THE UNIVERSITY OF CHICAGO

MEASUREMENT OF THE TOP QUARK MASS SIMULTANEOUSLY IN DILEPTON AND LEPTON + JETS DECAY CHANNELS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
WOJCIECH T. FEDORKO

CHICAGO, ILLINOIS
DECEMBER 2008

ACKNOWLEDGEMENTS

This is the hardest section to write as I am sure to forget someone important. Whomever you may be: the fact that the part of my brain, that was responsible for holding your representation, recently died form the lack of sleep, does not mean that you are not appreciated.

The inestimable Young-Kee is surely a proof positive of invalidity of what we thought were unbreakable physical laws. You know- the ones about speed of light travel, causality etc. Young-Kee, thank you for always spurring me on with a healthy dose of encouragement and great advice. I'm sure the time and energy has been borrowed from some parallel universe.

Erik Brubaker is a fantastic teacher - somehow never getting annoyed at a constant string of my stupid questions and misconceptions. Thank you for leading our analysis group and getting several results and the publication out the door, defusing the inevitable crises right before the blessing notes were due.

I feel very lucky that Jahred Adelman and I were at UofC at the same time. It was great to learn from you and work with you. I hope we will continue to do something fun together using even bigger toys, and I can't wait to have some bière with you on the other side of the pond.

It is really impressive how quickly Hyunsu Lee has plugged himself into out motley crew, navigated all the million lines of our incomprehensible code and started doing great work quietly and efficiently. Thank you Hyunsu for helping out so many times. I am really glad that you will lead this analysis and I am sure it will continue to be fun (if touch frustrating at times). I am also sure it will be improved under your lead.

I'm pretty certain that Mel Shochet just knows everything. Mel, thank you for your patient advice and guidance. This I hope will continue, because even though I learned so much from you, I still feel like I know, well not much.

Henry Frish can always inject some under or overstated humor into any situation. Henry, your perspective, on physics problems and otherwise has been worth more than... osmium.

Vadim "can you do it in an hour?" Rusu- getting Level 2 to work with you was a blast. I am still wondering why the end of store always comes at 3 am though. I have my money on you finding the Higgs faster than the LHC (if it's there).

I would also like to thank the whole Level 2 crew: Kristian, Sakari, Shawn, Ted, Daniel, Burkard, Cheng-Ju, Chris and Konstantin. No one ever got annoyed at me when I called their cell phone in the middle of the night, when I couldn't fix the thing myself. I am actually still surprised something didn't catch fire and that it is actually taking data.

Un-Ki Yang lead the top mass template group when I arrived and he continued to help out throughout several preliminary results. We still call him up at Manchester in a pinch. Thanks for always having some clever trick up your sleeve Un-Ki and I hope that you'll continue to share these.

I hope I wasn't too much of a disappointment for Steve Levy when we tried the loose tagger (and it took way too long to find out it wouldn't do any good for top mass). Steve thanks for putting up with me when I was totally green.

Thanks go to all members of the UChicago CDF group, past and present: Carla, Dan, Ivan, Kohei, Colin, Sasha, Satomi, Scott, Wes and Zaid. The pool of talent in our group is amazing. There is always someone that knows how to do the thing you need or to give you criticism on your work.

Pekka- thank you for introducing me to High Energy Physics, it has really been a lot of fun so far and it looks like more is coming. Thank you also and for continuing to give advice and support.

My classmates and friends have kept my sanity mostly intact (but then how am I supposed to judge?). I will remember fondly not only parties and outings to jazz clubs and bars but also QFT and solid state problem sets and studying for the qualifier. Thank you Robert, Robert, Matthew, Dimitris, Andy, Iggy, Raphael, Martina, Vassilis.

When things were not that great the people I know I could always count on, unconditionally, are my family. Thank you Mom, Dad, Peggy and John.

ABSTRACT

We present the first measurement of the top quark mass using simultaneously data from two decay channels. We use a data sample of $\sqrt{s} = 1.96$ TeV collisions with integrated luminosity of 1.9 fb^{-1} collected by the CDF II detector. We select dilepton and lepton + jets channel decays of $t\bar{t}$ pairs and reconstruct two observables in each topology. We use non-parametric techniques to derive probability density functions from simulated signal and background samples. The observables are the reconstructed top quark mass and the scalar sum of transverse energy of the event in the dilepton topology and the reconstructed top quark mass and the invariant mass of jets from the W boson decay in lepton + jets channel. We perform a simultaneous fit for the top quark mass and the jet energy scale which is constrained *in situ* by the hadronic W boson resonance from the lepton + jets channel. Using 144 dilepton candidate events and 332 lepton + jets candidate events we measure:
$\text{M}_{\text{top}} = 171.9 \pm 1.7$ (stat. + JES) ± 1.1 (other syst.) GeV/c^2=171.9 ± 2.0 GeV/c^2
The measurement features a robust treatment of the systematic uncertainties, correlated between the two channels and develops techniques for a future top quark mass measurement simultaneously in all decay channels. Measurements of the W boson mass and the top quark mass provide a constraint on the mass of the yet unobserved Higgs boson. The Higgs boson mass implied by measurement presented here is higher than Higgs boson mass implied by previously published, most precise CDF measurements of the top quark mass in lepton + jets and dilepton channels separately.

TABLE OF CONTENTS

Chapter

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Often very simple questions that we normally neglect to ask can yield surprising insights. The question that we want to address here is "Why do things have mass?". This question is naive only at a first sight. Answering it would imply that we understand the mechanism of unification of the electromagnetic and weak interactions akin to Maxwell's unification of electricity and magnetism.

The problem of creation of mass needs to be approached from different angles and the top quark is a very important piece of the puzzle. In this chapter we explore the role of the top quark in the Standard Model and how knowledge of the top quark mass helps predict the mass of the Higgs boson - a putative particle at the center of the electroweak symmetry breaking mechanism. Putting constraints on the Higgs boson mass is important to the experimental particle physics community searching for the Higgs boson. This is because the Higgs boson production modes and decay channels depend on it's mass. If the Higgs boson is indeed found knowledge of the top quark mass and other electroweak observables will help discern if the observed particle fits within the Standard Model predictions or if the Standard Model needs to be extended.

We also examine some hints that the top quark may be giving us about what questions to ask next. Future particle physics experiments will try to address the so called "hierarchy problem" of the Standard Model. Measurements of the top quark mass and the W boson mass are constraining the theories constructed to solve this problem.

Finally we outline and motivate the methodology adopted here to measure the top quark mass. In contrast to the traditional averaging approach, the top quark mass measurement simultaneously in two decay channels does not need to assume the values of correlations in systematic uncertainties or the form of the likelihood in either of the channels being combined. The methods developed here will be directly applicable in a future top quark mass measurement simultaneously in three decay channels, with a robust treatment of correlated systematic uncertainties. A correct treatment of the systematic uncertainties becomes increasingly important as the systematic uncertainties start to dominate the statistical uncertainties on the top quark mass measurements using the newly available large Tevatron data sets.

1.1 Top Quark Within Standard Model

The Standard Model of particle physics is the latest step in the reductionist march of science. The foundations of the Standard Model have first emerged as an answer to the plethora of hadrons observed in the particle physics experiments of the 50' and 60'. The model describes the constituents of matter. Incorporation of gravity into the theory has proven elusive, but electromagnetic and nuclear forces are well described. In the model we have three generations of fermions: the quarks:

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$

and leptons:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$$

The symmetries associated with electromagnetic and weak interactions are broken through inclusion of a Higgs field into the theory. The Higgs acquires a vacuum expectation value (VEV) transforming the massless gauge bosons of weak and electromagnetic interactions into massive $W^{\pm}$, Z^0 boson and a massless photon - force carriers for the unified electroweak interaction. The strong interaction is mediated by a gluon and affects quarks but not the leptons.

It is not possible to add naive mass terms to the Lagrangian of the theory such as $m_u \bar{u}_L u_R$ since they break the gauge invariance of the Lagrangian. Fermion masses must be generated through some other mechanism and the electroweak symmetry breaking is again thought to be responsible. The interaction of the Higgs field with the fermions is allowed under gauge invariance of the theory. When Higgs acquires the VEV the interaction terms become mass terms with masses of the quarks and leptons proportional to the coupling constants and the VEV. This mechanism fits very well with many experimental results. The one crucial piece missing is observation of the Higgs boson.

In this picture the top quark is not particularly distinguished from other quarks. This is because masses of fermions are proportional to fermion couplings to the Higgs boson - these couplings are free parameters of the theory. The mass of the top quark turns out to be extraordinarily large: $\mathrm{M_{top}} \sim 170$ GeV/c^2. The next heaviest quark, the b, has mass of only 4 GeV/c^2. Such high mass was the reason why the top quark was discovered only just over a decade ago at CDF and DØ detectors [1]. In the next section we will see why the high mass of the top quark makes it particularly interesting to study.

1.2 Impact of the Top Quark Mass Measurement

The mechanism of fermion and boson mass generation outlined above is compelling, but it has yet to be fully confirmed. A huge effort in experimental particle physics today is directed at a search for the Higgs boson, which has to be observed for this explanation of creation of mass to make sense. Moreover the mass of the Higgs has to satisfy indirect constraints imposed by the precision electroweak measurements. Here is where the mass of the top quark enters.

The mass of the W boson obtained directly from the Higgs mechanism is:

$$M_{W0} = \sqrt{\frac{\pi\alpha}{\sqrt{2}G_F sin^2(\theta_w)}} \tag{1.2.1}$$

In the equation above G_F is the Fermi constant and θ_w is the Weinberg angle. There are higher order corrections modifying the W boson mass as:

$$M_W = M_{W0} \times \frac{1}{\sqrt{1-\Delta r}} \tag{1.2.2}$$

The quantity Δr depends on the Higgs boson mass as well as the top quark mass. The corrections can be computed from Feynman diagrams shown in Figure 1.1. The contribution to Δr from the Higgs boson depends weakly on it's mass: $\Delta r_{Higgs} \propto -log(M_{Higgs})$. The dependence on the top quark mass is much stronger: $\Delta r_{top} \propto \mathrm{M}_{\mathrm{top}}^2$ [2]. We can invert the relations and knowing the experimental values of the top quark mass obtain a constraint on the Higgs boson mass as shown in Fig. 1.2 [3]. This assumes that the Standard Model is correct. In the diagram the dotted ellipse indicates the CDF Run II W boson mass measurement [4] and the most

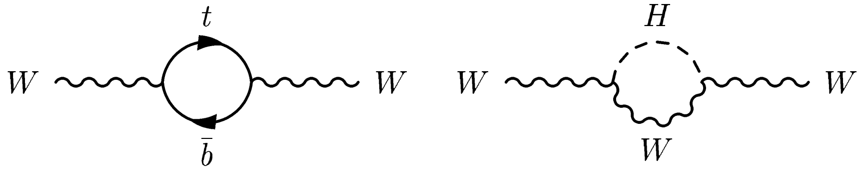


Figure 1.1: Feynman diagrams for loop corrections to W boson mass

precise dilepton channel top quark mass submitted for publication [5]. The dashed ellipse indicates the most precise published top quark mass measurement in the lepton + jets channel [6] and the same W boson mass measurement as in the dotted ellipse. The diagonal curves indicate a constant value of the Higgs boson mass. The LEP experiments have excluded the Higgs boson with mass lower than 114 GeV/c^2 [7] corresponding to the highest diagonal curve in the figure. The diagram demonstrates the need for further improvements in both the W boson mass and the top quark mass measurements so that an indirect, precise constraint on the Higgs boson mass can be obtained. Such a constraint is a powerful hint as to where to look for the Higgs boson. The Higgs boson production and decay branching fractions vary as a function of it's mass so it is important to design searches sensitive in the low mass region. In addition as apparent from Fig. 1.2 some tension starts to appear between the preferred Higgs boson mass and the direct search exclusion region. If the Higgs boson is eventually discovered comparison of it's mass to the one preferred by other electroweak observables will help to discern if in fact it's nature is as assumed in the Standard Model.

