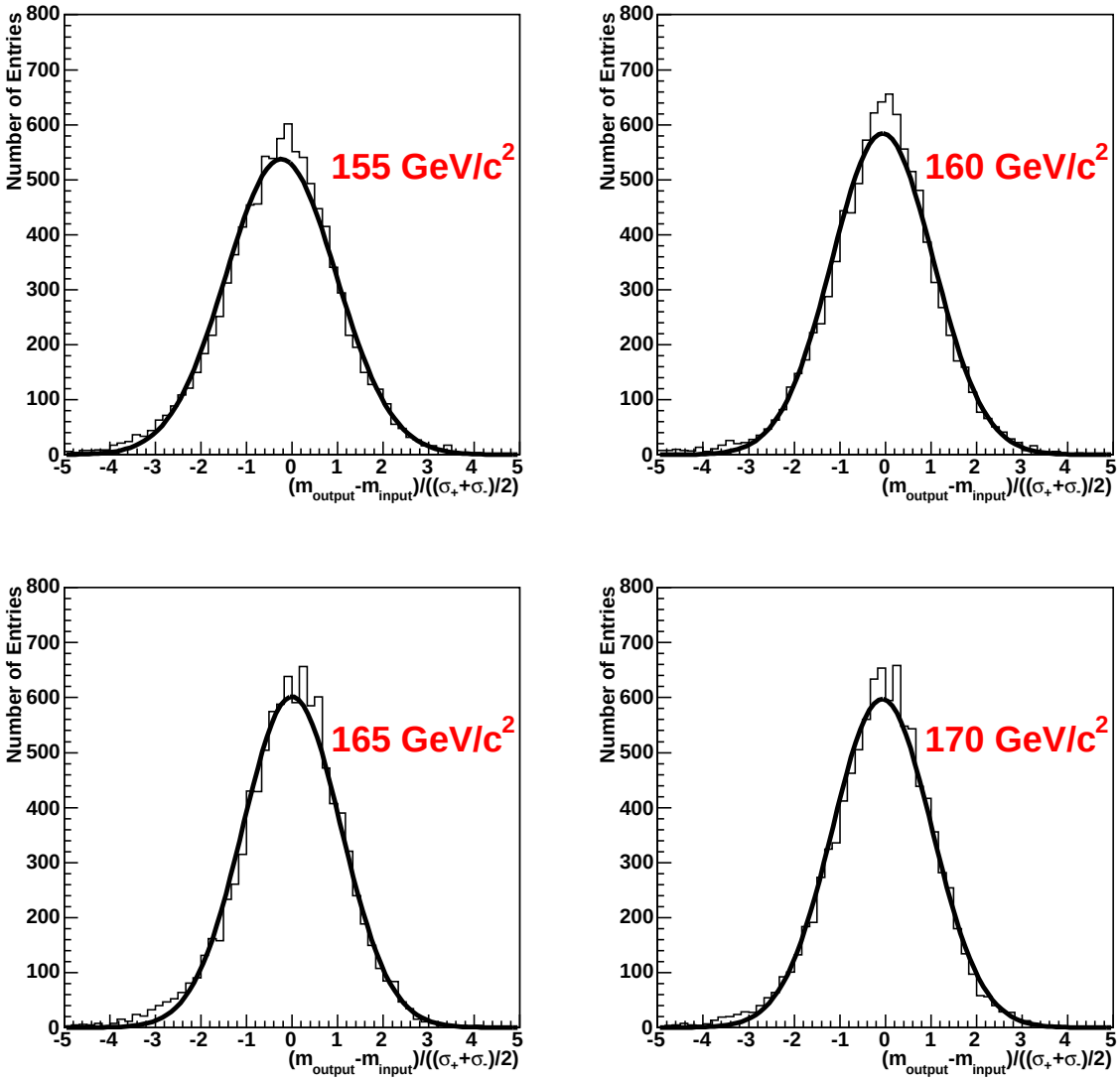

Figure 5.12: The "pull" distribution $\frac{M_{reco}-M_{generated}}{(\sigma_+ + \sigma_-)/2}$ for generated top quark masses from 155 GeV/c^2 to 170 GeV/c^2.

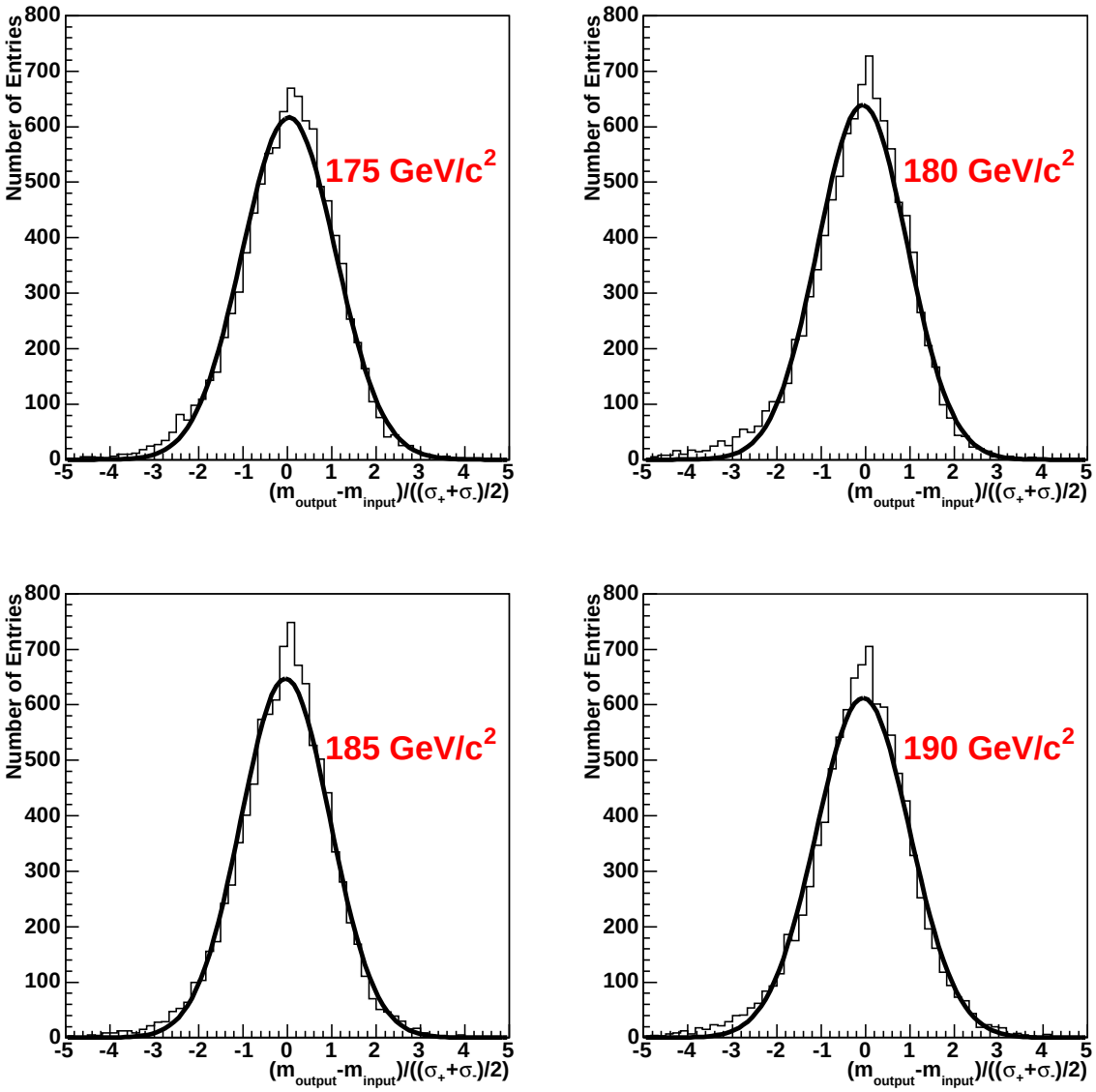

Figure 5.13: The "pull" distribution $\frac{M_{reco}-M_{generated}}{(\sigma_+ + \sigma_-)/2}$ for generated top quark masses from 175 GeV/c^2 to 190 GeV/c^2.

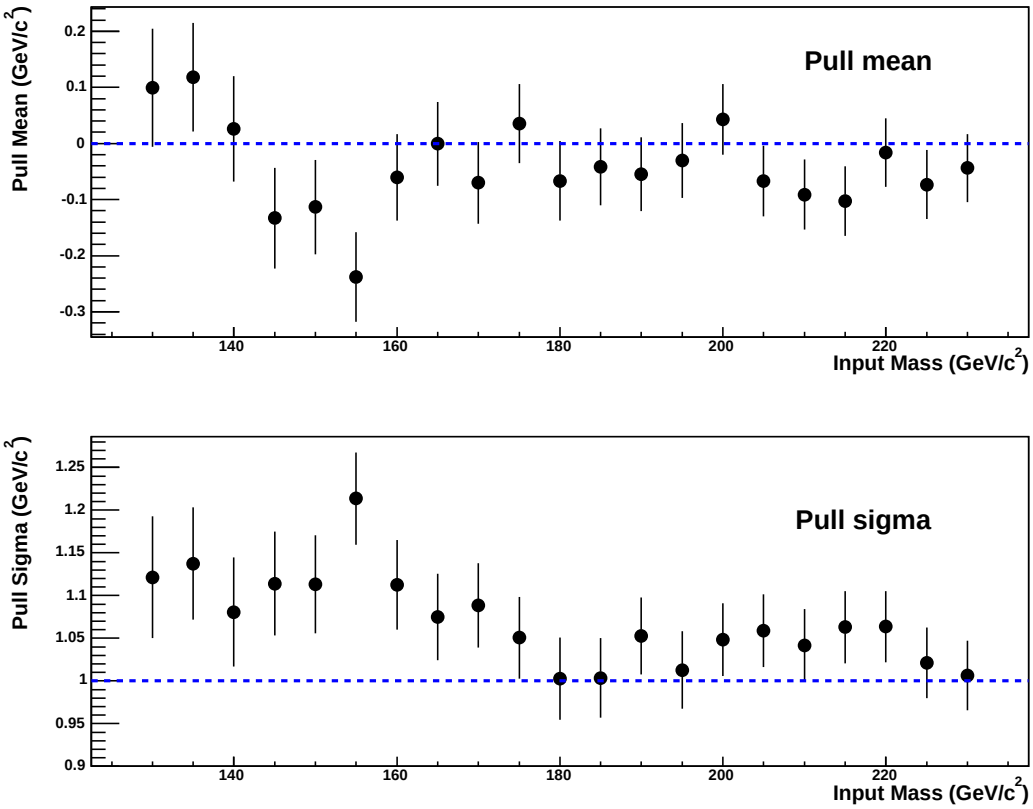

Figure 5.14: Average and sigma of the "pull" distributions.

5.6 Data result

We have 46 events passing the selection cuts in our 359 pb^{-1} of CDF II data. 30 of those events have an identified central electron. Of those which don't have an identified electron, 14 have an identified central muon. The remaining 2 have an identified plug electron. Of the 46 events, 45 have a solution with the NWA method.

We select from each event the most probable top quark mass (shown in Figures 5.17-5.19). We use the Gaussian background constraint of 14.1 $\pm$ 3.5 events as the input parameters of the likelihood function (Equation 5.15). Applying the likelihood minimization procedure described above, we measure a top quark mass of 170.7 GeV/c^2.

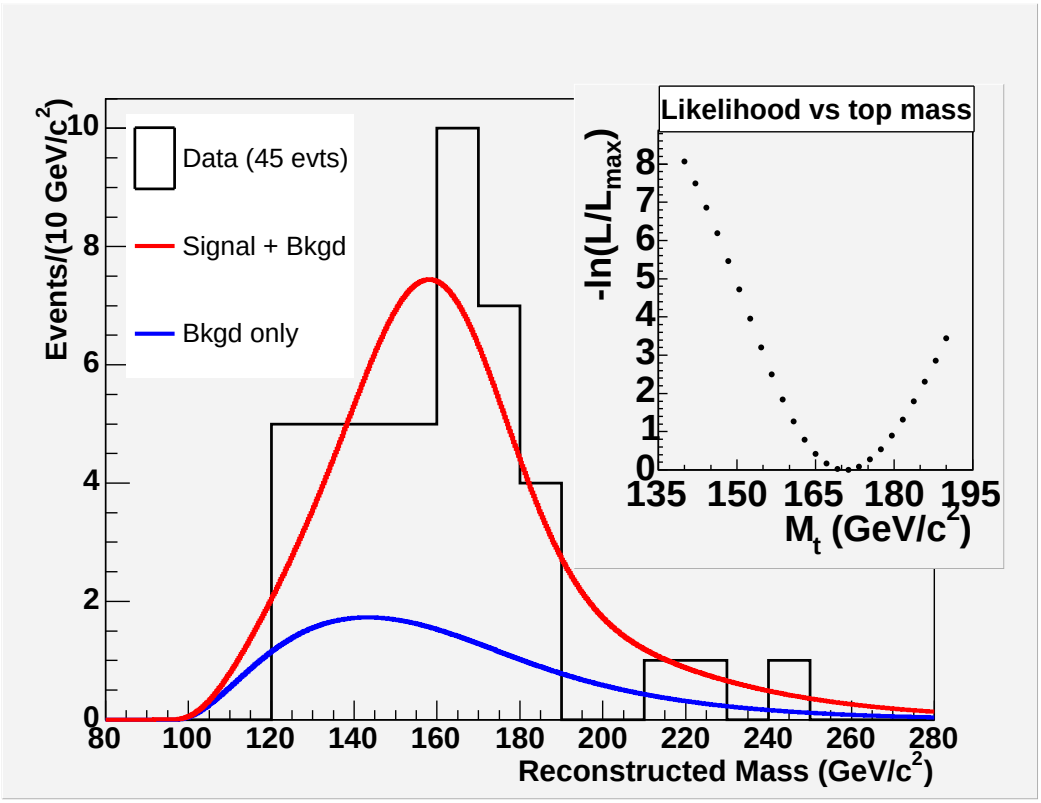

Figure 5.15: Reconstructed top quark mass for the 45 NWA-solved data events (histogram) and the fitted signal and background distribution functions (curves). The likelihood function is also shown (top right).

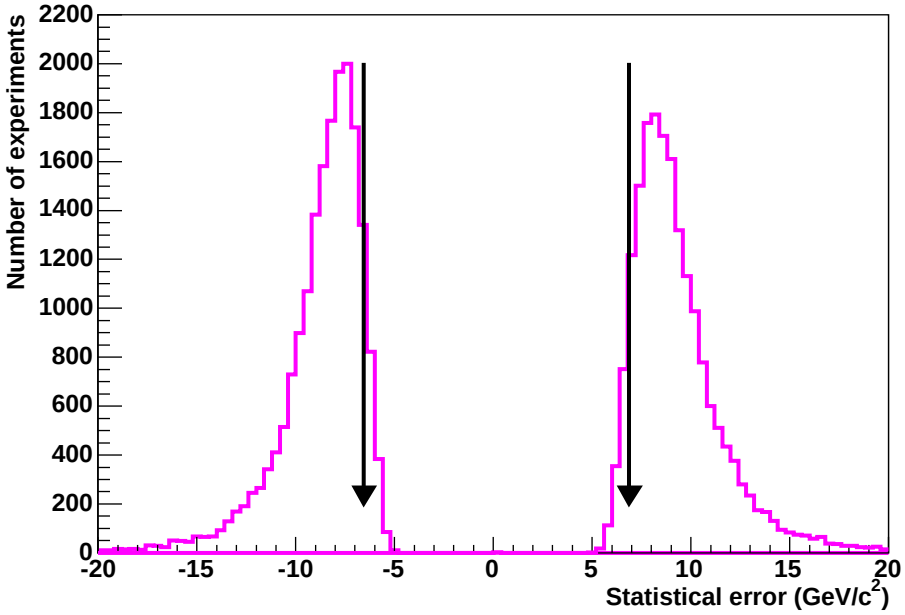

Figure 5.16: Comparison of the observed statistical uncertainties (arrows) from the 45 data events with the statistical uncertainty predicted by Monte Carlo using the 170 GeV/c^2 signal and for which the total number of events fluctuated to 45. 9% of pseudo-experiments give a better statistical uncertainty than the one obtained in data.

Figure 5.15 shows the reconstructed top quark mass and the likelihood function. By taking the width at $-ln(\mathcal{L}_{max}) + 0.5$, we get a statistical uncertainty of $+6.5$ and -6.2 GeV/c^2.

The mean of the "pull" distribution is close to 0 around the obtained mass (see Figure 5.14), and the size of the offset (see Figure 5.10) is much smaller than the obtained statistical uncertainty. However, there is still a small negative bias in the central value of the reconstructed top quark mass. As seen in Figure 5.10, in the $M_{top} < 160$ GeV/c^2 region, the bias is between 1 and 2 GeV/c^2. This is due to the fact that the background is easily confused with the signal in that region. Indeed, if we reduce the background to

a negligible level, this discrepancy is not observed. We study the region where we found the top quark mass $M_{top} > 160$ GeV/c^2 separately. The flat distribution in that region, when fitted with a line, yields a negative bias of 0.1 GeV/c^2. We therefore correct the central value of the top quark mass measurement to 170.8 GeV/c^2.

Since the sigma of the "pull" distribution is consistently higher than 1 around the obtained mass (see Figure 5.14), we need to study the possibility that the statistical uncertainty is underestimated. We scale our uncertainty with a factor such that 68% (area under $\pm 1\sigma$ for a normalized Gaussian) of our pseudo-experiments contain the central value. Using the 170 GeV/c^2 signal template with a corresponding $t\bar{t}$ cross section of 7.8 pb, we find a scale factor of 1.06. Thus, we correct for this and obtain a top quark mass of $170.8^{+6.9}_{-6.5}$ GeV/c^2 (statistical uncertainty only).

The likelihood fit returned a number of signal events of 32.4 ± 7.4 and a number of background events of 13.4 ± 3.5. If we remove the Gaussian constraint on the background (Equation 5.17), the fit converges as a pure signal configuration (45.0 ± 6.7 signal events and 0 ± 69 background events). The resulting top quark mass value is 168.3 ± 4.9 GeV/c^2. Thus, the constrained and unconstrained fits return consistent top quark mass results (given the statistical uncertainty).

The statistical uncertainty measured in the data is lower than the most probable expected uncertainty. We estimate the probability of achieving this result by running pseudo-experiments using the 170 GeV/c^2 signal template with corresponding $\sigma_{t\bar{t}} = 7.8$ pb and $n_{bkg} = 14.1 \pm 3.5$. We select pseudo-experiments where the true total (signal + background) number of events has fluctuated to 45. Figure 5.16 shows the comparison between the statistical uncertainty for Monte Carlo pseudo-experiments and data. Given that we observe 45 events, the probability of obtaining this statistical uncertainty or better is estimated to be 9%.

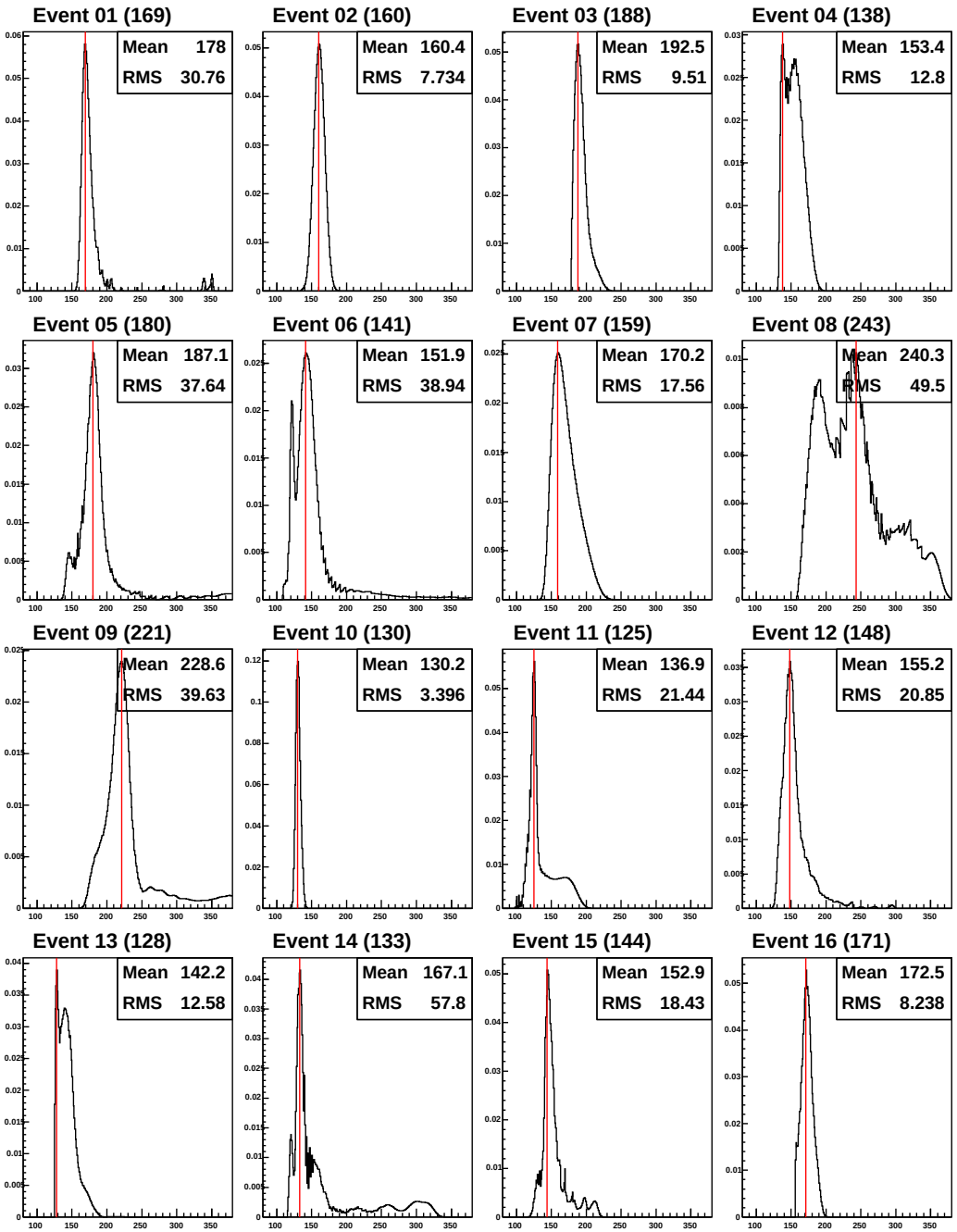

Figure 5.17: Top quark mass weight distribution from NWA for data events 1-16. The vertical red line shows the most probable top quark mass.

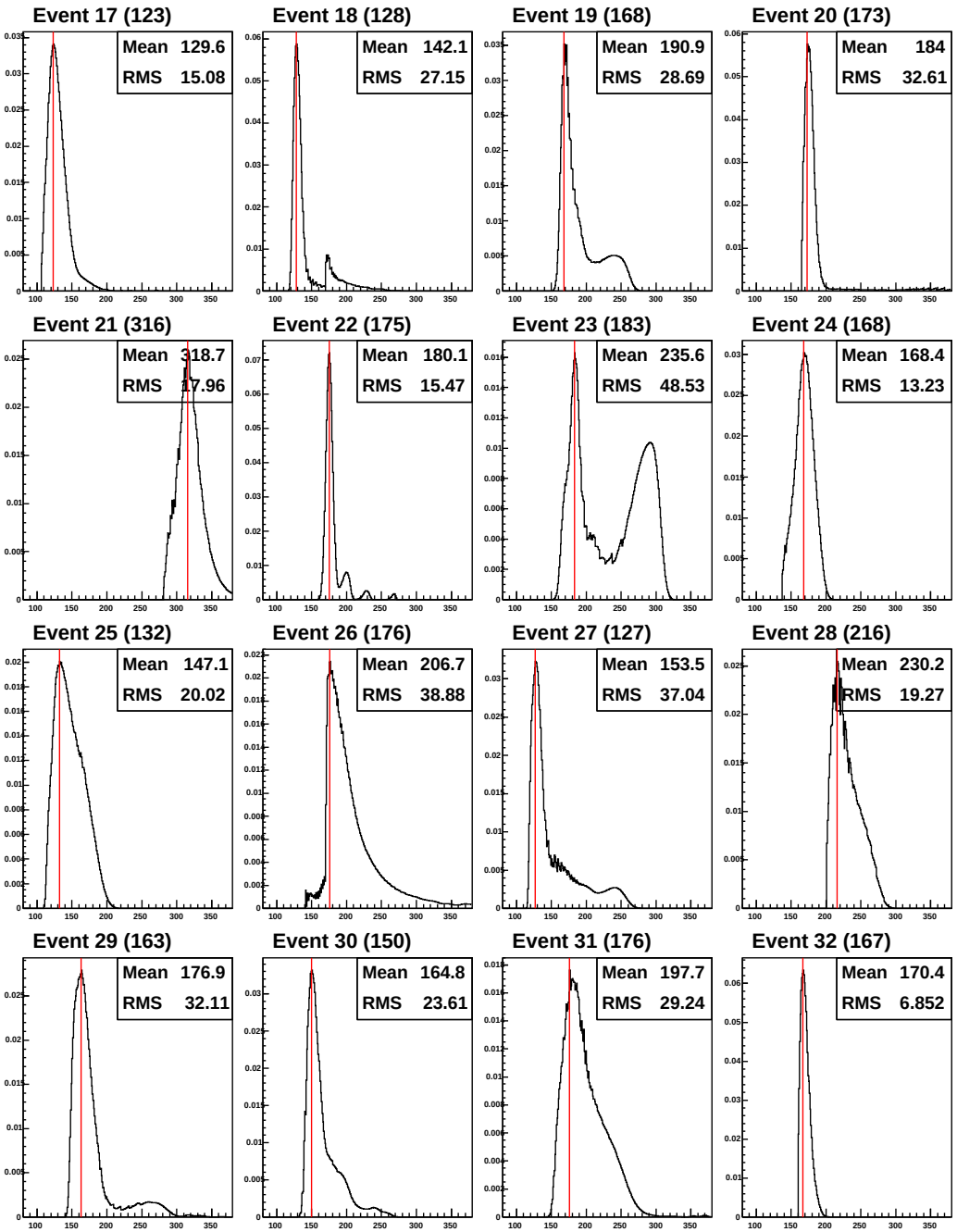

Figure 5.18: Top quark mass weight distribution from NWA for data events 17-32. The vertical red line shows the most probable top quark mass.

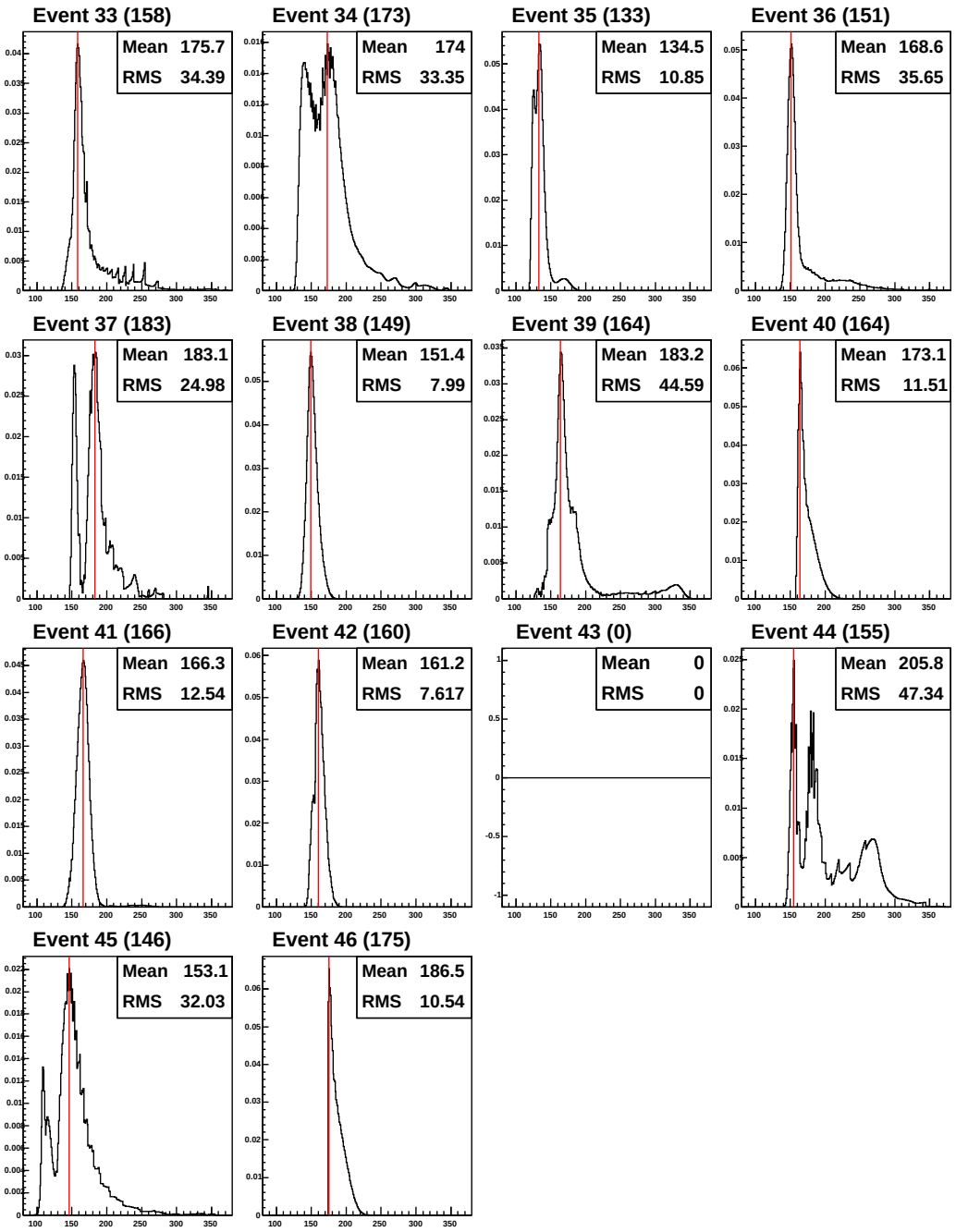

Figure 5.19: Top quark mass weight distribution from NWA for data events 33-46. The vertical red line shows the most probable top quark mass. For event 43, no solution was found.

Chapter 6

Systematic uncertainties

In this chapter, we estimate systematic errors on the top quark mass measurement from a variety of sources. These systematics come from uncertainties in the measured particle kinematics, in the Monte Carlo simulation of $t\bar{t}$ signal and background events, and in the assumptions made when implementing the NWA procedure. The estimates are based on pseudo-experiments drawn from a multitude of control and systematically modified Monte Carlo samples. The systematic uncertainties have been evaluated assuming the *a priori* top quark mass of 178 GeV/c^2 and the corresponding $t\bar{t}$ theoretical cross section of 6.1 pb. We have verified that the result is consistent when we assume other masses and cross sections, such as the *a posteriori* mass of 170 GeV/c^2 and the corresponding $t\bar{t}$ theoretical cross section of 7.8 pb.

6.1 Jet energy scale

In Section 3.3.2, we discussed how the jet energy is corrected. Here we discuss the systematic difference between the jet energy in data and in Monte Carlo simulations. To each jet energy correction is associated a jet energy systematic uncertainty [13].

6.1.1 η relative systematic uncertainty

The η relative jet energy correction aims to compensate for the non-uniformity of the calorimeter response in η. Though we corrected the calorimeter response to that of the central region separately in data and Monte Carlo using the balance of dijets (see Section 3.3.2), a systematic difference remains due to event selection bias. We varied the selections on the 3rd most energetic jet p_t and on the missing transverse energy significance: $\not\!\!E_T/\sum_{i\,jets} E_t$ which are designed to ensure the balance of the two jets. Deviations of the response are taken as a systematic uncertainty. These uncertainties are listed in Table 6.1.