The LEP Electroweak Working Group taking into account the most recent preliminary results on the W boson mass [2] and the top quark mass [8] as well as measurements of other electroweak variables computes a 95% confidence level upper limit on the Higgs boson mass to be 154 GeV/c^2 and 185 GeV/c^2 when the exclusion from the direct LEP search is taken into account

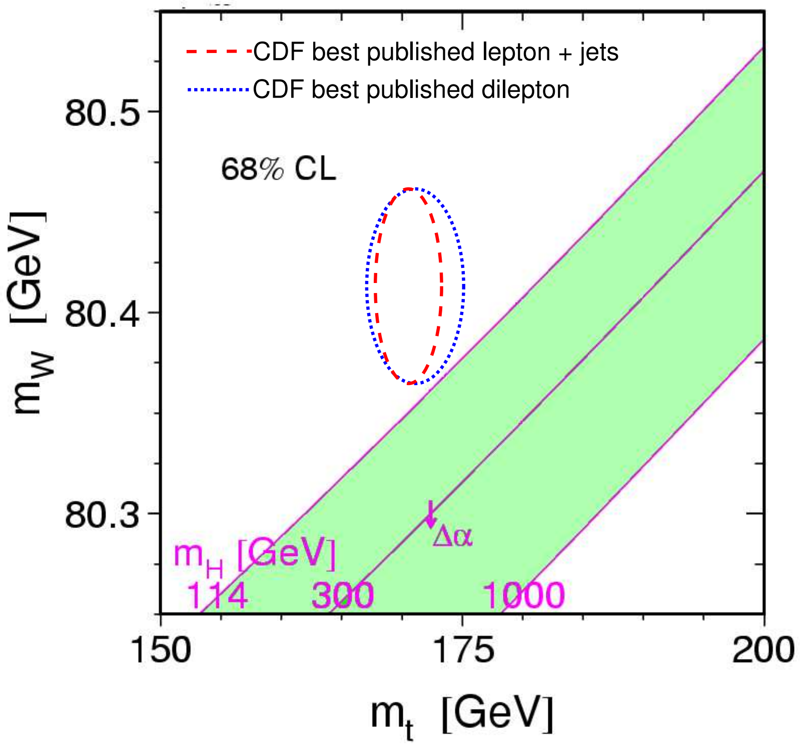


Figure 1.2: Dependence of the Higgs boson mass on the top quark mass and the W boson mass. Adapted from [3]. Ellipses indicate CDF II W boson mass measurement [4] and most precise CDF II (published or submitted) dilepton [5] (dotted) and lepton + jets [6] (dashed) top quark mass results. One dimensional 68% confidence levels are indicated by the ellipses.

The top quark mass measurements may be offering us clues as to what might lie beyond the Standard Model. The Standard Model has had a remarkable experimental success, however it requires very precise tuning in the loop corrections to the Higgs boson mass. This is the so called "hierarchy problem". One very elegant solution proposed to address this is introduction of "supersymmetry" - symmetry between fermions and bosons. This requires existence of supersymmetric partners to the particles observed to-date. If this model is true then the hierarchy problem is solved by natural cancellations of the one loop corrections to the Higgs boson mass. The top quark mass and the W boson mass may help us to constrain the parameter space of such a model or rule it out. Shown in Figure 1.3 (taken from References [9, 10]) are again the contours corresponding to the measured W boson and top quark masses. Assuming the Standard Model is correct authors scan over the value of the top quark mass and the Higgs boson mass solving for the W boson mass at each point. This generates the lower shaded region. The upper shaded region is obtained by scanning the top quark mass, generating random parameters of the Minimally Supersymmetric Standard Model (MSSM) and solving for the W boson mass. The two regions overlap in the shaded band in the middle. According to this analysis the current published measurements of the masses of the top quark and W boson mass indicate agreement with MSSM. Naturally many assumptions had to be made in making of this diagram so it is no way a proof, but it is a reminder that any new model will have to satisfy very stringent constrains from the electroweak measurements- top quark mass measurement presented here being one of them.

1.3 Top Quark Production and Decay

At the Tevatron the top quark is mainly produced in $t\bar{t}$ pairs. The processes contributing to the production are quark-antiquark annihilation shown in Fig. 1.4 and gluon fusion shown in Fig. 1.5. The gluon fusion is expected to occur between 10 and 20% of the time [15]. Total theoretical $t\bar{t}$ production cross section[1] is $\sigma_{t\bar{t}} = 6.7$ pb [15].

---

[1] The theoretical cros section depends on the value of the top quark mass. The quoted value assumes $\mathrm{M_{top}} = 175\ \mathrm{GeV}/c^2$

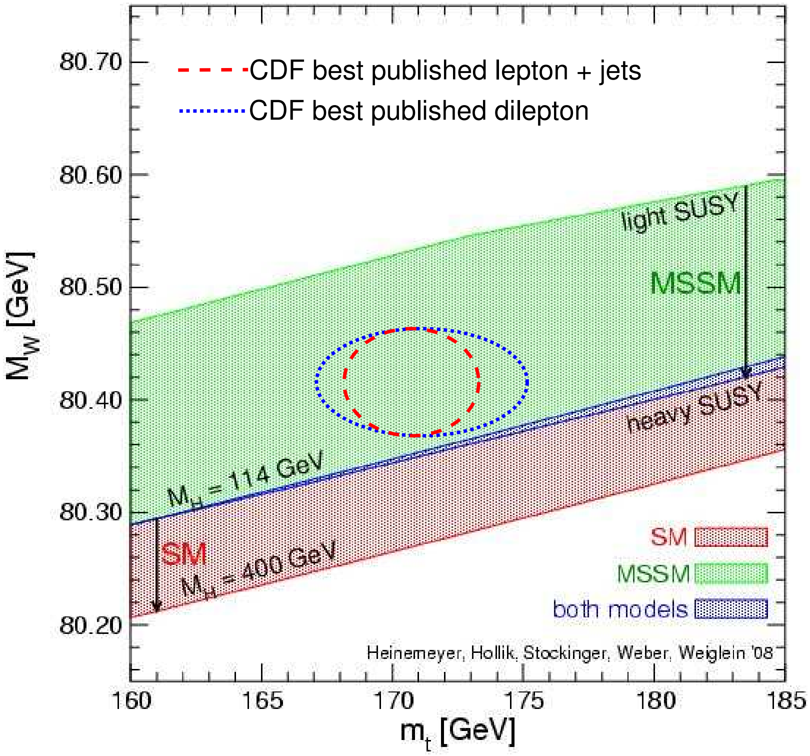


Figure 1.3: Impact of the top quark mass and the W boson mass measurements on the MSSM parameters. Taken from [9, 10] including two-loop corrections for the precision observables [11, 12, 13, 14]. Ellipses indicate CDF II W boson mass measurement [4] and most precise (published or submitted) CDF II dilepton [5] (dotted) and lepton + jets [6] (dashed) top quark mass results. One dimensional 68% confidence levels are indicated by the ellipses.

A single top quark can also be also produced in association with jets with a small cross section of $\sigma_{single\,t} \sim 1.4$ pb [16]. Leading order diagrams for s channel and t channel production are shown in Figure 1.6. We will not attempt to extract mass information from such events. We treat single top production as background.

Top quark decays essentially 100% of the time into a W boson and a b quark. This is because the CKM element $|V_{tb}|$ is close to unity. This CKM element describes the flavour mixing of the top quark and thus rate of top quark decays into W and a b quark. Direct measurements of the decay branching fractions of the top quark do not

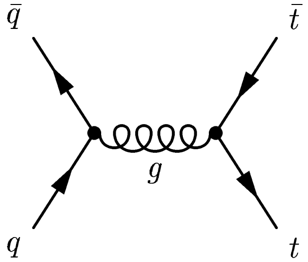


Figure 1.4: Production of $t\bar{t}$ pair through quark-antiquark annihilation

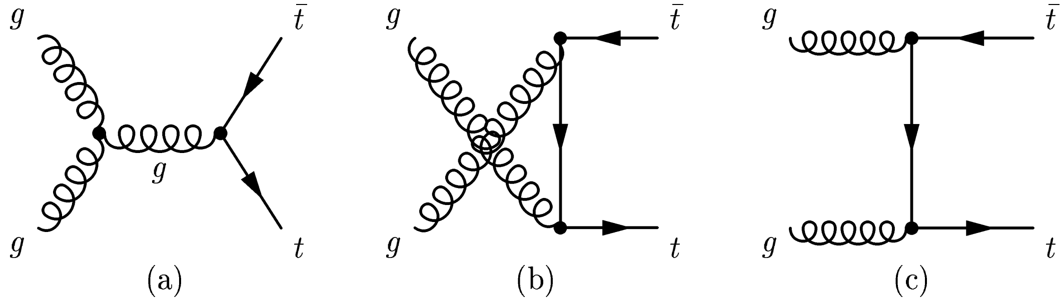


Figure 1.5: Production of $t\bar{t}$ pair through gluon fusion: the s channel (a) and the t channel (b) (c)

constrain $|V_{tb}|$ well, however using measurements of the other CKM matrix elements and the assumption of a three-generation Standard Model give $|V_{tb}| \sim 99.9\%$ [2].