$\lvert\eta\rvert$ range	0.0 - 0.2	0.2 - 0.6	0.6 - 0.9	0.9 - 1.4	1.4 - 2.0	2.0 - 2.6	2.6 - 3.6
$p_t < 12$ GeV/c	1.5%	0.5%	1.5%	2.5%	1.5%	5.0%	7.5%
$12 \leq p_t < 25$ GeV/c	1.5%	0.5%	1.5%	1.5%	1.5%	3.0%	6.0%
$25 \leq p_t < 55$ GeV/c	1.0%	0.5%	1.0%	1.0%	0.5%	1.5%	6.0%
$p_t \geq 55$ GeV/c	0.5%	0.5%	0.5%	0.5%	0.5%	1.5%	6.0%

Table 6.1: Jet energy scale systematic uncertainty on individual jets due to η dependence for different jet η and p_t ranges.

6.1.2 Multiple interactions systematic uncertainty

When products from other $p\bar{p}$ interactions fall in the jet reconstruction cone, it artificially inflates the jet energy. If the Monte Carlo doesn't properly reproduce this effect, it will cause a systematic uncertainty in the jet energy scale. In particular, the vertex reconstruction efficiency and the probability of mistakenly reconstructing fake vertices depend on the topology of the hard interaction. We compare the average transverse energy that falls in a randomly selected cone of 0.4 in minimum bias data, lepton triggered data,

and jet triggered data. The comparison shows no indication of a systematic difference between the various samples. However, due to statistical limitations in this study, we take a systematic uncertainty of 15% of the multiple interaction correction (see Section 3.3.2).

6.1.3 Absolute jet energy scale systematic uncertainty

In the calorimeter, hadrons collide inelastically with nuclei and part of their energy is lost. Therefore, the calorimeter response is smaller than the true energy of the incoming hadron. Due to the creation of π^0s and η^0s, which decay to a pair of photons, a fraction of the hadronic shower is electromagnetic. Since the number of steps in the cascade is dependent on the energy of the incoming hadron, the electromagnetic fraction will depend on it as well. For this reason, the calorimeter response is not linear with the energy of incoming hadrons. There are two consequences to this, which can systematically affect the top mass determination. First, the modeling of the calorimeter response to individual particles can be systematically different in the data and Monte Carlo. Second, the way jets are fragmented affects the simulated energy. For example, we will have a different response between 10 particles of 2 GeV each and 2 particles of 10 GeV each. In this section, the calorimeter response to individual particles and the fragmentation are studied as two separate effects. Figure 6.13 shows a summary of the absolute jet energy scale systematic uncertainties.

Electrons and photons interact electromagnetically with the calorimeter, and therefore their response is closer to unity. A systematic difference between their response in data and Monte Carlo is also studied here.

Calorimeter response to individual particles

The calorimeter response to charged particles is measured as the ratio of the energy deposited in the calorimeter and the momentum of its associated track E/p.

To measure the calorimeter response to charged hadrons, we study isolated tracks in a minimum bias data sample. We measure the track momentum in the COT and the energy deposited in the two central calorimeter compartments (CEM and CHA). Figure 6.1 shows the calorimeter response E/p for data and Monte Carlo. The systematic difference is extrapolated to entire jets using the particle energy distribution in Monte Carlo jets.

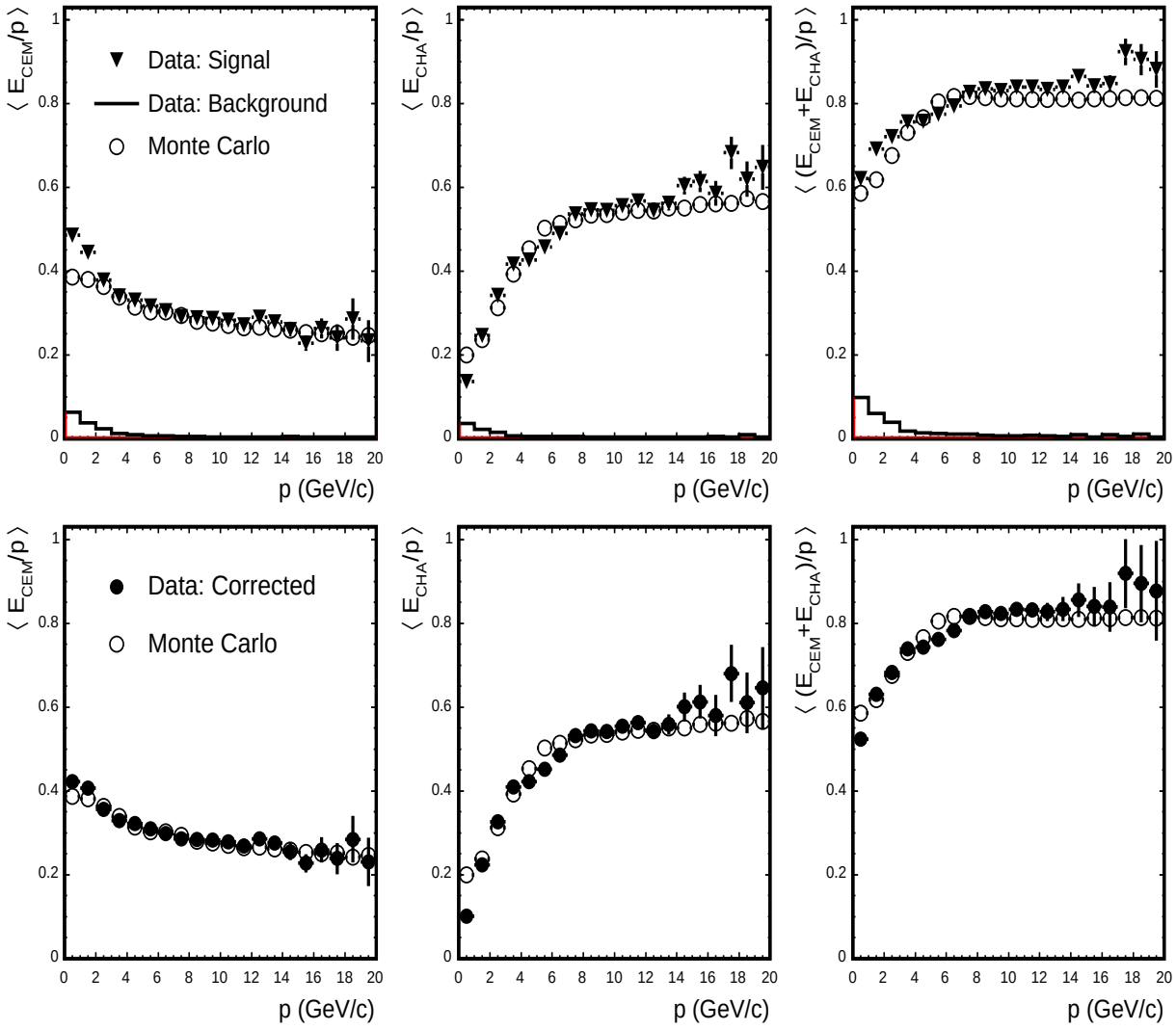

Figure 6.1: Calorimeter response E/p to minimum bias isolated tracks for data and Monte Carlo as a function of track momentum. The top plots show the signal and the background, and the bottom plots show the corrected signal.

To measure the calorimeter response to particles that interact electromagnetically, we study $J/\Psi \to e^+e^-$ and $Z \to e^+e^-$ events. We again compare the track momentum with the energy deposited in the calorimeter. Figure 6.2 shows a good match between data and Monte Carlo. Therefore, the systematic uncertainty on the jet energy scale due the calorimeter response to particles that interact electromagnetically (electrons and photons) is negligible.

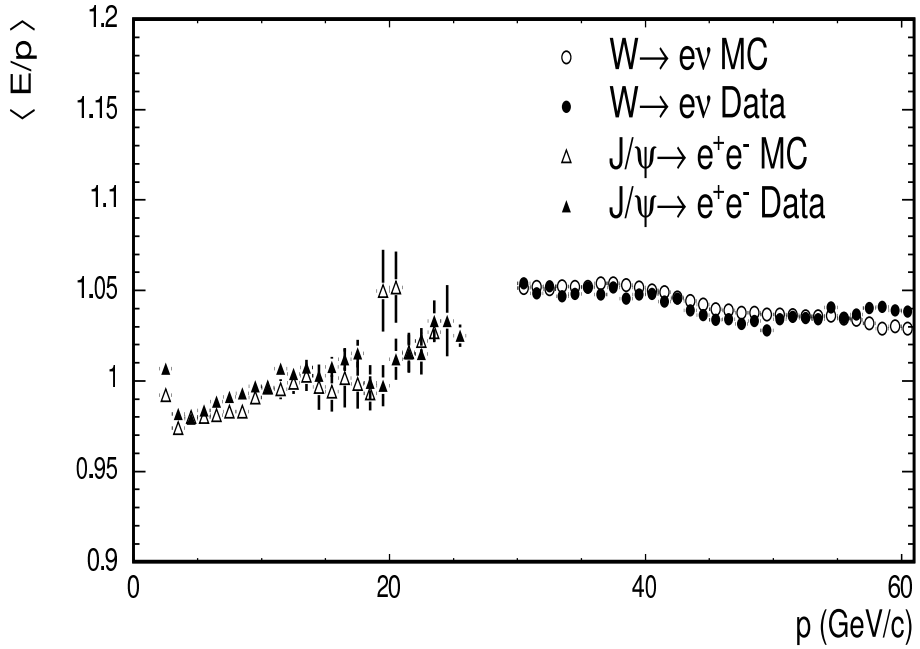

Figure 6.2: Average calorimeter response E/p to electrons in data and Monte Carlo. This result is also taken as the photon calorimeter response.

Jet fragmentation

We are interested in the uncertainty in our modeling of jet fragmentation and its impact on the jet energy scale. We study the fragmentation of jets using the tracks left by their charged particles in the COT. The momentum of the charged tracks that fall inside the jet reconstruction cone informs us about the charged portion of the jet. Due to the high density of COT hits in jets, the tracking efficiency in jets is much lower than that for isolated tracks. In order to properly study the charged portion of jets using the COT,

the first step is to measure the tracking efficiency in jets. This way, we will know how many tracks we lost, and compensate for them.

We use a track embedding technique to estimate the track reconstruction efficiency. The track embedding method consists of embedding Monte Carlo COT tracks in data jets. Hits are simulated for that track and embedded in the event. The event is reconstructed with these additional hits, and the projected trajectory of the embedded charged particle is compared with the reconstructed tracks.

For the efficiency of the embedded track to represent that of an actual data track, it must have the same properties. The tracks are tuned to emulate data tracks in terms of the following characteristics [24]:

- **Hit residual:** Uncertainties on the COT drift time and multiple scattering in the inner tracking system cause uncertainties on the hit position. The hit resolution influences the tracking efficiency, facilitating it when hits are simulated exaggeratedly close to the actual trajectory of the particle. The hit position of the Monte Carlo tracks is tuned using a smearing scale factor such that the hit resolution matches that of the data. On Figure 6.3, we see the track residual for data tracks in jets, embedded track to which this correction has been applied, and, for comparison, tracks in Monte Carlo PYTHIA jets (with no corrections applied).

- **Single hit efficiency:** The hit inefficiency for isolated tracks is caused by COT imperfections and pattern recognition failure. Since the pattern recognition is the same in data and Monte Carlo, we need to disentangle it from the COT imperfection effect. In the default simulation, COT imperfections are neglected, and it is assumed that, for an isolated track, each hit has 100% chance of being read by the sense wire. To measure the single hit efficiency, we looked at the number of hits per superlayer in cosmic tracks that are fiducial to all superlayers of the COT. For each superlayer, Table 6.2 shows the average number of hits found in cosmic tracks. To isolate the COT imperfection effect from the pattern recognition failure

effect, we study isolated Monte Carlo tracks. We found the hit efficiency due to pattern recognition failure to be $\sim 98.5\%$, independently of the superlayer. We use this factor as a denominator of the single hit efficiency and apply the single hit efficiency factor per superlayer to the embedded tracks.

- **Hit merging:** Hit merging occurs when two hit pulses overlap, or if they are close enough in time so that the electronic system can't differentiate them. Due to the high density of COT hits in jets, hit merging can't be neglected when measuring the tracking efficiency in jets. This effect depends on the hit pulse width and the hit merging distance (the distance below which two hits pulses are too close to be differentiated by the electronic readout system). The simulated drift model was tuned to have the same hit pulse width shape as that of the data. The hit merging distance corresponds to a drift time of 10 ns. We apply this correction in the simulation of our embedded track.

A good measure of a track's quality is the number of hits in the COT superlayers. This greatly influences the chances of success or failure to reconstruct a track. Figures 6.4, 6.5, and 6.6 show the total number of hits in the axial and stereo COT superlayers for tracks in jets. Data tracks, uncorrected Monte Carlo tracks, and tracks for which the above corrections have been applied are shown.

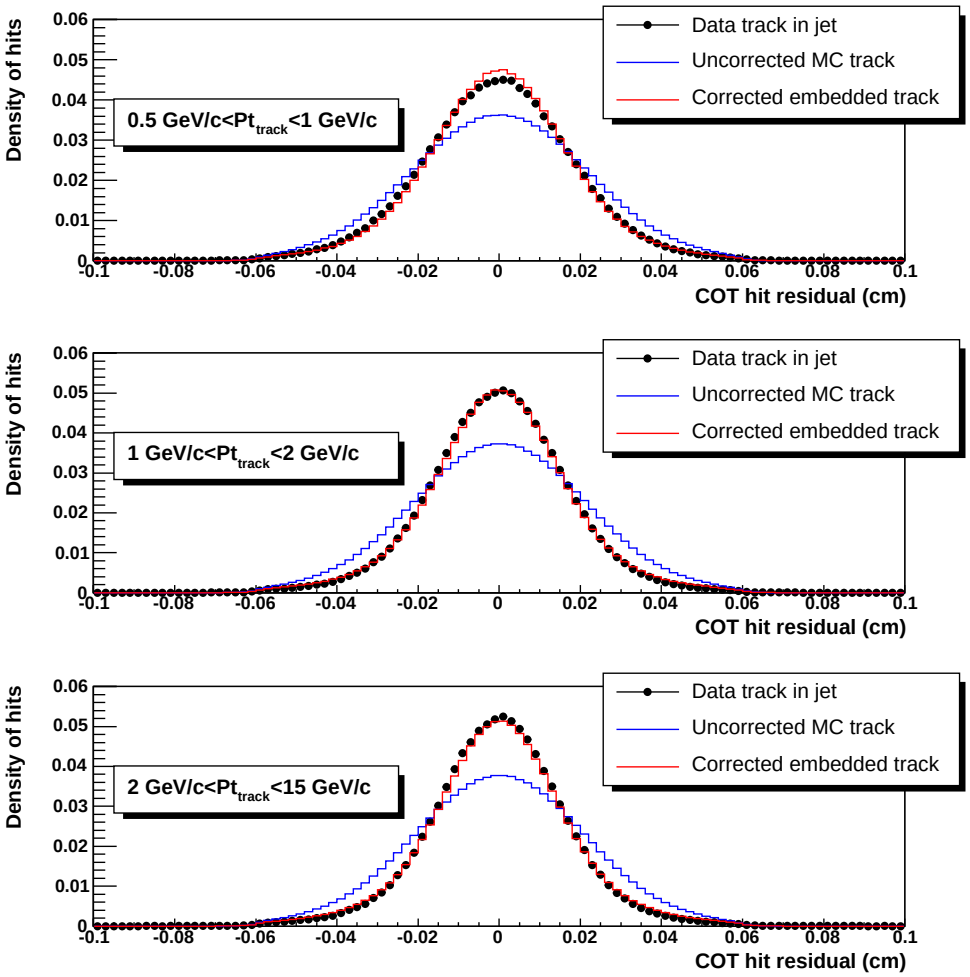

Figure 6.3: Track hit residual for data tracks, uncorrected tracks and tracks with corrected hit position resolution

Superlayer	Average number of hits
1	21.7 ± 0.3
2	22.9 ± 0.3
3	22.2 ± 0.3
4	23.0 ± 0.3
5	22.9 ± 0.3
6	23.1 ± 0.3
7	23.0 ± 0.3
8	23.1 ± 0.3

Table 6.2: Number of hits in cosmic tracks. Note that cosmic tracks traverse the entire COT, doubling the maximum number of hits per superlayer to 24.

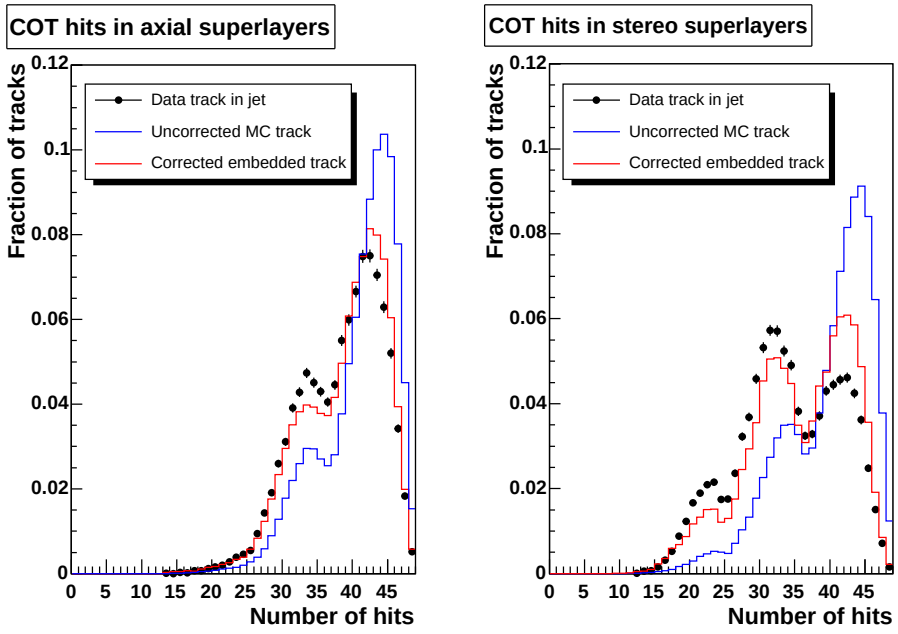

Figure 6.4: Sum of all hits in the four COT axial superlayers (left) and the four COT stereo superlayers (right) for tracks with 0.5 GeV/c $< p_t <$ 1 GeV/c in jets. We show data tracks, uncorrected tracks, and track for which the hit position resolution, the hit pulse width, the single hit efficiency, and the hit merging distance have been corrected.

Once the embedded tracks are tuned to emulate data tracks, they can be used to measure the tracking efficiency. This is done separately in data and Monte Carlo, since the simulation doesn't account for the high occurrence of hit merging in a high hit density area. In Figures 6.7, 6.8, and 6.9, we show the tracking efficiency in jets as a function of jet transverse energy, track transverse momentum, and angular distance $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi_0)^2}$ between the track and the jet centroid.

To count the number of tracks in a jet, we correct for tracks that have not been observed by weighting the observed tracks by one over their expected efficiency. Figure 6.10 show the distribution of the transverse momentum of tracks inside jets. We see an overall agreement between data and MC, for both PYTHIA and HERWIG.

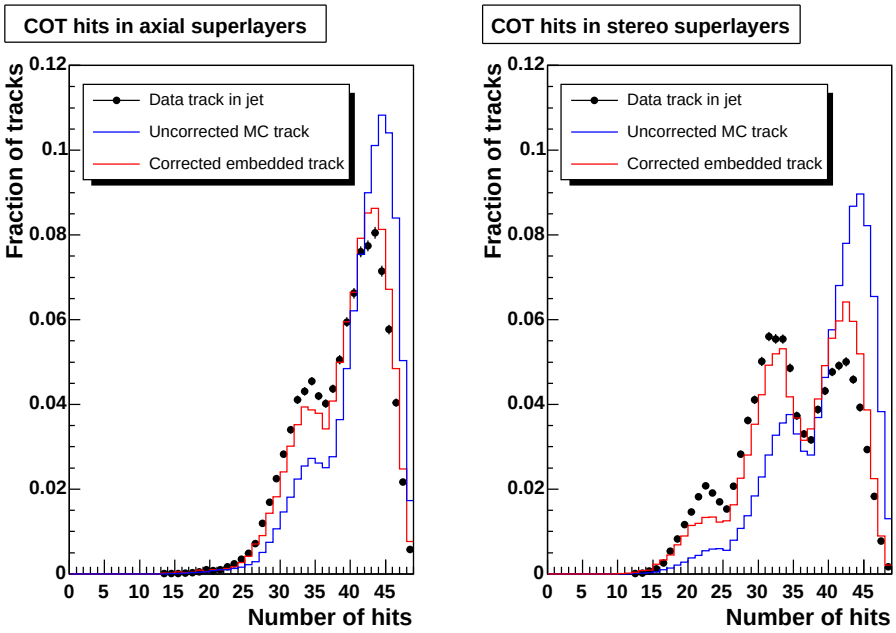

Figure 6.5: Sum of all hits in the four COT axial superlayers (left) and the four COT stereo superlayers (right) for tracks with 1 GeV/c $< p_t <$ 2 GeV/c in jets. We show data tracks, uncorrected tracks, and track for which the hit position resolution, the hit pulse width, the single hit efficiency, and the hit merging distance have been corrected.

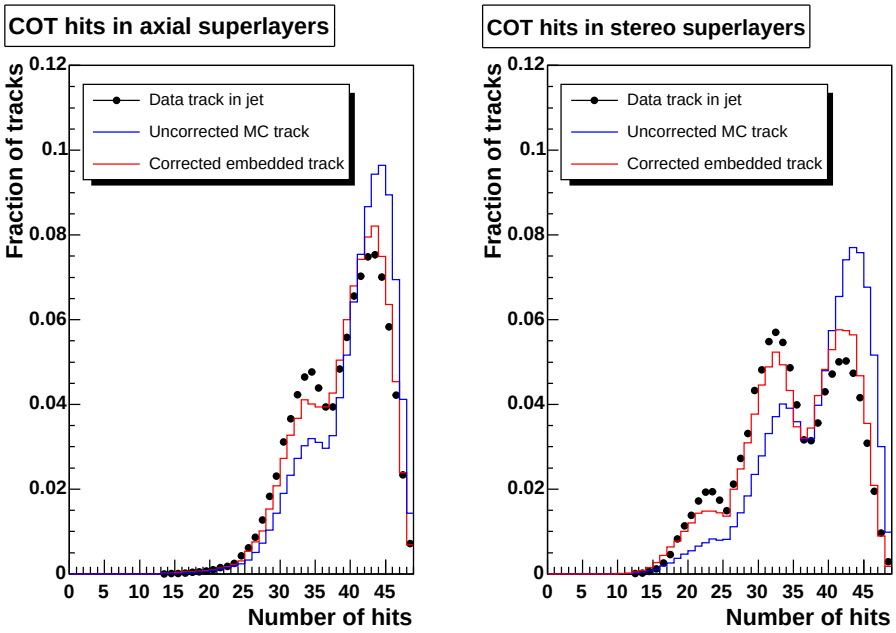

Figure 6.6: Sum of all hits in the four COT axial superlayers (left) and the four COT stereo superlayers (right) for tracks with 2 GeV/c $< p_t <$ 15 GeV/c in jets. We show data tracks, uncorrected tracks, and track for which the hit position resolution, the hit pulse width, the single hit efficiency, and the hit merging distance have been corrected.

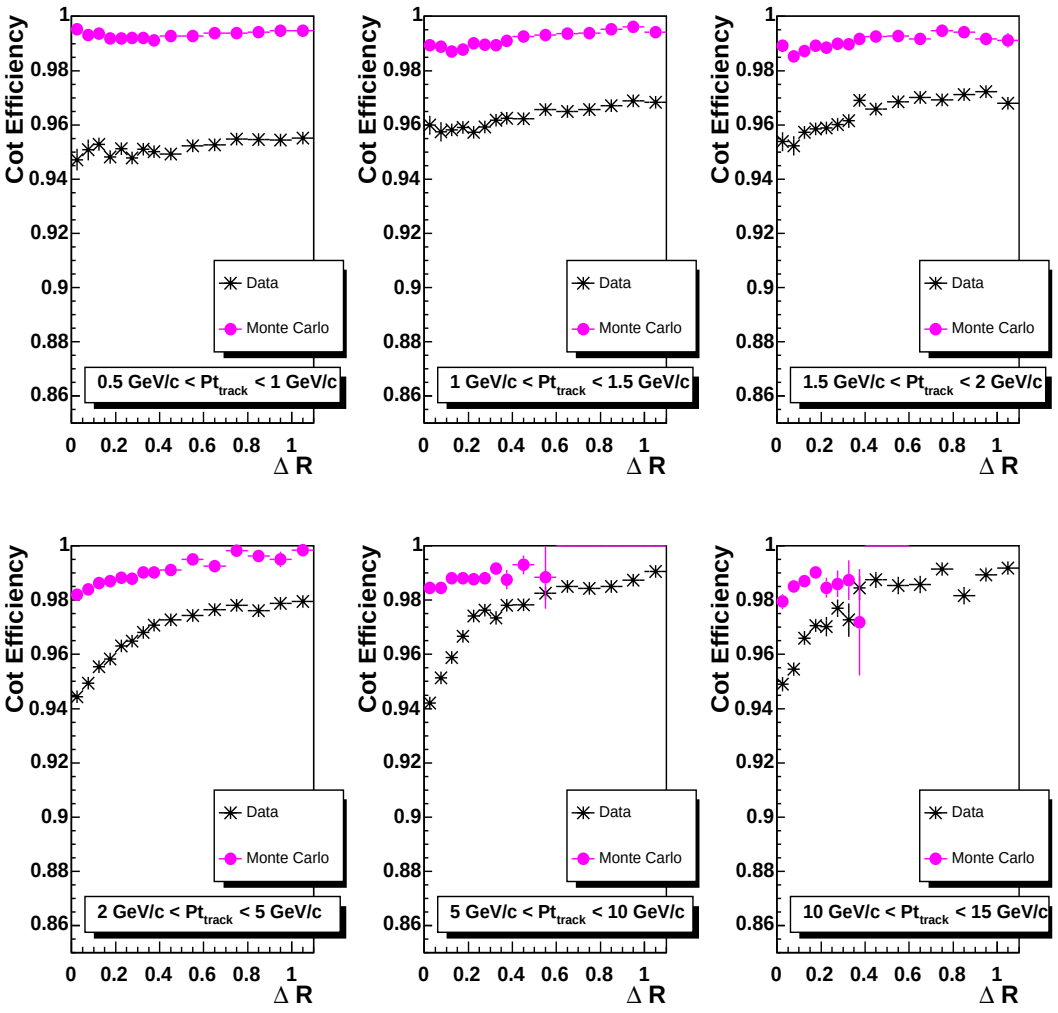

Figure 6.7: Tracking efficiency in jets with 20 GeV $< E_t <$ 30 GeV. $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi_0)^2}$ is the angular distance between the track and the jet centroid.

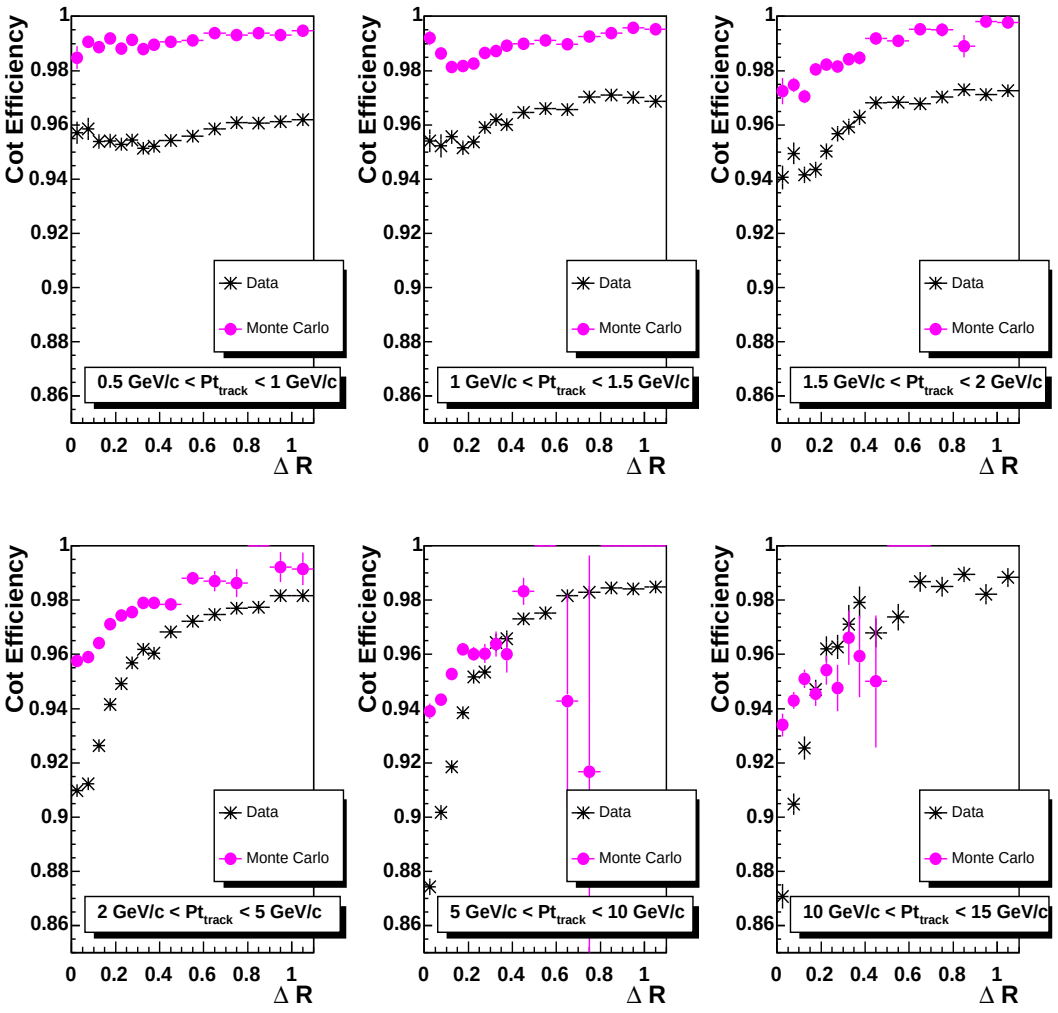

Figure 6.8: Tracking efficiency in jets with 60 GeV $< E_t <$ 70 GeV. $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi_0)^2}$ is the angular distance between the track and the jet centroid.