Recently an upper limit on the top quark width has been established experimentally to be $\Gamma_{top} < 13.1$ GeV in an agreement with the Standard Model prediction of approximately 1.5 GeV. To first order the top quark width is given by [17]:

$$\Gamma_{top} = \frac{G_F \mathrm{M}_{\mathrm{top}}^3}{8\pi\sqrt{2}}, \tag{1.3.1}$$

where the mass of W boson and b were neglected in the calculation. This gives the top quark a lifetime of the order of 10^{-25} s. Such short lifetime does not permit the top quark to form bound states, it decays as a free quark. This is a unique property among the quarks. It makes measurements dependent on the spin of the quark possible.

Figure 1.7 shows a decay of a $t\bar{t}$ pair. As mentioned above each top quark decays

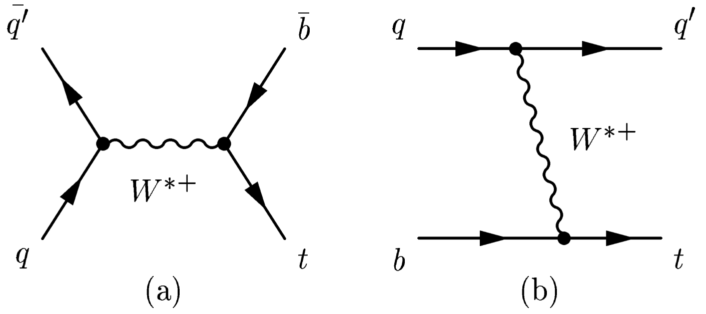


Figure 1.6: Production of single top events in s the channel (a) and the t channel (b)

into a W boson and a b quark. W bosons can then decay leptonically into a lepton-neutrino pair or hadronically into a quark-antiquark pair. It is useful to classify the decay channels of the $t\bar{t}$ pair based on how both W's decay.

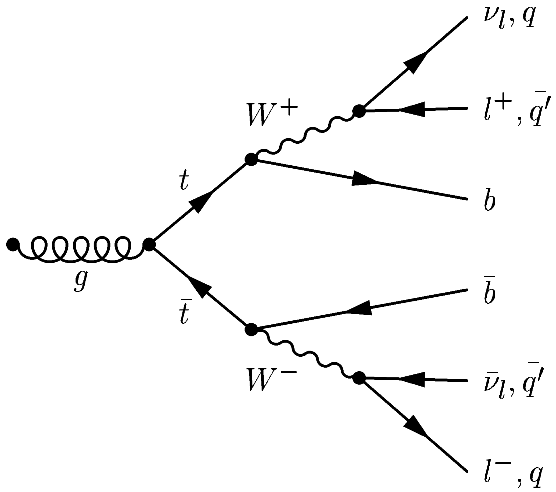


Figure 1.7: Decay of $t\bar{t}$ pair

We have the following possibilities:

- All-hadronic channel

  In this channel both W bosons decay into quark-antiquark pairs. This decay happens most often: about 44% of the time. Quarks have color quantum numbers therefore the probability for W to decay into quarks is approximately three times larger than for W to decay into electron muon or τ lepton. Six jets, collimated showers of particles, are present in the final state. Because of high jet multiplicity and a lack of lepton this type of events are easily faked by QCD multijet production. Since these processes have very high rate at a hadron collider overwhelming background is a challenge. We do not include this decay channel in our analysis, however future analysis may incorporate it.

- Lepton + jets channel

  A convention we will follow from now on is to call a muon or an electron a lepton, treating τ leptons separately. In the nomenclature the lepton + jets channel refers to the decay where one of the W boson decays into quarks and the other W boson decays into an electron or a muon with the associated neutrino. Therefore we expect to observe four jets and a lepton in this type of event. Neutrino will escape undetected but we can infer it's presence by summing the momenta of the observed objects. If this sum is not equal to zero we associate the "missing" momentum with the neutrino. About 35% of $t\bar{t}$ pairs decay into the Lepton+Jets channel. Presence of a lepton essentially eliminates the QCD backgrounds. The signature of this type of an event can be mimicked by an event where one W boson is present e.g. a W+jets production event shown in Fig. 1.8

- Dilepton channel

  Here both W bosons decay into leptons (electrons or muons) and neutrinos yielding observation of two jets, two leptons and a momentum imbalance. Only about 5% of $t\bar{t}$ events fall into this category. Other processes having a similar signature have relatively low rates. The dilepton events may be mimicked by e.g. Drell-Yan production of leptons with jets resulting from radiation emitted by the colliding quarks (Fig. 1.9). Since two neutrinos escape the detector we do not have enough information to calculate the four-vectors of the decay products based on the quantities measured in the detector. This is a drawback of the dilepton channel.

- Tau lepton events

  Events in which at least one of the W boson decays into a τ lepton are very challenging to include in a top quark mass analysis. This is due to complicated nature of the decay of the τ, which produces neutrinos and may produce leptons or quarks. Currently $t\bar{t}$ events with τ leptons are not used in any top quark mass measurements.

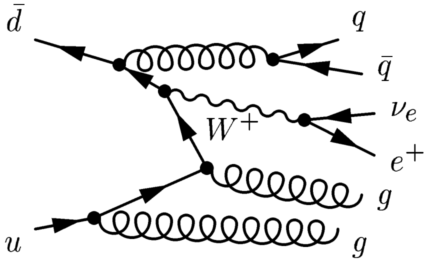


Figure 1.8: Example of W boson production with jets

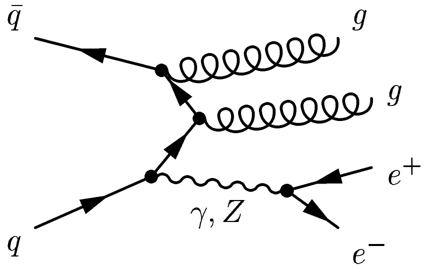


Figure 1.9: Example of Drell-Yan production with jets from initial state radiation

1.4 Measurement Strategy

In the analysis presented here we follow a template method. In this section we give a very brief overview of the analysis. The details of each component will be elucidated in the following chapters.

We start by designing event selection that maximizes number of $t\bar{t}$ events in our datasets and minimizes the contribution of backgrounds. We use two decay channels: dilepton and lepton + jets. Selection is separate for the two. Events must have two energetic jets, two leptons of opposite sign and large momentum imbalance to enter into the dilepton sample. We attempt to reject backgrounds by imposing various topological cuts. For example to reject Drell-Yan events such as the one shown in Fig. 1.9 we require that the two leptons form invariant mass different from the Z boson mass. In the lepton + jets channel we require four jets, one lepton and momentum imbalance. Since this channel has higher statistics we can make more stringent criteria. An important example is identification of b jets. A requirement that at least one jet is identified as coming from a b quark decay helps reject a majority of background events. In particular contribution of W+jets (Fig. 1.8) events is greatly suppressed.

In the $t\bar{t}$ events we measure numerous quantities: momenta of jets and leptons and overall momentum imbalance. We reduce all this information into a variable expected to be highly dependent on the top quark mass. The best choice is the invariant mass of the top quarks produced in the collision. In the dilepton channel we do not have enough information to calculate it based on the measurements from our detector. We integrate over the unknowns to form a proxy variable, the m_t^{NWA}. To improve the precision of the measurement we will also use a second observable sensitive to $\mathrm{M_{top}}$: the H_T- a linear sum of energies of all objects present in an event. In the lepton + jets channel we have more information than is required to extract the invariant mass of the top quarks. We use a χ^2 fitter to choose the most likely invariant mass of the top quarks in a given event: the m_t^{reco}.

Top quark mass measurements are very sensitive to modelling of jets. Formation of jets from partons (quarks or gluons) is a very complicated process. The partons do not propagate freely but form hadrons. Energy can be radiated away from the jet direction and additional energy can enter a jet from other processes in the event. Finally the detectors have nonlinear response and are fragmented in a complex geometry. We apply series of corrections, translating measured energy of a jet into the energy of the originating parton. This calibration is called the "jet energy scale" (JES). The difference in all the effects mentioned above between Monte Carlo simulations and the data is modelled by the parameter Δ_{JES} measured in units of total systematic uncertainty on JES: σ_c. The nominal value of $\Delta_{\mathrm{JES}} = 0\ \sigma_{\mathrm{c}}$ indicates no difference in JES effects between the data and simulation. All the observables described in the previous paragraph rely on jet energies and so are sensitive to jet modelling. However in the lepton + jets channel one of the W bosons decays hadronically, forming a pair of jets. Invariant mass m_{jj} formed by these two jets is not sensitive to $\mathrm{M_{top}}$ but it is very sensitive to the Δ_{JES}. We therefore will employ m_{jj} to calibrate modelling of the jet energy scale effects *in situ* - directly using the $t\bar{t}$ data.

We obtain a probability density function (PDF)(template) of the observable pair (m_t^{NWA}, H_T in the dilepton channel and m_t^{reco}, m_{jj} in the lepton + jets channel) as a function of the top quark mass $\mathrm{M_{top}}$ and Δ_{JES}. In order to do so we generate multiple Monte Carlo (MC) samples with the parameters $\mathrm{M_{top}}$ and Δ_{JES} taking on values in ranges $120 - 240$ GeV/c^2 and $-3.0 - +3.0\ \sigma_{\mathrm{c}}$. Final step in generating the template is to smooth the MC data so that the PDF is a continuously varying function of observables and parameters. The probability density function is extracted using non-parametric techniques, without assuming the functional form for the observables in a MC sample or their dependence on the parameters. We similarly construct templates for the background models. These naturally are not functions of $\mathrm{M}_{\mathrm{top}}$ but do depend on Δ_{JES}.

The final step is to compare the distributions of the observables obtained in the data to the PDFs obtained as described above. This is done by means of a single likelihood simultaneously using the dilepton and lepton + jets data. The fit searches for such combination of the background and signal PDFs that fits the data best. The $\mathrm{M}_{\mathrm{top}}$ and Δ_{JES} of the most likely parent distribution is reported as the final result. Important feature of the fit is that the Δ_{JES} calibration extracted from the lepton + jets channel is applied in a natural way to both channels. As a consistency check we also perform measurements of the top quark mass using the dilepton and lepton + jets data sets separately. The measurement including only lepton + jets data has been presented before in [18]. Revised treatment of systematic uncertainties is presented here for that measurement. Since in the dilepton events there is no hadronically decaying W boson *in situ* JES calibration is not possible when data from this channel alone is used. Thus in the dilepton only fit we fix Δ_{JES} to it's nominal value of $0\ \sigma_{\mathrm{c}}$.