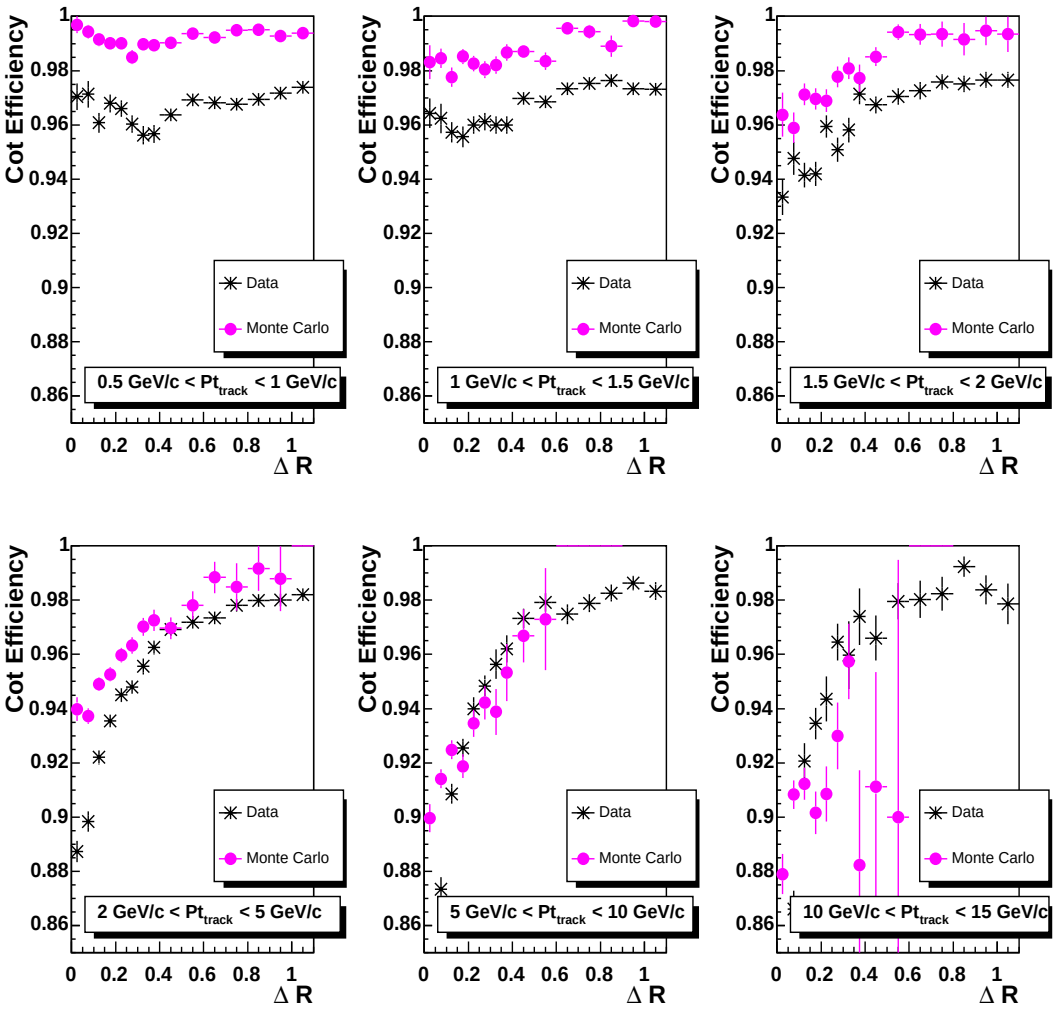

Figure 6.9: Tracking efficiency in jets with 100 GeV $< E_t <$ 120 GeV. $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi_0)^2}$ is the angular distance between the track and the jet centroid.

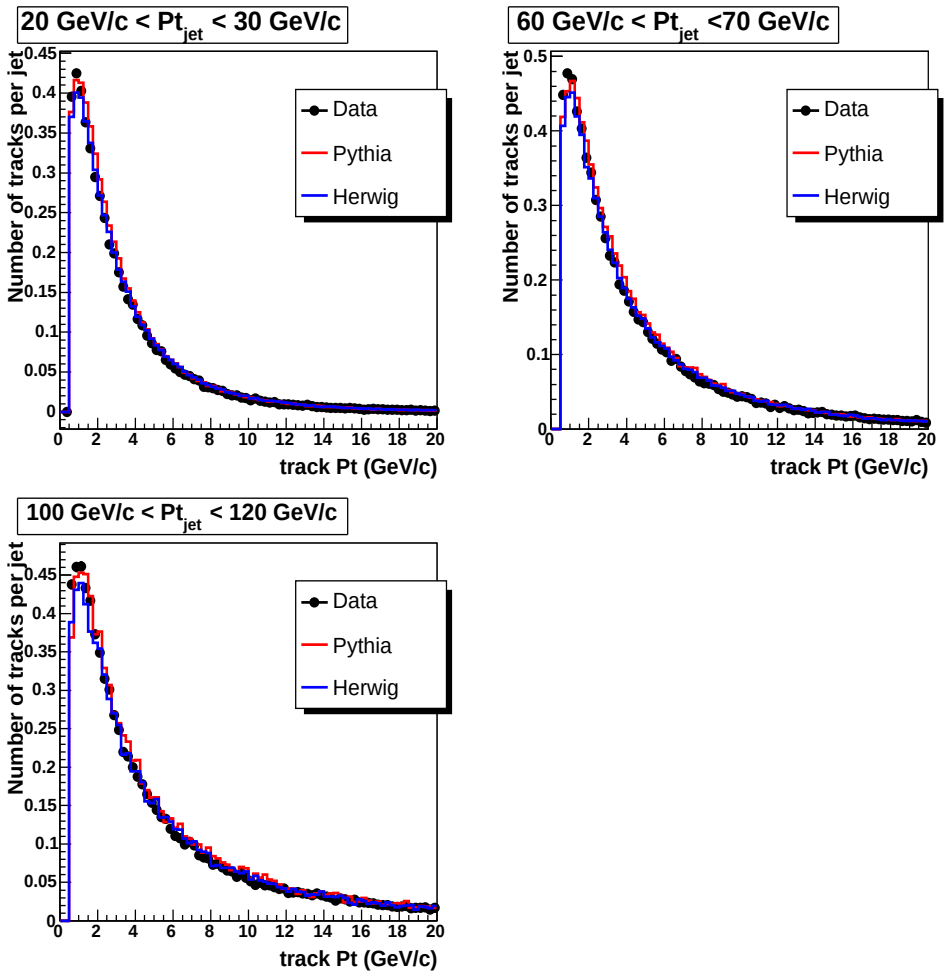

Figure 6.10: Transverse momentum distribution of charged particles in jets inside a cone of 0.4 for selected jet transverse momentum ranges

The absolute jet energy is corrected by accounting for the loss of energy due to the calorimeter response (see Section 3.3.2). The fragmentation uncertainty comes from the difference in loss of energy between MC and data for a measured and assumed calorimeter response E/p see Figure 6.1.

If we approximate that the calorimeter response to electromagnetic neutrals is 1.0, the energy loss is given by:

$$loss = \sum_{charged} p_t - \sum_{charged} p_t(E/p) \tag{6.1}$$

Relative to the jet momentum, we define the fractional loss as [25]:

$$fractional\,loss = \frac{\sum_{charged} p_t - \sum_{charged} p_t(E/p)}{CAL\,jet\,p_t} \tag{6.2}$$

where E/p is the calorimeter response for a track of momentum p and $CAL\,jet\,p_t$ is the jet momentum calculated from the calorimeter measurement. The jet fragmentation uncertainty is defined as the difference between the fractional loss in data and MC. Figure 6.11 shows the energy loss for data, PYTHIA and HERWIG. Figure 6.12 shows the fractional loss, as well as the difference between data and MC [26]. For jets with 20 GeV/c $< P_t <$ 140 GeV/c, we see an average difference below 1% for both PYTHIA and HERWIG.

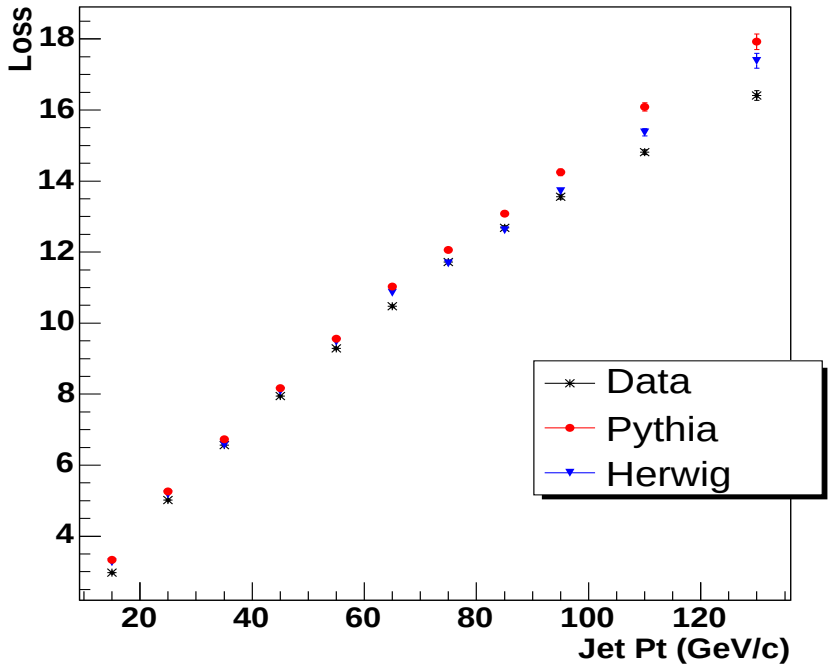

Figure 6.11: Absolute jet energy loss (GeV) as a function of the jet P_t measured in the calorimeter in data and Monte Carlo

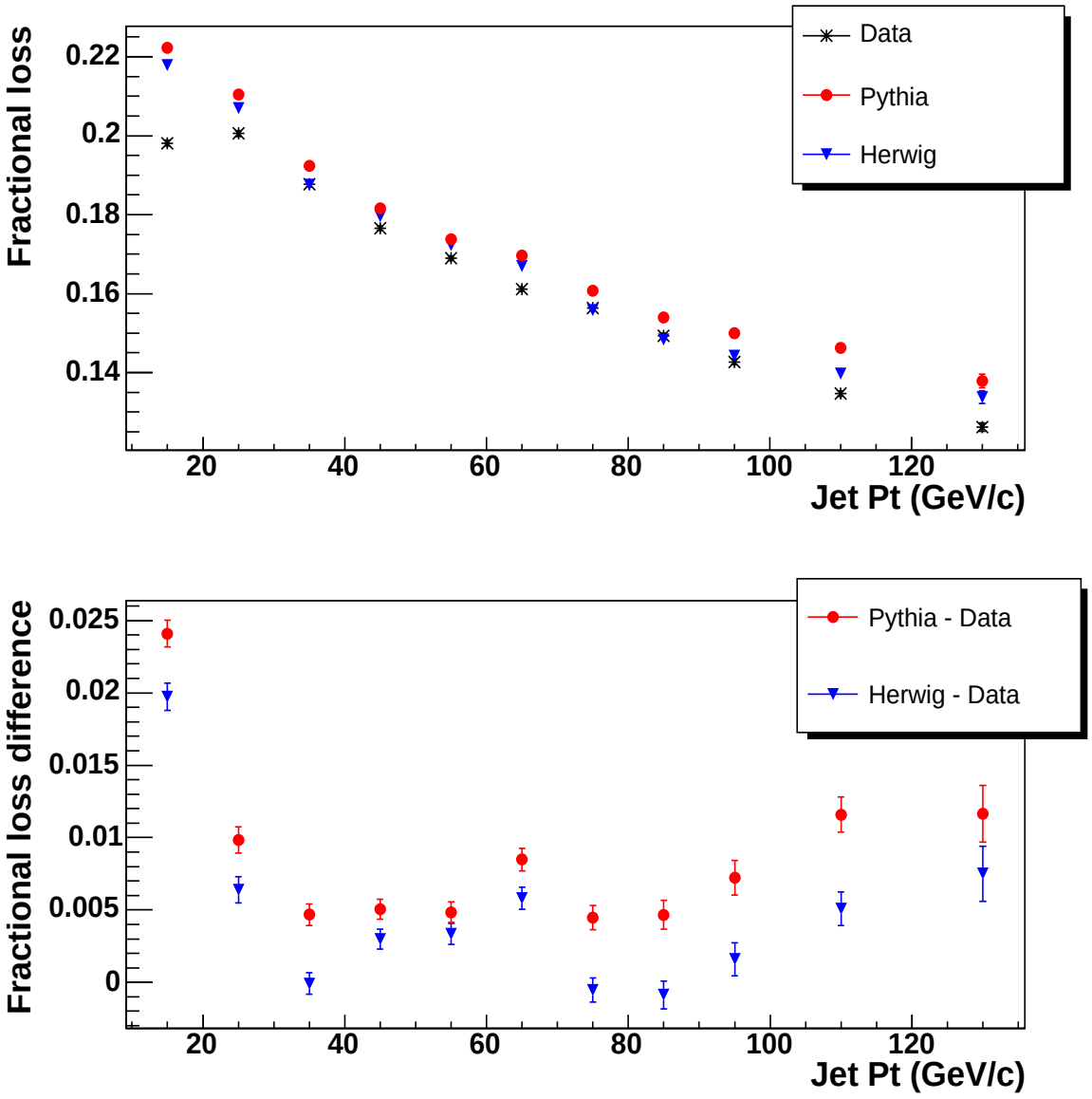

Figure 6.12: Fractional jet energy loss as a function of the jet P_t measured in the calorimeter in data and Monte Carlo (top) and difference between Monte Carlo and data (bottom). This difference is taken as the fragmentation jet energy systematic uncertainty.

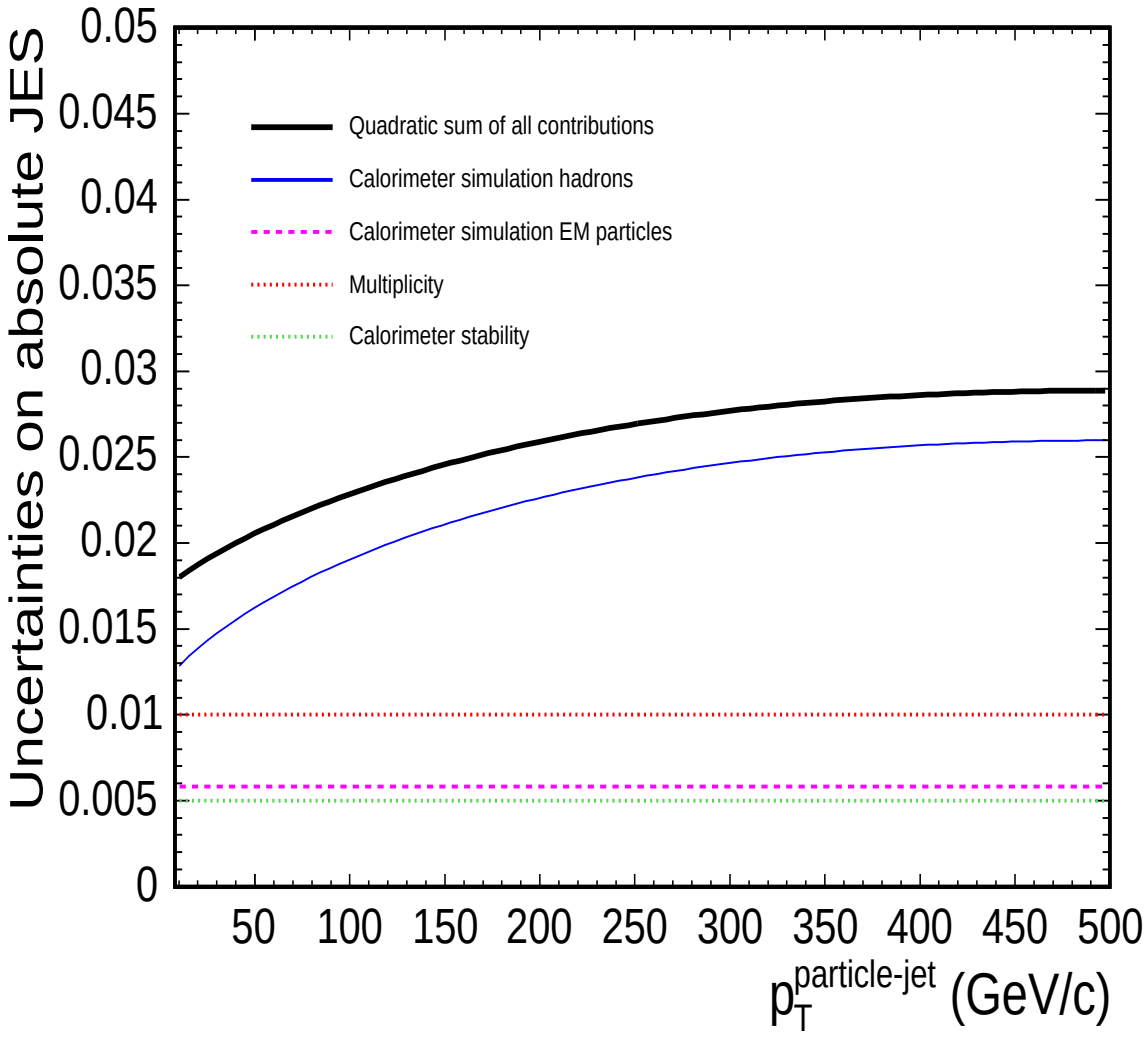

Figure 6.13: Summary of the absolute jet energy scale systematic uncertainties

6.1.4 Underlying event systematic uncertainty

In the same $p\bar{p}$ collision as the hard interaction of interest, interactions can occur between the spectator quarks. The products of these interactions can artificially contribute to the measured jet energy. We study the systematic difference between the underlying event in data and Monte Carlo jet events. We estimate the underlying event effect using all tracks with $p_t > 0.5$ GeV/c in the region transverse to the jet defined as $60^\circ < \Delta\phi(jet, track) < 120^\circ$. The measured average transverse momentum is normalized to the jet reconstruction cone size (0.4). The Monte Carlo is consistent with the data within 30% for all jet P_t. Therefore, we take 30% of the underlying event correction (see Section 3.3.2) as a systematic uncertainty.