This is the first measurement of the top quark mass where multiple channels are being used simultaneously, providing an intrinsically correct treatment of correlations between the two decay channels in systematic uncertainties. Traditional method of combining the data from different channels is to have dedicated separate analyses in each of those channels. Following this a combination technique such as BLUE [19] is used to obtain a single number. There are two main reasons why our analysis is more robust than such techniques. BLUE and other combination methods rely on the knowledge of the likelihood form in each of the component datasets. This is usually assumed to be Gaussian. While this is most likely not a bad assumption it may induce a bias. In this analysis the likelihood in the dilepton sample is non-Gaussian even in relative vicinity of the central value. Even bigger problem with combining the results from separate measurements is in calculation of systematics. The correlations in systematic effects between different channels must be assumed. In most cases reasonable assumptions can be made. For example we may assume that systematic effect due to initial state radiation is 100% correlated between lepton + jets and dilepton channel. Practically never is it possible to actually measure the value of the correlation. Removing such assumption is clearly beneficial. Of special interest is the uncertainty due to jet energy scale. The dilepton channel used alone does not have an *in situ* JES calibration. To combine the two channels (for example as in Reference [8]) one has to divide the JES uncertainty remaining after the calibration in the lepton + jets channel into the statistical uncertainty on JES arising from *in situ* JES calibration alone and the JES uncertainties resulting from the JES effects outlined before. Value of 0 is assigned for the first component in dilepton channel and 100% correlation is assumed between the other components in the combination. As we can see the procedure is rather convoluted and it's hard to be convinced that all the correlations have been properly accounted for. Our method is free of those potential problems. The correlations in systematic effects are not known but are treated intrinsically in a combined likelihood fit making our method a more robust approach.

The systematic uncertainties start to dominate the total uncertainty on top quark mass measurements due to a large Tevatron data set size available today. In this environment a correct treatment of correlations in systematic effects between different decay channels will be increasingly important. Techniques developed here lay the foundation for a top quark mass measurement with a uniform treatment of systematic uncertainties, where events from all $t\bar{t}$ decay channels are used.

CHAPTER 2

EXPERIMENTAL SETUP

In this chapter we briefly describe the machinery needed to create and study the proton anti-proton collisions that take us back in time to moments after the universe came to be. The time machine is the Tevatron accelerator complex which accelerates beams of protons and antiprotons to the speed of light. The microscope that we use to study what happens when the beams collide is the CDF II detector.

2.1 Tevatron Accelerator Chain

The series of accelerators at Fermilab accomplishes a seemingly impossible task of converting a single “C” size hydrogen bottle into a year’s supply of most energetic $p\bar{p}$ collisions available to high energy physicists today. We describe briefly each stage in this process below. A diagram of the Fermilab accelerator complex is shown in Fig. 2.1 [20].

2.1.1 Proton Acceleration

The first stage of creating a proton beam is the preaccelerator. This electrostatic Cockroft-Walton design accelerator ejects negatively ionized hydrogen into the Linac accelerator with energy of 750 keV. Upon injection to the Linac the continuous H^- beam is separated into bunches. The Linac accelerates the ions to energy of 400 MeV. The long pulse length (20 ms) of the linac is compressed in the Booster accelerator

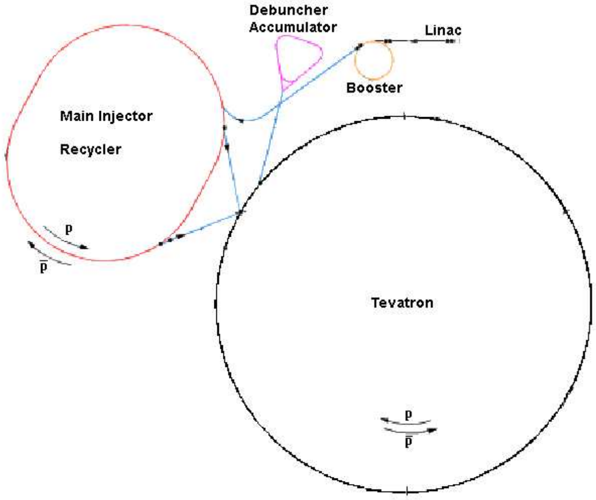


Figure 2.1: Tevatron accelerator chain

through multi-turn injection. At this step the proton beam already in the Booster is merged with the H^- beam by passing both beams through a dipole magnet from two different directions. The combined p-H^- beam is passed through a carbon foil which strips off the electrons from hydrogen ions. Once pulses from the Linac have been injected the beam is accelerated to 8 GeV within 67 ms. The beam is then extracted from the Booster and transferred to the Main Injector.

2.1.2 Main Injector

The Main Injector is a synchrotron 3.3 km in circumference. It accepts protons at 8 GeV from the booster and accelerates them to 150 GeV to inject into the Tevatron or to 120 GeV to send them to the antiproton source or fixed target experiments. The accelerator also can accept antiprotons from the Accumulator and Recycler accelerators, boost them to 150 GeV and inject into the Tevatron.

2.1.3 Manufacturing Antiprotons

To produce antiproton beam a multi-stage process is carried out. First a batch of about 5×10^{12} protons is accelerated to 120 GeV in the Main Injector. A series of radio frequency manipulations are applied in order to narrow the bunches in time, this however increases their momentum spread. Such arrangement of bunches results in maximal phase space density of antiproton beam. The proton batch is extracted and transported through a series of transfer lines to a target enclosure where it is tightly focused on a nickel target. Showers of secondary particles are struck out of the target and hit a collection lens. The lens is made out of Lithium - the lowest density conductor. Current of up to 670000 A is passed through the lens to create a magnetic field that focuses the negative particles and renders them parallel. A magnet placed behind the lens bends out the particles with momenta of 8 GeV/$c\pm$350 MeV/c into the aperture of the Debuncher accelerator located in the triangular tunnel. Only about 10^8 antiprotons are captured. In the Debuncher another series of RF manipulations is performed to minimize the momentum spread of the bunches. Antiproton beam is stochastically cooled in the Debuncher to minimize the transverse beam size and momentum spread. In this technique a feedback loop is created where signals from pickups measuring beam parameters are used to fire kicker electrodes. Over two seconds of antiproton beam orbiting the Debuncher the beam's 6-dimensional phase space density increases by a factor of 500. Before the next pulse of antiprotons arrives the unbunched beam is injected into the Accumulator accelerator located in the same tunnel. Radio Frequency manipulations and further cooling are then applied to bunch and merge (or "stack") the antiproton beam with the antiproton beam already in the Accumulator. The stacking process may go on continuously for days until the beam is extracted and transfered into the Main Injector and the the Recycler accumulator located in the same tunnel as the Main Injector. The 8 GeV antiproton beam is cooled using “electron cooling” where a continuous beam of 4.3 MeV electrons (of the same average velocity as the 8 GeV antiproton beam) is temporarily mixed with the antiproton beam. The energy is transfered to the electron beam “cooling” or reducing the momentum spread of the antiproton beam.[21], [22]

2.1.4 Tevatron

The Tevatron is currently the most powerful accelerator in the world, capable of accelerating proton and antiproton beams up to 980 GeV. During the procedure called “shot setup” the Main Injector transfers 3 trains of 12 proton bunches at 150 GeV into the Tevatron. The 36 proton bunches travel in clockwise direction around the ring. After the protons have been injected the antiprotons from the Recycler are transferred into Main Injector, brought to 150 GeV and injected into Tevatron in counterclockwise direction. Once the beams are injected we have a “store”. Two beams have different orbits but the same bunch structure. Following injection Tevatron ramps to 980 GeV and beams are brought to collision at two points around the ring: at the CDF II and DØ detectors. Bunches collide at a rate of 2.7 MHz. Approximately 10^{13} and around 0.3×10^{13} antiprotons are initially traveling in the Tevatron. Instantaneous luminosities achieved by the Tevatron accelerator exceed $3 \times 10^{32}\ cm^2 s^{-1}$ and have been raising over time as shown in Fig. 2.2. As particles are lost and the beam emittance increases the instantaneous luminosity drops. Once it reaches about $0.5 \times 10^{32}\ cm^2 s^{-1}$ over a period of about a day the store is aborted. Figure 2.3 shows that the integrated luminosity delivered to the CDF II by the Tevatron is nearly 5 fb^{-1}. This provides the experimenters with a wealth of data to test the prevailing theories and search for traces of new physics. Due to inefficiencies in data taking from detector failures as well as dead-time the acquired dataset is approximately 4.0 fb^{-1}. This analysis uses dataset collected up to May 2007 with integrated luminosity of 1.9 fb^{-1}.

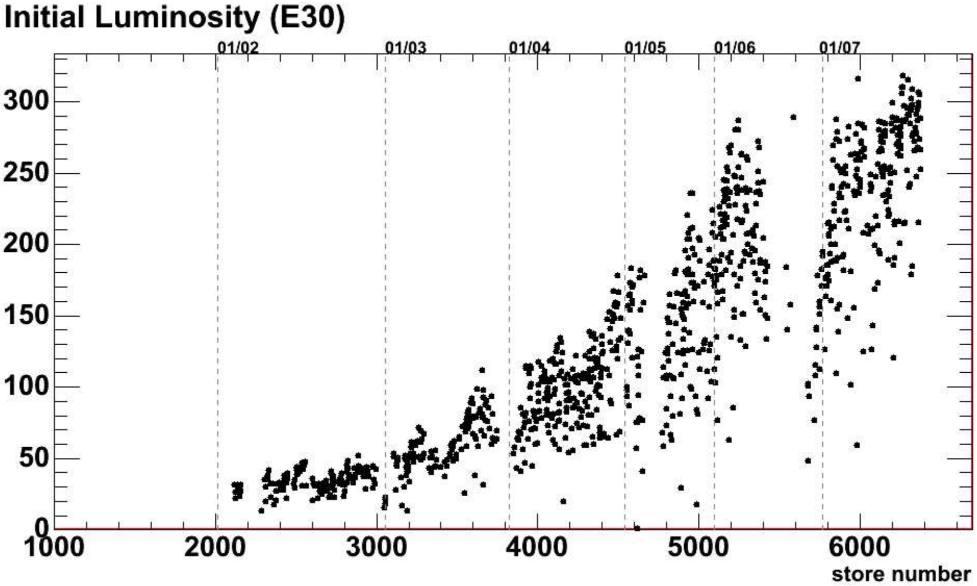


Figure 2.2: Initial instantaneous luminosity at the beginning of store. The abscissa is store number and the vertical lines indicate beginning of a year. Units of the ordinate are cm^2/s

2.2 The CDF II Detector

The Collider Detector at Fermilab is located at one of the two collision points along the ring of the Tevatron accelerator. The detector has a cylindrical geometry. The beam line runs along the axis with proton and antiproton beams brought to collision in the center of the detector. The elevation view of the detector is shown in Figure 2.4. The origin of the coordinate system is at the geometrical center of the detector. The z axis direction runs along the the proton beam direction (due East). Tevatron radius defines the x coordinate direction while the y coordinate points up. It is often more convenient to use angular coordinates. The azimuthal angle ϕ is the angle from