6.1.5 Out-of-cone systematic uncertainty

Several effects can cause systematic differences in the quantity of energy falling outside of the reconstruction cone. For instance, the fragmentation and gluon radiation in the final state influence the fraction of the energy outside the cone. In addition, low p_t tracks can be bent away from the jet reconstruction cone. If any of these effects is different in data and Monte Carlo, it will cause a systematic uncertainty in the jet energy. We calculate the out-of-cone systematic uncertainty using events where we have a photon and a jet back-to-back in ϕ. The energy of photons is well measured in the calorimeter and is taken as a reference of the true energy of the balancing jet. We study the energy that falls in an annulus that surrounds the jet reconstruction cone, and that is limited by a cone of 1.3. The difference between data and Monte Carlo is shown in Figure 6.14 and taken as a systematic uncertainty.

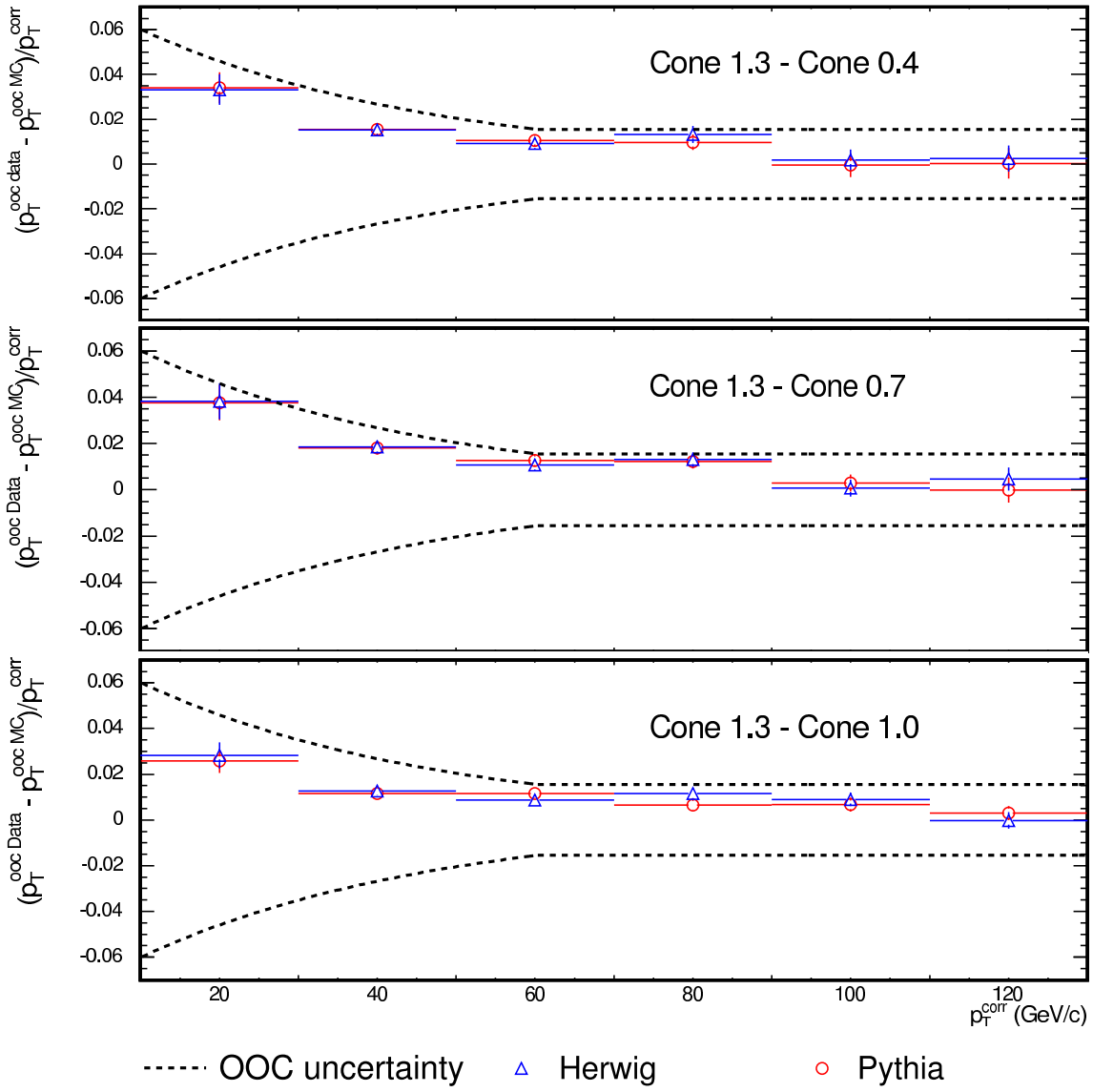

Figure 6.14: Out-of-cone systematic uncertainty for different jet reconstruction cone sizes. p_t^{cor} is the assumed true transverse momentum of the jet taken from the balancing photon energy measurement

6.1.6 Splash out systematic uncertainty

The out-of-cone energy is defined as the energy within $0.4 < \Delta R < 1.3$. It has been estimated in PYTHIA simulations that, from the jet, an additional 0.5 GeV is deflected outside the out-of-cone region. We define this energy as the "splash-out" energy and take half of this value, 0.25 GeV, as our systematic uncertainty.

6.1.7 b-jet modeling systematic uncertainty

Since the above energy corrections and corresponding systematics are obtained from generic (light-quark dominated) jets, additional uncertainties arise from extrapolating this procedure to b-jets. These are the three main sources of discrepancies between generic jet and b-jet energy [27, 28]:

- Color flow: energy transfer between quarks (0.3% effect).
- Fragmentation: due to b-quark mass, b-jets can fragment differently (0.4% effect).
- Semi-leptonic decays: b-quark decaying to a hadron, a charged lepton, and a neutrino (0.4% effect).

6.1.8 Jet energy scale systematic effect on the top quark mass

For a given jet, we calculate the variation in jet energy caused by a $+1\sigma$ and -1σ fluctuation of the jet energy scale. For a given systematic effect on the jet energy, we shift all the jets in an event in the same direction ($+1\sigma$ or -1σ), perform the event selections and apply the NWA algorithm to obtain a systematically shifted mass. This is done for the signal samples and for each background sample. We thus obtain positively and negatively shifted templates for both signal and background. While running the pseudo-experiments, signal and background events are selected from the shifted templates, but they are compared to the parameterization of the nominal (non-shifted) templates in the

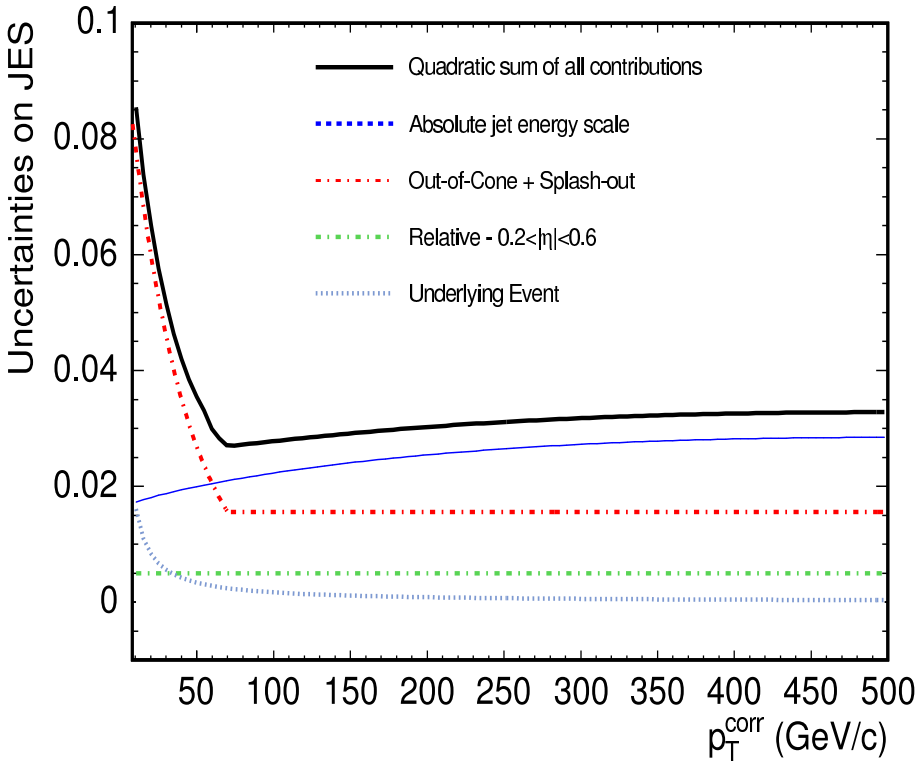

Figure 6.15: The total jet energy systematic uncertainty

likelihood function (Eq. 5.16). We take half the difference between the median of the positive and negative reconstructed mass distribution as our uncertainty. An alternative method is to add all the systematic effects on a given jet in quadrature and to apply them all at once.

Table 6.3 lists the systematic uncertainties on the top quark mass due to the $\pm 1\sigma$ shift of each jet systematic uncertainty effect. The result obtained when we apply all systematic effects at once is also shown. As seen in the Table, the sum of the individual effects is larger than the case where we apply all effects at once. The jet energy scale systematic uncertainties are uncorrelated and added in quadrature. We take 3.4 GeV/c^2 as our jet energy systematic.

Source	Systematic uncertainty on M_{top} (GeV/c^2)
All effects at once	2.6 ± 0.2
η relative response	0.6 ± 0.2
Multiple interactions	0.2 ± 0.2
Absolute jet energy scale	3.0 ± 0.2
Underlying event	0.2 ± 0.2
Out-of-cone energy	1.4 ± 0.2
Splash out energy	0.1 ± 0.2
b-jet energy	0.6 ± 0.3
total	3.4 ± 0.6

Table 6.3: Jet energy systematic uncertainties with the quadrature sum. Also shown is the effect on the top quark mass when we shift the jet energy for all effects at once.

6.2 Parton distribution function

The systematic error in the NWA measurement due to the uncertainty in Monte Carlo parton distribution functions (PDF) is estimated by applying an event re-weighting method [29]. In this method, variations in the parameterization of quark and gluon PDF's are tested using the same Monte Carlo sample by weighting each event according to its generated information, which avoids the need to generate large Monte Carlo samples for each variation. In our case, we apply this technique to the $t\bar{t}$ PYTHIA sample, and construct templates which are weighted according to different PDF fits:

- **CTEQ5L:** fit by CTEQ group [30], default PDF used in this analysis
  - **CTEQ6M:** $\pm 1\sigma$ variation of 20 CTEQ eigenvectors based on experimental uncertainties in the modeling of CTEQ
- **MRST72:** fitted by MRST group [31], with $\Lambda_{QCD} = 228$ MeV
- **MRST75:** fitted by MRST group, with $\Lambda_{QCD} = 300$ MeV

Half the difference between each CTEQ6M eigenvector $\pm 1\sigma$ variation (see Figure 6.16) is added in quadrature to determine an overall CTEQ PDF uncertainty. The difference between the CTEQ and MRST PDF groups, the difference between different MRST Λ_{QCD} settings, and the CTEQ PDF uncertainty are added in quadrature. This results in a total PDF systematic uncertainty of 0.4 ± 0.7 GeV/c^2, as shown in Table 6.4.

6.3 Gluon radiation

In $t\bar{t}$ events, the occurrence of high energy gluon radiation before the decay of the top quark (initial state radiation or ISR) or after the decay of the top quark (final state radiation or FSR) contributes to the shape of the signal templates. This is mainly due to the presence of additional jets, which, if energetic enough, can be confused with the

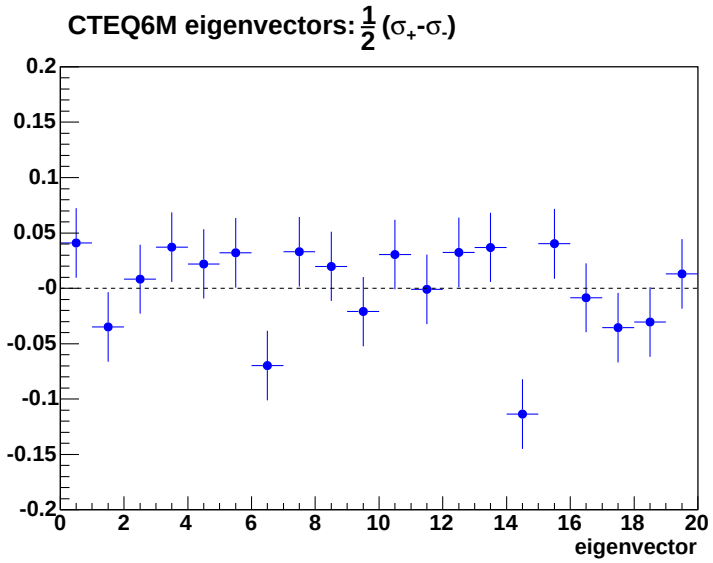

Figure 6.16: Half-differences between $\pm 1\sigma$ CTEQ6M eigenvector variations used in determining the PDF systematic uncertainty on the top quark mass.

Estimator	$\boldsymbol{\Delta m_{top}}$(GeV/c^2)
CTEQ6M ($\pm\sigma$)	0.2 $\pm$ 0.2
MRST72-CTEQ5L	0.2 $\pm$ 0.5
MRST75-MRST72	0.3 $\pm$ 0.5
total	0.4 $\pm$ 0.7

Table 6.4: Summary of the PDF systematic uncertainties on the top quark mass

mode	Δm_{top}(GeV/c^2)
"more" ISR	$+0.4 \pm 0.5$
"less" ISR	-0.8 ± 0.5
"more" FSR	-0.5 ± 0.5
"less" FSR	-0.2 ± 0.5

Table 6.5: Summary of the ISR and FSR systematics uncertainties on the top quark mass.

b-jet from the top decay. Therefore, a mis-modeling of the ISR or FSR can result in a systematically shifted template and cause a systematic uncertainty on the top quark mass.

To estimate the uncertainty due to the modeling of the ISR, we use Drell-Yan events where the (Z/γ) decays to a pair of charged electrons or muons. Since there is no jet in this final state, we can isolate the ISR effect from the FSR effect. The ISR activity will influence the p_t distribution of the charged leptons. Figure 6.17 shows the p_t distribution of the charged leptons for PYTHIA and data [32]. In PYTHIA, we vary the parameters of the DGLAP gluon radiation evolution function [33] such that the charged lepton p_t will be shifted by an amount corresponding to its uncertainty. Thus, we obtain an upper and lower limit on the ISR activity. The result is then extrapolated to the Q^2 of $t\bar{t}$ events. Using these shifted DGLAP parameters, $t\bar{t}$ samples with "more" or "less" ISR are created.

Since FSR is fundamentally the same process as ISR, we can estimate the FSR effect from the ISR effect. We vary the same DGLAP parameters on the final state portion of the process to create $t\bar{t}$ samples with "more" or "less" FSR.