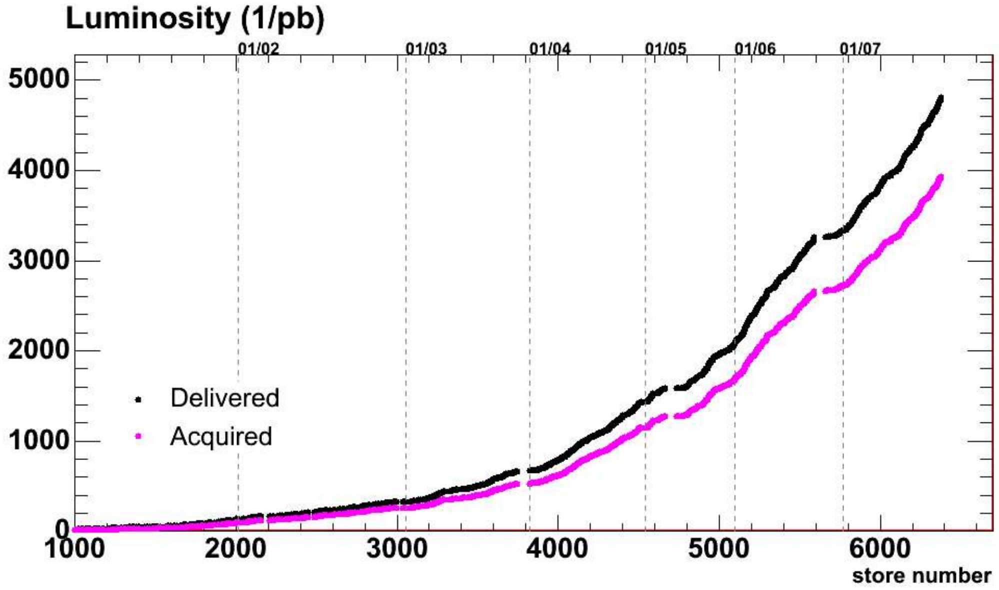


Figure 2.3: Integrated luminosity delivered to and and acquired by the CDF II detector. The abscissa is store number and the vertical lines indicate beginning of a year.

the x axis in the plane transverse to the beamline. The polar angle θ is the angle from the proton beam direction. The pseudorapidity $\eta = -ln(tan(\frac{\theta}{2}))$ is a quantity numerically close to rapidity for highly relativistic particles. The collisions occur along the beamline and are distributed about the center of detector with a Gaussian with a width of 30 cm, therefore we distinguish between the polar angle and pseudorapidity defined with respect to the $z = 0$ point ($\theta_{\rm det}$ and $\eta_{\rm det}$) and those defined with respect to the collision point (θ and η). A detailed description of the CDF II detector is provided in [23]. In this section we briefly introduce the detector subsystems relevant to this analysis starting with the detectors closest to the interaction point.

2.2.1 Tracking

The charged particles escaping the interaction point will be observed in the silicon tracking detectors. The innermost silicon detector, layer 00 lies $1.4 - 1.6$ cm from the

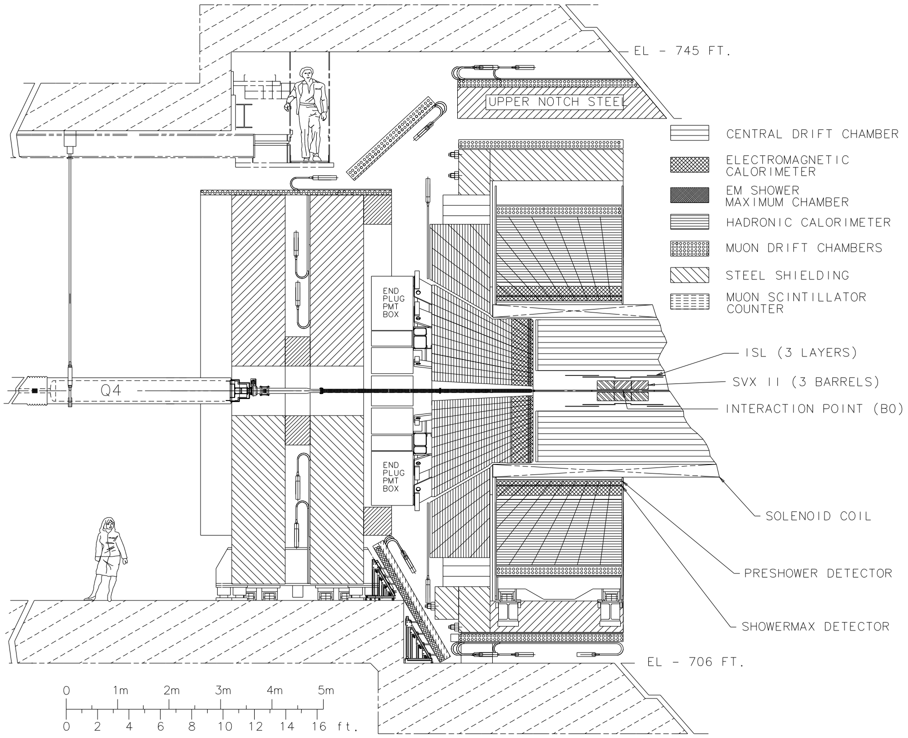


Figure 2.4: Elevation view of the CDF II detector

beamline [24]. Layer 00 is a single-sided silicon strip detector mounted directly onto the beryllium beampipe, proving axial tracking information only [25]. The silicon vertex detector (SVX II) consists of five double sided silicon strip detectors lying at radii from 2.5 cm to 10.6 cm from the beamline [26]. All layers provide track measurement in $r - \phi$ plane as well as $r - z$ plane through inclusion of axial strips and stereo strips on each layer. First, Second and Fourth layer have 90^{o} stereo strips and the other layers have 1.2^{o} small angle stereo strips [25]. The outermost silicon tracking stations are the intermediate silicon layers (ISL) - a set of double sided silicon strip detectors with axial and small angle stereo (1.2^{o}) tracking. In the pseudorapidity region $|\eta_{\text{det}}| < 1.0$ we have one ISL layer at radius of 22 cm and for the region $1.0 < |\eta_{\text{det}}| < 2.0$ we have two ISL layers at 20 and 28 cm radii [26]. The best tracking position resolution achieved by the silicon detector system is 9 μm.

Tracking of charged particles in radii between 43 and 142 cm in pseudorapidity region $|\eta_{\text{det}}| < 1.0$ is provided by a cylindrical open cell drift chamber, the central outer tracker (COT). The COT's sense wires run in z direction between the endplates of the COT and are grouped in 8 superlayers. Each superlayer is made up of 168 to 480 supercells comprising the sense wires, field shaping wires and cathode sheets shared between the neighboring supercells. Four of the superlayers have a 2^{o} small angle stereo configuration and are arranged alternately with the axial superlayers. The supercells are tilted from the radial direction by 35^{o} to ensure azimuthal drift direction. The chamber operates with a equal parts mixture of argon and ethane. This choice of gas mixture provides a fast drift time ($\sim 50\ \mu\text{m}/\text{ns}^{-1}$). Such fast drift times ensure that the electron avalanches will be collected on the sense wires before the next bunch crossing occurs. The COT achieves 140 μm resolution on position of a hit. The tracking detectors are immersed in a 1.4 T solenoidal magnetic field enabling momentum measurement. Momentum resolution achieved by the COT measured using cosmic rays is $\sigma_{p_T}/p_T^2 = 0.17 \times 10^{-3}$. Note that the resolution decreases for increasing momentum due to progressively worse curvature resolution for almost straight tracks. [27] [28]

2.2.2 Calorimetry

Energy of the particles escaping the collision is measured by the calorimeter system. CDF uses lead-scintillator and steel-scintillator sampling devices for the electromagnetic and hadronic calorimetry. The calorimeter system is comprised of central electromagnetic calorimeter (CEM) [29], central (CHA) and wall (WHA) hadronic calorimeter [30] covering pseudorapidity region $|\eta_{\text{det}}| < 1.0$ and plug electromagnetic (PEM) [31] and hadronic (PHA) covering pseudorapidity region $1.1 < |\eta_{\text{det}}| < 3.6$.

All detectors are divided into projective towers each spanning 15^{o} in ϕ and approximately 0.1 in η_{det} with the pseudorapidity segmentation growing for high η_{det}. The ϕ segmentation is doubled for six plug calorimeter towers lowest in η_{det}. The segmentation in the hadronic calorimeters matches the segmentation in the electromagnetic calorimeters. A single tower can be composed of towers in several calorimetry detectors e.g. in CEM, CHA and WHA detectors. Shower maximum detectors are embedded in the central and plug electromagnetic calorimeters at approximately 6 radiation lengths from the collision point [29, 32]. The central electromagnetic strip chamber (CES) provides position measurement for the shower maximum with resolution of 2 mm for 50 GeV photons and electrons [29]. The plug electromagnetic shower maximum detector consists of two layers of scintillator strips providing position resolution for electrons and photons of 1 mm [32]. Energy resolution in the CEM as determined using $Z \rightarrow ee$ events is $13.5\%\sqrt{E_T} \oplus 2\%$[1]. The PEM resolution has been measured using test beam data at $16\%\sqrt{E}$. The test beam data has been also used to determine single pion response of the hadronic calorimeters. The CHA detector achieves resolution of $75\%\sqrt{E_T}$, WHA $80\%\sqrt{E_T}$ and PHA $80\%\sqrt{E_T} \oplus 5\%$. All units are in GeV. [28]

2.2.3 Muon Chambers

The muons created in the interaction will leave a track in the inner tracking detectors, then pass through the calorimeter volume as minimum ionizing particles. They can be detected then in drift chambers located beyond the calorimeters.

The muon detectors used in this analysis are the Central Muon Detector (CMU), Central Muon Extension (CMX) and Central Muon Upgrade (CMP). All detectors

---

[1] The transverse energy E_T is defined as $E_T \equiv sin(\theta) \times E$

are comprised of single cell wire chambers of similar design operated in proportional mode. The CMU detector is symmetric in azimuth. The CMU modules have four layers of wire chambers attached to the outside of the central calorimeter wedges. Detector's wires run in the z direction. The measurement of ϕ coordinate of passing muon is obtained by converting the hit arrival time to distance from the wire, knowing the drift velocity. The z coordinate measurement is provided by sense wire readout through a resistive wire. CMU detector provides muon coverage in the pseudorapidity range of $0 < |\eta| < 0.6$. The CMP chambers form a box structure of four layers of drift tubes mounted outside the CMU chambers. In addition 3 interaction lengths of steel absorber lie between the chambers to reduce contamination from hadronic showers leaking out of the calorimeters. Because of it's lack of symmetry the pseudorapidity coverage varies for the CMP detector extending up to $|\eta| < 0.6$. The CMX detector is formed by 8 layers of drift tubes arranged in a conical shape extending the muon coverage to pseudorapidity of $|\eta| < 1.0$. Scintillator detectors (CSP, CSX) are attached to the outside of muon chambers. They are not used in the offline reconstruction but the CSX system is used for trigger decisions. Muons above threshold energy of 3 GeV will be detected in all of the central muon systems. [28] [26]

2.2.4 Luminosity Counters

The Cherenkov Luminosity Counters (CLC) measure the number of inelastic $p\bar{p}$ collisions in a single bunch crossing within few percent. This quantity is used together with the knowledge of $p\bar{p}$ interaction cross section and detector acceptance to calculate the instantaneous and integrated luminosity. The CLC consists of two sets of gas Cherenkov counters on both sides of the CDF detector. Each set comprises 3 rings of 16 conical counters pointing back at the interaction point and covering the pseudorapidity range of $3.75 < |\eta| < 4.75$. The counters are up to 180 cm long. Particles from $p\bar{p}$ collisions will travel trough the length of the counter emitting Cherenkov radiation which is collected by photomultiplier tubes at the end of the cone. The number of particles observed in the CLC is translated into luminosity measurement with an uncertainty of 5.9%. [33]

2.2.5 Trigger System

The trigger system is crucial to maximizing the usefulness of data collected by the CDF detector for physics analyses. The bunch crossing occurs at a rate of 2.7MHz. The charge of the trigger system is to reduce this flood of data so that it can be stored and analyzed. Only events containing objects with high transverse momenta will be accepted. The trigger is divided into three levels each providing successively more sophisticated event reconstruction. The first level features custom electronics. At this stage we can apply cuts based on presence of calorimeter towers above certain threshold. A coarse calculation of $\not{E}_T$ is also performed. A tracking system (XFT) can extrapolate tracks to calorimeter towers and muon chambers. Muon momentum threshold can be applied using timing information. Data rate at the output of Level 1 is approximately 40 kHz.

Level 2 is a combination custom hardware-PC system. More sophisticated event reconstruction algorithms such as calorimeter cluster finding can be applied at level 2 providing further event rejection. A hardware subsystem is capable of detecting displaced vertices consistent with b quark decays. Maximum accept rate of the Level 2 trigger is around 800 Hz.

The last stage before the data is stored on disk is the Level 3 trigger consisting of a cluster of commercially available computers. The same event reconstruction algorithms are applied here as in final data analysis including tracking, electron and muon identification and jet clustering. About 100 events per second are written out.

2.2.6 Monte Carlo Generation

This method of measuring the top quark mass relies on having accurate models of signal and background processes. With few exceptions those models are obtained from Monte Carlo simulations. The signal ($t\bar{t}$) samples are generated with a leading-order generator PYTHIA 6.216 [34]. We generate $t\bar{t}$ samples with the input top mass ranging from 120 GeV/c^2 to 240 GeV/c^2. Additional $t\bar{t}$ samples are generated using HERWIG 6.510 [35] for evaluation of systematics. We also simulate diboson production (WW, WZ, ZZ) using the PYTHIA generator. The MADEVENT [36] generator is used in simulation of single top quark events in s-channel and t-channel.

To simulate Drell-Yan process as well as $W +$ jets production we use ALPGEN 2.10 [37]. We generate separate samples based on number of additional light flavor and heavy flavor partons produced with the W or Z boson, so that we don't rely on multiple jet production purely from showering. We further divide the Drell-Yan production based on the invariant mass of the dilepton system into three regions: below, near and above the Z boson mass. This scheme allows us to reduce the overall number of events that we need to simulate. The dilepton selection is designed to reject Drell-Yan events and only events in the tails of the Z peak have a chance of passing the selection. By simulating events in the tails separately we can fully explore the kinematics of the events entering our samples without cost of the simulation becoming prohibitive. Similarly we simulate $W+$ heavy flavor production separately from $W+$ light flavor production since we have cuts designed to reject $W+$ light flavor from our final lepton + jets samples. We then combine the samples with separate final states according to the cross section reported by ALPGEN. Output of ALPGEN generator is interfaced with PYTHIA 6.325 [38] for parton showering. In order to ensure that the showering doesn't induce over-coverage of the phase space (e.g. when W + 1 parton event produces an extra jet becoming a W + 2 parton event) we use the MLM [39] matching scheme. In addition since we generate samples with different flavor content separately we check if b and c quarks do not appear after showering in the samples generated with only light flavor at the matrix element level. We also check if b quarks do not appear in samples generated only with light flavor and c at the matrix element level.

To model possible multiple proton-antiproton interactions occurring in a single bunch crossing we add to the events simulated for a particular process, interactions where no partons with high transverse momenta are produced. Those "minimum bias" collisions are simulated with PYTHIA. The number of minimum bias interactions added to a given event depends on the instantaneous luminosity profile of the data run of this event. The instantaneous luminosity profile is matched between Monte Carlo samples and data only for the first 1.2 fb^{-1} of integrated luminosity. This incorrect model is a source of bias which is evaluated in Sec. 7.2

The event generators are interfaced with CDF II simulation software that models the detector response to the passage of the particles through detector volumes. Response of the calorimeters as well as gas ionization model are parametrized in order to shorten the simulation time required. [40]

In some cases it is possible to generate new samples with different underlying parameters by modifying the reconstruction or assigning different weights to events. For example in order to fit for the jet energy scale parameter Δ_{JES} (*cf.* Sec. 1.4) we need to generate about 2000 $t\bar{t}$ samples with a range of input M_{top} and Δ_{JES}

parameters. In practice we simulate just the samples with nominal Δ_{JES} and having those we construct the others by shifting jet energies (prior to event selection or reconstruction).

CHAPTER 3

RECONSTRUCTION OF PHYSICS OBJECTS

3.1 Track Fitting

In $t\bar{t}$ events decaying in dilepton and lepton + jets channels one or two high momentum leptons will be found. Their presence is characterized by observation of a high momentum track in the COT. In this section we describe the algorithm used to find tracks in the CDF detector.

The tracking algorithm starts by finding track segments in the axial superlayers of the COT. The segment is seeded by hits on 3 consecutive sense wires. The timing of the hits must be such that the angle of the track that would deposit those hits is less than 50^{o} from the radial direction. The timing is calculated using the innermost and outermost sense wire hits and cross checked with the middle wire. Next a seed is grown in an iterative procedure to a track segment by searching for hits that are consistent with forming a straight line.

After the track segments have been found in the axial superlayers the track segments matches in the two outer superlayers are found and used to compute track parameters. These are then used to find matching track segments in the inner axial superlayers and track segments in the outer superlayers not associated with any tracks.

An alternative algorithm called “histogram linking” is also performed to find axial tracks. Algorithm selects a given segment and using the beam position calculates and histograms the curvature of hits in other superlayers. A track will be comprised of hits accumulated in the same histogram bin. Parameters in the $r-\pi$ plane of the axial tracks found using both algorithms are determined using a χ^2 fit.

Both collection of the axial tracks are subsequently used in stereo pattern recognition where segments and hits in the stereo superlayers are matched to the axial tracks. A χ^2 fit is performed to find parameters of the track in $r-z$ plane, followed by a fit to determine parameters of the track in 3-d space simultaneously.

To improve the momentum resolution for muon and electron reconstruction the stereo tracks are refit using the beam position as a constraint. The transverse momentum resulting from this fit is used in muon reconstruction (described later in Sec. 3.3), but the track itself is not used when the silicon hits are being added.

Silicon detector hits are added to the tracks using an iterative “outside in” algorithm. A COT track is extrapolated into the outermost silicon layer and associated hits are found. Uncertainties on the track parameters are used to determine the search region. For each of the silicon hits found a new track is refit with the added information. Two best candidate hits are kept. The two new tracks are then extrapolated into the next layer and procedure is repeated finally resulting in a small tree of tracks. The track with the best χ^2 is kept. [40]

3.2 Electrons

Electrons used in this analysis include central (“CEM”) and plug electrons (“PHX”). Plug electrons are only used in the dilepton dataset while central electrons are used in both lepton + jets and dilepton datasets.

To reconstruct the CEM electrons clusters of at most two towers adjacent in the η direction are found in the central calorimeter. The transverse electromagnetic energy deposited in the cluster must be greater than 20 GeV. Since electrons are expected to leave most of their energy in the electromagnetic compartment we require that the ratio of energy deposited in hadronic and electromagnetic compartments in the cluster is less than $0.055 + 0.00045 \times \mathrm{E(GeV)}$. The dependence of the cut on total energy of the electron captures possibility of electromagnetic shower spilling into the hadronic compartment. A COT track with p_T of at least 10 GeV/c must point into the location of the cluster. We require that the track z intercept is within 60 cm of the origin, and that the track has hits in at least 3 axial superlayers and 2 stereo superlayers. The ratio of transverse energy of the cluster to the transverse momentum of the track must be less than 2 for the tracks with $p_T < 50$ GeV/c. Position of the electromagnetic shower measured in the CES detector must match the extrapolated track position within 10 cm in the z direction. Difference between the extrapolated track and the shower in the CES detector in the direction along the azimuth (δX) must be $-3.0\text{ cm} \leq q \times \delta X \leq 1.5\text{ cm}$. The cut is asymmetric due to the bremsstrahlung radiation and depends on the charge q of the track. Lateral shower profile measured by the calorimeter and shower maximum detectors must match that measured in the test beam electrons.

Plug electrons (so called "PHX electrons") are constructed by combining information from the plug calorimeter (PEM), shower maximum detector (PES) and the silicon tracker (SVX). We require transverse electromagnetic energy in a 2×2 tower cluster to be at least 20 GeV. The ratio of hadronic to electromagnetic energy in the cluster must be less than 0.05. A cluster in the scintillator shower maximum detector must be found within 3 cm of the location of the calorimeter cluster. The PES cluster must lie within pseudorapidity region $1.2 \leq \eta_{det} \leq 2.8$. The shower profile in the PEM and PES detectors must compare well to the profiles measured in the test beam. The plug electrons lie in the pseudorapidity region not covered by the central tracking therefore an independent tracking mechanism had to be devised. This algorithm constructs a track helix based on the location of the $p\bar{p}$ interaction, location of the cluster in shower maximum detector and transverse energy measurement in the plug calorimeter cluster. This seed helix is then used to construct a track using hits in the silicon detector. To accept the plug electron candidate we require that such track matches the electromagnetic cluster, has at least 3 hits in the silicon detector and it's point of closest approach to the beamline is less than 60 cm from the origin.