While using the nominal signal and background templates in the likelihood, we perform pseudo-experiments drawing signal generated from the "more" or "less" ISR and

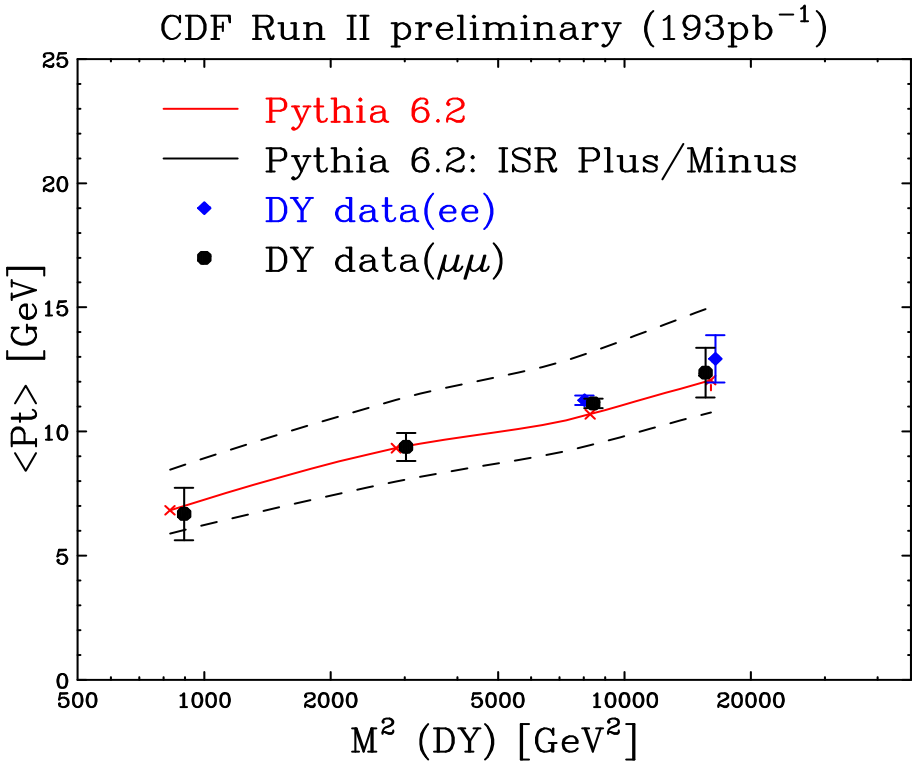

Figure 6.17: Average p_t of electrons and muons in data and Monte Carlo Drell-Yan events. The dashed curve shows the upper and lower limits of ISR activity and are used to create the systematically shifted templates.

FSR samples. The results from these pseudo-experiments are compared with the nominal PYTHIA sample (see Table 6.5). For the case of ISR, we take half the difference between "more" and "less", 0.6 GeV/c^2 as our systematic uncertainty. In the case of the FSR, since both the "more" and the "less" go in the same direction, we take the largest of the two, 0.5 GeV/c^2, as our systematic uncertainty.

6.4 Monte Carlo generator

HERWIG was used to generate the $t\bar{t}$ Monte Carlo signal. The jet fragmentation, the gluon radiation, and the underlying event are simulated at generator level. These effects have already been accounted for separately. However, as a cross check, we investigate here if a different choice of Monte Carlo generator would have resulted in a different top quark mass result. We compare the reconstructed top quark mass when we draw pseudo-experiments from signal templates generated with samples using PYTHIA and HERWIG. The mean reconstructed top quark masses differ by 0.1 ± 0.5 GeV/c^2. Therefore, we consider this uncertainty negligible.

6.5 Missing transverse energy resolution

If our knowledge of $\sigma(\not\!\!E_T)$ doesn't represent the data, the weights assigned to a certain event's configuration will be changed (see Eq. 5.9) and cause a systematic error. We perform pseudo-experiments by drawing events from templates constructed with half and double the nominal $\sigma(\not\!\!E_T) = 15$ GeV/c^2, and compare these to the nominal templates. Assuming that the true resolution can uniformly be anything between these values, we take the difference between the resulting values over $\sqrt{12}$. These extreme values should account for the possibility that $\sigma(\not\!\!E_T)$ is non-Gaussian. Comparing with the nominal HERWIG sample, we obtain a systematic uncertainty of 0.0 ± 0.2 GeV/c^2. Therefore, we consider this effect negligible.

6.6 Template statistics

Limited statistics in the signal and background templates result in a systematic uncertainty on the parameterizations used in the likelihood function (Eq. 5.16). For each signal template, we fluctuate the number of events in each bin following a Poisson distribution and create a new template. We then fit the probability density function to this "fluctuated" set of templates and perform 10 000 pseudo-experiments by drawing events from the nominal (non-fluctuated) template, repeating this procedure 100 times. We take the root mean square of this distribution as the systematic uncertainty due to the signal template statistics, giving 0.2 ± 0.2 GeV/c^2. We repeat the same procedure for each bin of each background component, and obtain a systematic uncertainty due to the limited background template statistics of 1.3 ± 0.2 GeV/c^2.

6.7 Background shape

Our background templates are all obtained using certain Monte Carlo generators and the CDF detector simulation. One systematic uncertainty associated with the total parameterized background template reflects the imperfect knowledge due to limited Monte Carlo statistics, as described above. Another source, discussed here, reflects our uncertainty in modeling the background shape, even in the limit of infinite statistics. We disentangle the background shape uncertainty into two sources. The first is the relative composition of each background source in the combined background template. The second is the mis-modeling of the kinematic quantities of the individual background samples.

6.7.1 Relative composition

This systematic comes from uncertainties in the numbers of background events from each source, as listed in Table 4.1. We base the systematic uncertainty on the top quark mass on variations of the two largest and most disparate backgrounds, the Drell-Yan

and "fakes". Indeed, in Figure 5.7, we see that the Drell-Yan template has a mean of $M_{top} = 165.6$ GeV/c^2 and the "fake" template has a mean of $M_{top} = 150.4$ GeV/c^2. We create two systematically shifted templates by varying the weight of the Drell-Yan and the "fake" contributions in opposite directions:

- Number of Drell-Yan $+1\sigma$ and number of "fakes" -1σ
- Number of Drell-Yan -1σ and number of "fakes" $+1\sigma$

We draw pseudo-experiments from the two templates. After likelihood minimization, we take half the difference between the systematically shifted templates and obtain a systematic uncertainty on the top quark mass of 2.2 GeV/c^2.

6.7.2 Drell-Yan $\not\!\!E_T$

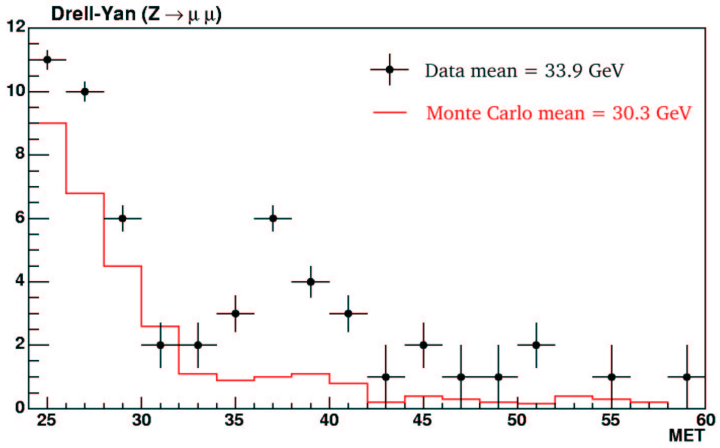

Figure 6.18: Comparison of data and Monte Carlo Drell-Yan $\not\!\!E_T$ after event selections. Monte Carlo is shown by the red line and data is shown by the black points.

The Drell-Yan events do not usually have high $\not\!\!E_T$, and only a mis-measurement of the $\not\!\!E_T$ can lead them to pass our selection requirements. An incorrect $\not\!\!E_T$ simulation (for example, if we do not properly simulate the mis-measurement of a jet in a calorimeter

crack) could lead to a significant change in the background shape. We estimate this mis-modeling from the difference in the means of the data and Monte Carlo Drell-Yan $\not\!\!E_T$ distributions (Figure 6.18) and obtain 12%. We therefore increase the $\not\!\!E_T$ by 12% for the Drell-Yan component of the background, and perform pseudo-experiments drawing events from this systematically shifted template. Comparing with the nominal HERWIG sample, we obtain a systematic uncertainty of 0.0 GeV/c^2. This shows that the systematic uncertainty on the background shape due to the effects of modeling the $\not\!\!E_T$ in Drell-Yan events is negligible.

6.7.3 Isolated track momentum in "fakes"

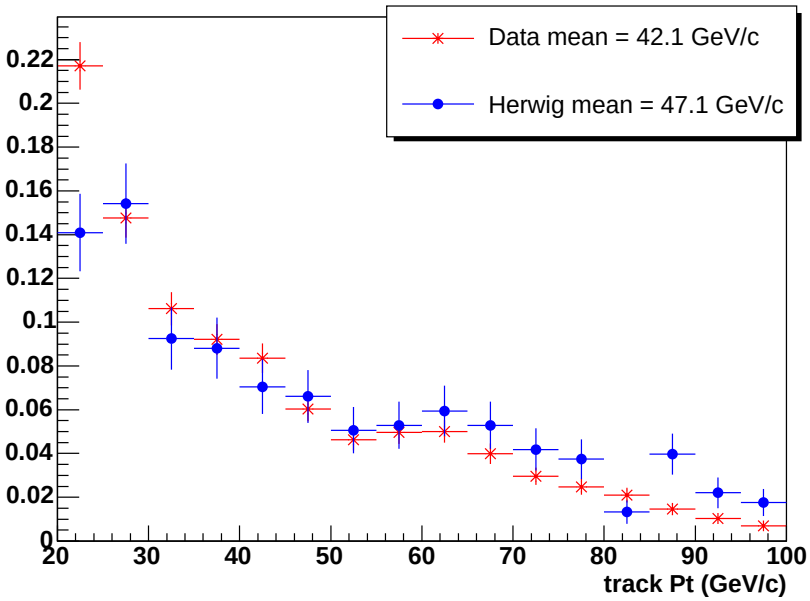

Figure 6.19: Comparison isolated track p_t in jet triggered data and HERWIG dijet events.

"Fakes" are events where a jet as been misidentified as an isolated track. Since we use only the tracking information to measure the p_t of the isolated track, "fake" tracks are likely to be mis-measured objects. We investigate the effect of this potential

mis-modeling in the $W \to \ell\nu$+partons "fakes" background. From a comparison of jet data and HERWIG (see Figure 6.19), we observe a difference in mean p_t of roughly -12% (data-MC). We then create a new "fake" background template with the p_t of the isolated track reduced by this amount, and perform pseudo-experiments drawing from this template and comparing them to the nominal templates. We measure a systematic uncertainty on the top quark mass of 0.0 GeV/c^2, which shows that this effect is negligible.

6.7.4 Jet energy mis-modeling

While the effect of jet energy mis-measurement on the top quark mass has already been estimated in Section 6.1, here we investigate how mis-modeling in the Monte Carlo simulation of jet energy can alter the shape of the background template parameterization. For "fakes", we compare the E_T of the leading three jets between inclusive electron data and the $W \to \ell + 3\,jets$ HERWIG (see Figure 6.20), which shows an overestimation in the Monte Carlo of roughly 5%. We create a "fakes" background template where we decrease the jet E_T by this amount, and perform pseudo-experiments drawing from this template. We measure a systematic uncertainty on the top quark mass of 0.5 GeV/c^2.

Next, we study the simulation of jet energy for Drell-Yan events by comparing our PYTHIA sample with inclusive electron data, where two identified leptons with an invariant mass within 15 GeV/c^2 of m_Z have been selected (along with the usual 2 jets). From Figure 6.21 we see that the simulation underestimates the data by about 7%. As above, we create new Drell-Yan templates with this boosted jet energy and perform pseudo-experiments, leading to a systematic uncertainty on top quark mass of 1.3 GeV/c^2.

Since the diboson background is also simulated with PYTHIA, we assume the same 7% uncertainty, and find a systematic error on the diboson background shape due to jet energy mis-modeling of 0.1 GeV/c^2.

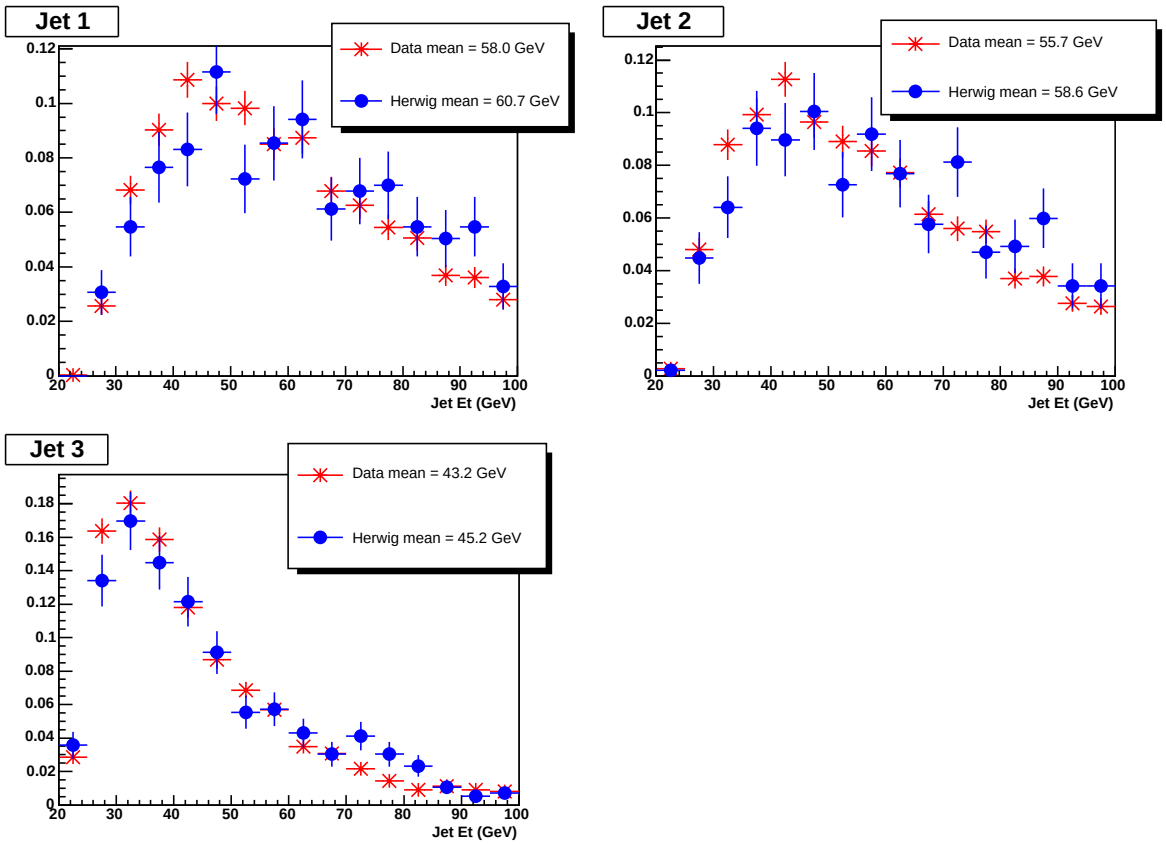

Figure 6.20: Comparison of $W \rightarrow \ell + 3jets$ HERWIG jet energy with that from inclusive electron data.

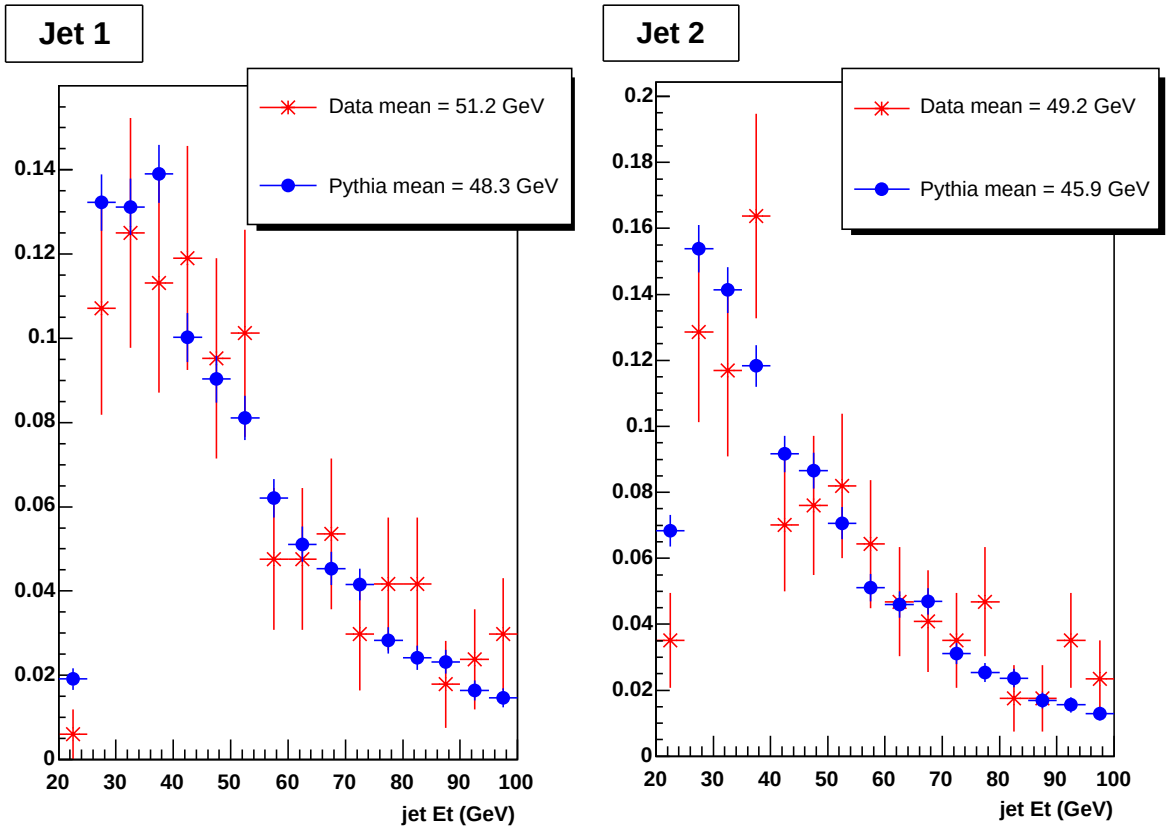

Figure 6.21: Comparison of jet energy from Drell-Yan ($Z \rightarrow ee$) PYTHIA with that of Z-decays selected from inclusive electron data.

6.7.5 Total background shape systematic

Combining the above systematic uncertainties on the background shape due to errors in relative composition and mis-modeling of the $\not\!\!E_T$ in DY, the isolated track momentum in "fakes", and the jet energy in all background samples, we arrive at a total background shape systematic uncertainty on the top quark mass of 2.6 GeV/c^2.

6.8 Summary

Source	Systematic uncertainty
Jet energy scale	3.4
PDF's	0.4
ISR	0.6
FSR	0.5
Background templates statistics	1.3
Background shape	2.6
total	4.6

Table 6.6: Summary of systematic uncertainties (in GeV/c^2) on the top quark mass.

Table 6.6 summarizes the systematic uncertainties on the top quark mass measurement. We estimate our total systematic uncertainty to be 4.6 GeV/c^2. Uncertainties of 0.2 GeV/c^2 and lower are considered negligible and were discarded.

Chapter 7

Conclusion

We measured the top quark mass in the dilepton channel. The event selection was based on the identification of a charged lepton and a well defined track. The technique used, the Neutrino Weighting Algorithm, consisted of weighting the direction of the two neutrinos according to how well it corresponds to the missing transverse energy. The data events were compared to a set of signal and background Monte Carlo templates using a likelihood fit. Through pseudo-experiment tests, we have shown that the method is consistent and robust. We obtained:

$$M_{top} = (170.8^{+6.9}_{-6.5}\,(stat) \pm 4.6\,(syst))\mathrm{GeV/c^2} \tag{7.1}$$

Figure 7.1 shows a summary of the top quark mass measurements made at the Tevatron. The dilepton Neutrino Weighting Algorithm result provides an important contribution to this collective effort of improving our knowledge of the top quark mass. At CDF, three other top quark mass measurements have been made in the $t\bar{t}$ dilepton channel using different techniques. All four results have been combined, yielding a result of: $M_{top} = 168.3 \pm 5.3\,(stat) \pm 3.3\,(syst)$ GeV/c^2 [37]. All the top quark mass measurements in CDF and D0 in all channels have been combined, yielding a result of: $M_{top} = 172.5 \pm 1.3\,(stat) \pm 1.9\,(syst)$ GeV/c^2.

Figure 7.2 shows the constraints on the Higgs boson mass from the Tevatron's top quark mass measurement and the LEP2 W boson mass measurements. The corresponding Higgs boson mass prediction has been constrained to: [36]:

$$M_{Higgs} = 91^{+45}_{-32}\,\mathrm{GeV/c^2} \tag{7.2}$$

$$M_{Higgs} < 186\,\mathrm{GeV/c^2}\ \ (95\%\,\mathrm{confidence\,level}) \tag{7.3}$$

Combined with exclusion region from the LEP experiments which tells us that $M_{Higgs} > 114\,\mathrm{GeV/c^2}$ (see Figure 7.2), the Higgs boson mass prediction is much narrowed down by the recent top quark mass measurements.

A parallel effort was also made to measure other properties of the top quark. In a separate analysis, using the same data events, the $t\bar{t}$ production cross section has been measured to be: $\sigma_{t\bar{t}} = 9.9 \pm 2.1\,(stat) \pm 1.3\,(syst) \pm 0.6\,(lum)$ pb [17].

Today, it is believed that we are a few years away from an era of major scientific breakthrough. The LHC will allow us to probe the subatomic world at higher energies than ever, with the highest luminosities ever attained for a hadronic collider. The experimental confirmation of the existence of the Higgs boson would change our understanding of the origin of mass. The work presented in this thesis can be seen as preparatory to the LHC experiment. If the Higgs boson is observed and its mass measured at the LHC, the constraint on its mass will provide a consistency check of the Standard Model. Our improvement of the measurement of the top quark mass has further tightened this constraint. In addition, top quark production is predicted to be an important background process at the LHC, and our increasing knowledge of the top quark will help the LHC analyses.

Meanwhile, until the LHC begins taking data in a few years, the Tevatron remains the only collider in the world capable of producing top quarks in pairs. With an increasing integrated luminosity, the statistical uncertainty on the top quark mass will be greatly reduced. Additional work will be required to reduce the systematic uncertainties further.

Notably, a more precise tuning of the jet simulation will allow us to reduce the jet energy scale systematic uncertainty. In addition, the identification of b-jets will allow us to significantly reduce the background and its associated systematic uncertainties.

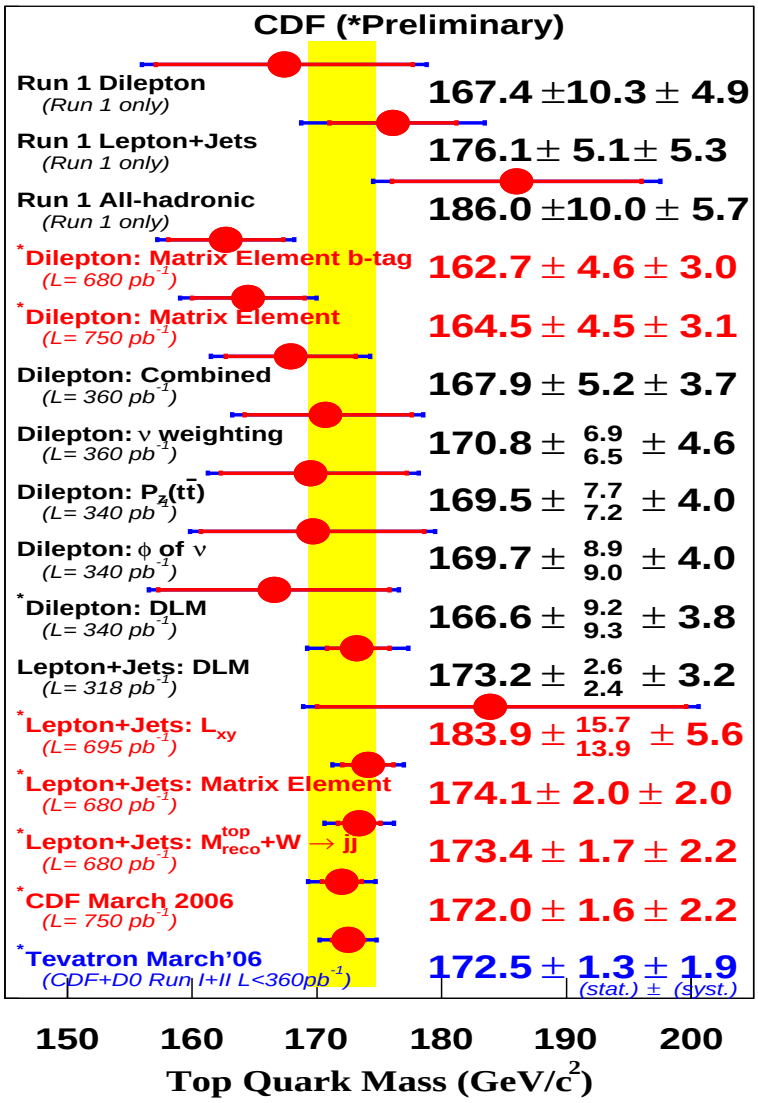

Figure 7.1: Summary of the CDF top quark mass measurements. The Tevatron (CDF and D0) combined result is also shown (blue)

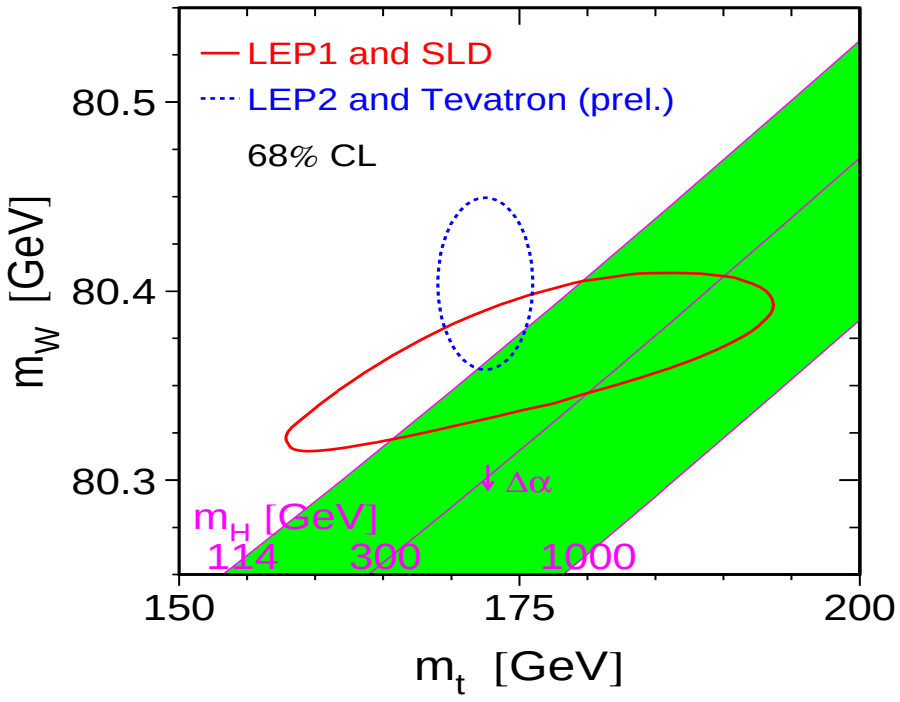

Figure 7.2: Constraints on the Higgs boson mass from the top quark and W boson mass measurements in CDF, D0 Run I and the LEP2 experiment (dashed blue line). Constraints from studies at the Z boson pole (red line) and from a direct search at LEP (green bands) are also shown.

Bibliography

[1] F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).

[2] S. Abashi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[3] I Bigi *et al.*,Phys. Lett. B **181**, 157 (1986).

[4] LEP Collaboration *et al.*, hep-ex/0312023 (2003).

[5] F. Abe *et al.*, Phys. Rev. Lett. **82**, 271 Erratum-ibid. **82**, 2808 (1999).

[6] R. Blair *et al.*, FERMILAB-PUB-96-390-E (1996).

[7] G. Corcella *et al.*, JHEP **0101**, 010 (2001).

[8] T. Sjostrand *et al.*, hep-ph/0108264 (2001).

[9] M. Mangano *et al.*, JHEP **0307**, 001 (2003).

[10] R. Brun *et al.*, CERN-DD-78-2 (1978).

[11] G. Grindhammer *et al.*, Nucl. Instrum. Meth. A **290**, 469 (1990).

[12] F. Abe *et al.*, Phys. Rev. D **45**, 1448 (1992).

[13] A. Bhatti, *et al.*, "Determination of the Jet Energy Scale at CDF", Nucl. Instrum. Meth. hep-ex/0510047 (2005).

[14] C. Hill *et al.*, "Electron Identification in Offline Release 5.3", CDF Note 7309 (2005).

[15] V. Martin *et al.*, "High-p_t Muon ID Cuts and Efficiencies for use with 5.3.1 data and 5.3.3 MC", CDF Note 7367 (2004).

[16] R. Bonciani *et al.*, Nucl.Phys.B **529**, 424 (1998).

[17] C. Hill *et al.*, "A Measurement of the Top Dilepton Cross-section Using The Tight Lepton and Isolated Track Sample", CDF Note 7679 (2005).

[18] D. Acosta *et al.*, Phys. Rev. Lett. **93**, 142001 (2004).

[19] T. Affolder *et al.* Phys. Rev. D **63**, 032003 (2001).

[20] K. Kordas *et al.*, "Gen5 Measurement of the Top Dilepton Mass using Neutrino Weighting", CDF Note 7548 (2005).

[21] J. Heinrich, private communication (2004).

[22] D. L. Nordstrom *et al.*, Phys. Rev. D. **66**, 010001-282 (2002).

[23] G. Arfken and H. Weber. "Mathematical Methods for Physicists". Harcourt/Academic Press (2000).

[24] S. Sabik, P. Savard, "Track reconstruction efficiency in jets", CDF Note 6894 (2004).

[25] A. Bhatti, private communication (2005).

[26] S. Sabik, P. Savard, "Measurement of the track multiplicity, momentum distribution and fragmentation uncertainty in jets", CDF Note 7346 (2004).

[27] P.K Sinervo, S. Xie, "The Effects of Colour Flow on B Jet Energies in Top Quark Events", CDF Note 7238 (2004).

[28] J.-F. Arguin, P.K Sinervo, "b-jets Energy Scale Uncertainty From Existing Experimental Constraints", CDF Note 7252 (2004).

[29] O. Gonzalez and C. Rott, "Uncertainties due to the PDFs for the gluino-sbottom search", CDF Note 7051 (2004).

[30] H.L. Lai *et al.*, Phys. Rev. D **51**, 4763 (1995).

[31] A. Martin *et al.*, Eur. Phys. J. C **C23:**, 73 (2002).

[32] Y. Kim U. Yang, "Initial state gluon radiation studies on Drell-Yan data for top-pair production in hadron collider", CDF Note 6804 (2004).

[33] V. Gribov, L.v Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).

[34] C. Mills, private communication (2005).

[35] J. Goldstein *et al.*, "Reconstructing the Plug Electron Energy in 5.3.3", CDF Note 7687 (2005).

[36] The LEP Electroweak Working Group,http://lepewwg.web.cern.ch/LEPEWWG/, to be published (2005).

[37] A. Abulencia, *et al.*, "Top Quark Mass Measurement from Dilepton Events at CDF II", to be submitted to Phys Rev. Lett. (2005).