Electrons are expected to deposit their energy in a narrow region of the calorimeter. To suppress contribution of hadrons faking electrons we define isolation as transverse energy in towers within distance R of 0.4 from the cluster (excluding cluster towers themselves) divided by the transverse energy of the cluster. If this ration is less than 0.1 we call the electron candidate "isolated". In the lepton + jets dataset we require that all electron candidates are isolated. In the dilepton sample we allow non-isolated central electrons.

Electron identification efficiencies are studied by selecting events in the data with two electrons from a Drell-Yan process. Efficiencies found are approximately 70-80%. The same technique applied in Monte Carlo samples gives slightly higher efficiencies leading to electron identification scale factor of 0.94 for PHX electrons and 0.97-0.98 for isolated and non-isolated CEM electrons.

Efficiency for a CEM electron to fire a trigger is evaluated in W candidate sample by separately studying Level 1, 2 and 3 requirements. Measured trigger efficiency is 98%.

3.3 Muons

Muons produced in the collision will leave a track in the COT and possibly SVX detectors, they will pass through the calorimeters leaving energy consistent with a minimum ionizing particle and finally pass through the muon chambers leaving hits which are then reconstructed as track segments or "stubs". We divide the muon types depending on which muon system they traverse. The muon types considered are: a CMUP muon - traversing CMU and CMP detectors, a CMU (CMP) muon traversing CMU (CMP) detector but not CMP (CMU) detector, a CMX muon and a CMIO muon which does not traverse any of the central muon detectors. We determine if muon can traverse any of this detectors by extrapolating the COT track. Track requirements for all the muon types are common. We require a that a track p_T is at least 20 GeV/c. We require that the track has segments in 3 axial and 2 stereo COT superlayers. Point of closest approach of the track to the beamline must lie within 60 cm of the origin and it's impact parameter must be less than 0.2 cm. If the track has hits in the SVX the impact parameter cut is tightened to 0.02 cm. For all muon types we require electromagnetic transverse energy deposition in the calorimeter tower matched to the track to be less than $2+max(0, p_T - 100) \times 0.0115$ GeV and the hadronic energy has to be less than $6+max(0, p_T - 100) \times 0.028$ GeV. In addition we require CMIO total energy deposition to be greater than 0.1 GeV. For a CMUP muon we require that difference in position in azimuthal direction of the track extrapolated to the depth of muon chambers and the muon stub is less than 7 cm for the CMU stub and 5 cm for the CMP stub. For a CMU muon this difference has to be less than 3 cm and for CMP muon less than 5 cm. For the CMX muon we require 6 cm or better match.

A variable equivalent to electron isolation for muons is ratio of transverse energy within radius of 0.4 from the muon traversal position to the track transverse momentum. Muon is considered isolated if this ratio is less than 0.1.

In the lepton + jets sample we consider isolated CMUP and CMX muons. To increase acceptance in the dilepton sample we allow CMU, CMP (isolated and non-isolated) muons as well as isolated CMIO muons.

Both trigger efficiencies for muons are studied using $Z \rightarrow \mu\mu$ events. Trigger efficiency for CMUP muon is about 92% and for CMX muon about 87%. There are small differences in muon identification efficiency in data and Monte Carlo. Generally Monte Carlo samples predict slightly higher identification efficiency resulting in identification scale factors of 0.93-0.98 for all muon categories except CMIO. The scale factor for CMIO muons is 1.03.

3.4 Primary Vertex

The luminous region at the CDF II detector lies close to the geometrical axis of the detector. It is not confined to a small spot in space, but rather spread along the beamline. The collisions occur according to a Gaussian distribution with standard deviation of approximately 30 cm. The location of the $p\bar{p}$ collision is obtained by finding the mean z coordinate of tracks in the COT. There may be more than one such "z-vertex" in an event. If this is the case we choose the z-vertex closest to the z intercept of an isolated lepton. In samples where no leptons are present the primary vertex is chosen based on the linear sum of transverse momenta of the tracks originating from that interaction. The location of primary interaction is used as the origin in calculating transverse energies of towers and jets. [40]

3.5 Jet Clustering and Jet Corrections

Partons produced in the collisions are not free to escape, but instead must condense into hadrons. This is due to the confining nature of the strong interaction. Hadrons may in turn decay into lighter hadrons and leptons. Experimentally we will observe collimated showers of particles in the detector carrying the momentum of the source parton. We will not attempt to identify each particle in such a cascade, but we will cluster them into an object called "jet". The algorithm JETCLU [41] employed in constructing jets proceeds as follows:

First, all calorimeter towers with E_T above 1 GeV are found and sorted according to their transverse energy. These towers are called "seed towers". Note that the pseudorapidity coordinate used to calculate the transverse energies is calculated with respect to the interaction vertex, not with respect to the center of the detector. The algorithm then forms preclusters by adding neighboring seed towers, but only within distance $R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ of 0.4. For each precluster centroid η and ϕ coordinates are calculated as weighted averages of individual tower coordinates. Tower transverse energies are used in the average. For each precluster, towers with $E_T > 0.1$ GeV, which lie within the distance R=0.4 of the precluster centroid form a cluster. If any two clusters share some calorimeter towers the towers will be uniquely assigned or the clusters will be merged. If the transverse energy in towers shared between two clusters is greater than a half of the transverse energy in the smaller cluster, the clusters are merged. Otherwise each shared tower is uniquely assigned to one of the clusters based on it's proximity to the centroid in the $\eta - \phi$ space. Centroids of all clusters are recalculated and towers added and subtracted from clusters until no towers are entering or leaving any clusters. Note that towers associated with reconstructed leptons are removed from the original tower collection in order not to reconstruct an electron as a jet and not to capture muon energy in one of the reconstructed jets. The cone radius R used for jets in this analysis is always 0.4 however a different radius might be used for a different analysis.

Each calorimeter tower of the jet is assigned a massless four-vector whose direction is the same as that of a unit vector pointing from the interaction point to the center of the tower. The energy component of this vector is set to the energy in the tower. A sum of vectors corresponding to all towers is assigned as a four-vector of the jet. Note that the electromagnetic and hadronic compartments are treated separately.

As we have seen in chapter 2.2.2 the calorimeter is a complicated machine. Different regions of calorimeter display different response and there are regions of the calorimeter which are not instrumented ("cracks"). In addition the response of a single calorimeter tower is non-linear and depends on the number of particles impacting the tower and their momenta. Spurious energy may be deposited in the jet due to soft interactions occurring in the same bunch crossing. To obtain the jet energy carried by the particles contained in this jet ("particle level jet") we need to compensate for these effects by applying a series of corrections. The particle level jet is not equivalent to the parton that originated the jet and further corrections are needed to obtain the energy of that parton. We will discuss the corrections as they are applied in sequence to jets. [42]

3.5.1 Relative Jet Energy Corrections

To make the energy response of the calorimeter uniform in pseudorapidity we apply the relative jet energy corrections. This correction is derived using dijet events using "trigger and probe" approach. A jet with detector pseudorapidity range of $0.2 < |\eta_D| < 0.6$ is the "trigger" jet and the jet falling outside of this region is the "probe" jet. A correction is derived to make the average p_T of the probe jet equal to that of the trigger jet. The correction is a function of the detector pseudorapidity of the jet and it's transverse momentum. Due to the calorimeter simulation in the crack and forward regions the correction has to be derived separately for data and Monte-Carlo events. In data, samples triggered on presence of high momentum jet are used. To derive the correction for Monte-Carlo events dijet samples generated with PYTHIA generator are used. Systematic uncertainty on the correction takes into account the inaccurate modelling of the dijet events, event selection biases, remaining non-uniformities in the jet response and inaccuracies of Monte-Carlo modeling of the dijet events. The uncertainty shown in Fig. 3.1 is the smallest in the central region (0.5%) and grows to 1.5% for the forward jets with high momentum and to 3% for the forward jets with low momentum.

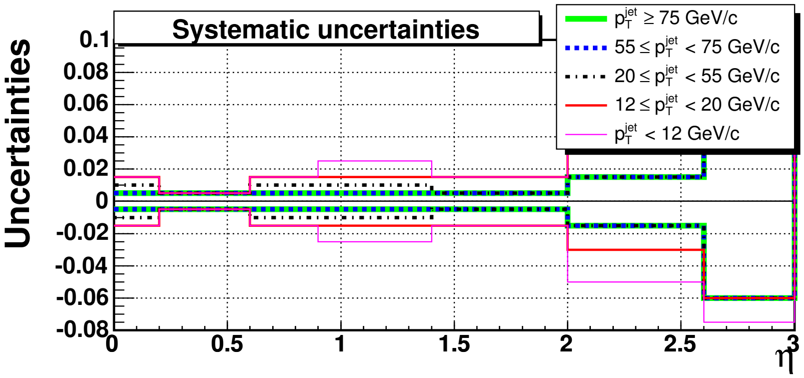


Figure 3.1: Fractional uncertainty on momentum of a jet due to differences in calorimeter regions as a function of η_{det}. Several ranges of jet p_T (after relative corrections) are shown.

3.5.2 Multiple Interaction Corrections

At the Tevatron peak instantaneous luminosities we can have as many as 10 $p\bar{p}$ interactions in a single bunch crossing.One of those interactions naturally gives rise to the event of interest. The other interactions produce particles which will add to the energy of the measured jets. To determine how much energy needs to be subtracted from the jets we use minimum bias data characterized by coincidence of hits in the forward beam counters. A random cone with a centroid with pseudorapidity region of $0.2 < |\eta_D| < 0.6$ is selected in each event and the transverse energy in that cone is computed. Average E_T deposited is parametrized as a function of number of $p\bar{p}$ interactions and taken as a correction. We count number of $p\bar{p}$ interactions in an event through number of so-called "z-vertices" (*cf.* Sec. 3.4)

For each additional interaction 0.6 GeV is expected to be deposited in a jet. Studies of data samples enriched in jets, W boson and binning the data in instantaneous luminosity bins yield a systematic uncertainty of less than 1% on the jet momentum as shown in Fig. 3.2.

3.5.3 Absolute Energy Corrections

The energy read out by the calorimeter does not necessarily correspond to the energy of an incoming particle. In general the calorimeter response will depend on the momentum and kind of the particle. For example the calorimeter response is different for two pions of 3 GeV/c momentum than for one pion with 6 GeV/c momentum. The nonlinearity in single particle response is especially pronounced for particles with momenta less than 6 GeV/c. This results in overall non-linearity in jet response. To estimate the size of correction needed, a jet clustering algorithm is applied to stable particle lists in PYTHIA Monte-Carlo dijet events. Such particle jets are matched to the calorimeter jets within R of 0.1. The correction is derived using a likelihood fit dependent on the momenta of the particle and calorimeter jets. Jets in central region ($0.2 < |\eta_D| < 0.6$) only are used. The energy of the jet has to be increased by as much

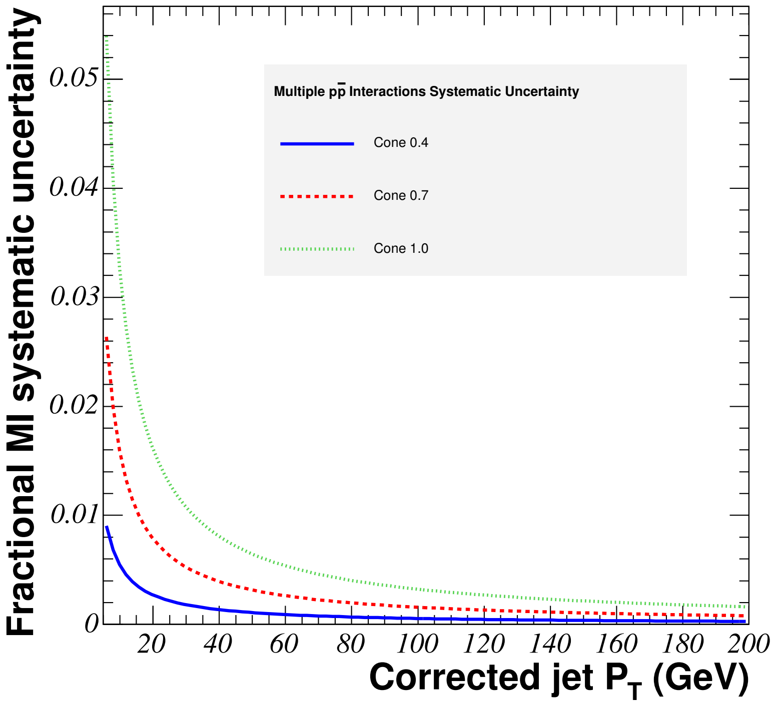


Figure 3.2: Fractional uncertainty on momentum of a jet due to multiple $p\bar{p}$ interactions. Uncertainty for three cone sizes are shown.

as 30% for a low energy jet ($p_T \sim 20$ GeV/c). The correction drops to approximatelly 10% for high energy jet. We refer to jet afer the absolute corrections as corrected to particle level.

The systematic uncertainties on the absolute energy correction arise form calorimeter response to single particles (given by their E/p ratio). Additional uncertainties are caused by modeling of particle momentum spectrum inside a jet (fragmentation) and as well as stability of calorimeter calibration and non-uniformity of calorimeter in ϕ coordinate. The uncertainty on the correction varies between approximately 2% for a jet with p_T of 20 GeV/c to 3% for a high energy jet as shown in Fig. 3.3.

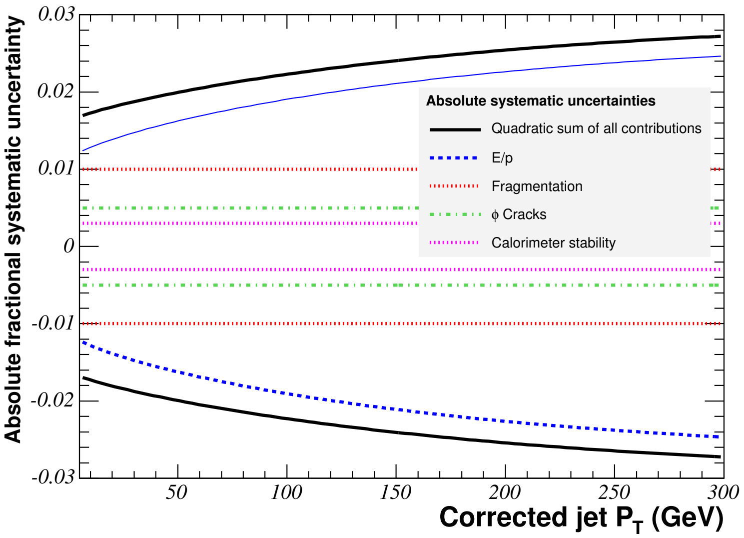


Figure 3.3: Fractional uncertainty on momentum of a jet (after absolute corrections) due to uncertainties on the absolute corrections. Components of the uncertainty are shown in dashed or dotted curves. The black solid curve gives the total absolute jet energy scale uncertainty.

3.5.4 Underlying Event Corrections

At a hadron collider such as the Tevatron the hard scattering of quarks and gluons is of the main interest, however in each $p\bar{p}$ interaction the proton-antiproton remnants, not involved in the hard scattering, will also interact. This spectator interaction will deposit some energy in jets originating from the hard scattering process. Underlying event corrections remove this contribution. The calculation of this contribution is performed in the same way as calculation of multiple interaction corrections but with only one interaction present in the events. The correction needed for cone 0.4 jets is 0.6 GeV with systematic uncertainty of 30%.

3.5.5 Out-of-Cone Corrections

To compute the energy of the parton escaping the collision based on the particle jet we need to take into account the energy flowing outside the cone of the jet. The procedure to determine the energy lost outside of the jet cone is similar to the procedure used to calculate the absolute energy corrections. Partons escaping the collision are matched to particle jets in PYTHIA dijet Monte Carlo events. We find that for a jet with p_T of 20 GeV/c approximately 12% of energy flows outside of the cone. This fraction decreases for higher energy jets. Jet after the out-of-cone corrections is refered to as parton-level jet since it's energy is our best estimate of the original parton's energy.

The systematic uncertainties on this correction are computed by comparing energy deposited in an annulus with radius R=1.3 surrounding the jet cone in data and Monte-Carlo γ+jet events (triggered on the presence of isolated electromagnetic cluster). For a 0.4 cone jet with transverse momentum of 20 GeV/c the uncertainty on the correction is 9% and about 2% for a jet with transverse momentum of 60 GeV. Uncertainty on the energy deposited beyond the radius R of 1.3, so called "Splash out", is evaluated using Monte Carlo samples to be 0.25 GeV per jet.

3.5.6 Total Jet Energy Scale Uncertainty

The jet energy calibration, or the series of corrections that give us the parton momentum given the jet momentum, is called the "jet energy scale". The uncertainties on the corrected jet momentum due to all the jet energy scale effects described above are shown in Fig. 3.4. The total uncertainty (shown in solid black curve) for jets with a low momentum is as much as 10%. For jets of high momentum the total uncertainty is approximately 3%.

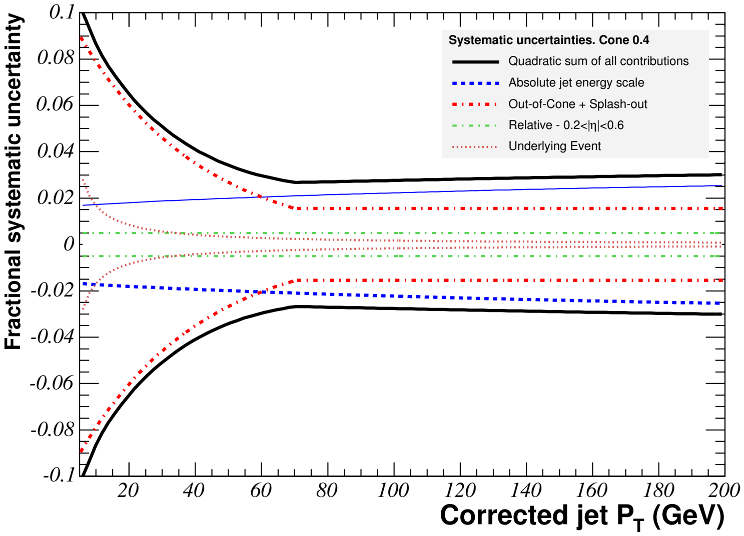


Figure 3.4: Total jet energy scale uncertainty is shown in solid black curve as a function of p_T of a jet corrected to parton level. Separate components of the uncertainty are plotted in dotted or dashed curves

3.6 Identification of b Jets

In all $t\bar{t}$ events two b jets are present. The identification of a jet as originating from a b quark suppresses background events and provides a handle to separate the sample into statistically independent subsamples. b jet identification or "tagging" relies on long lifetime of the b quark. SecVtx algorithm is used to find vertices within jets that are displaced from the $p\bar{p}$ interaction point in the $r-\phi$ plane. The algorithm attempts to construct a vertex using tracks contained inside the jet whose impact parameter d_0 is inconsistent with 0. If such a vertex is found it's signed distance in $r-\phi$ plane L_{2D} from the $p\bar{p}$ interaction point is computed. If the significance $L_{2D}/\sigma_{L_{2D}}$ is greater than 3.0 the jet is considered "tagged". If the significance is less than -3.0 we have a "negative tag". This is the case when the vector from the primary vertex to the secondary vertex points in opposite direction to the direction of the jet. Such an occurrence is due to poorly reconstructed tracks and is a useful handle for evaluating tagging rate for light flavor jets (or "mistags"). The mistag rate is measured in a sample acquired using a jet trigger. After subtracting contribution of heavy flavor jets having negative tags, light flavor long lived meson decays as well as contribution of negative tags resulting from detector material interaction the rate is parametrized in so called "mistag matrix". Mistag matrix is a function of jet E_T, η, ϕ and track multiplicity of a given jet as well as scalar sum of transverse energies of all the jets in a given event. Mistag rate is shown in Figure 3.5. The efficiency for tagging a b jet is measured in a data sample with low energy ($E_T > 9$ GeV, track $p_T > 9$ GeV/c) non-isolated electrons. Two jets are required, one encompassing the electron within it's cone. The "electron" jet is presumed to have resulted from a semileptonic decay of a heavy flavor quark. The fractions of jets containing b decays and c decays are estimated by studying D^0 decays or alternatively muons within the jets. Once the heavy flavor content of the jets is known the tagging efficiency can be computed. The tagging rate in $t\bar{t}$ events is shown in Figure 3.6. It is maximum for a b jet with E_T of ~ 75 GeV and is approximately 43% with mistag rate of about 1%. The tagging rate estimated in Monte Carlo samples is slightly higher therefore a scale factor of 0.977 ± 0.035 is applied whenever a probability of tagging a heavy flavor jet is needed. [40]

3.7 Missing Transverse Energy

Neutrinos produced in $t\bar{t}$ decays escape the detector unobserved. However we can infer their presence using conservation of energy-momentum. In the collisions at the Tevatron we know that the total momentum of $p\bar{p}$ system in the direction along the beamline is zero. However we do not know event-by-event what is the momentum fraction carried by individual partons, therefore it is impossible to constrain the z

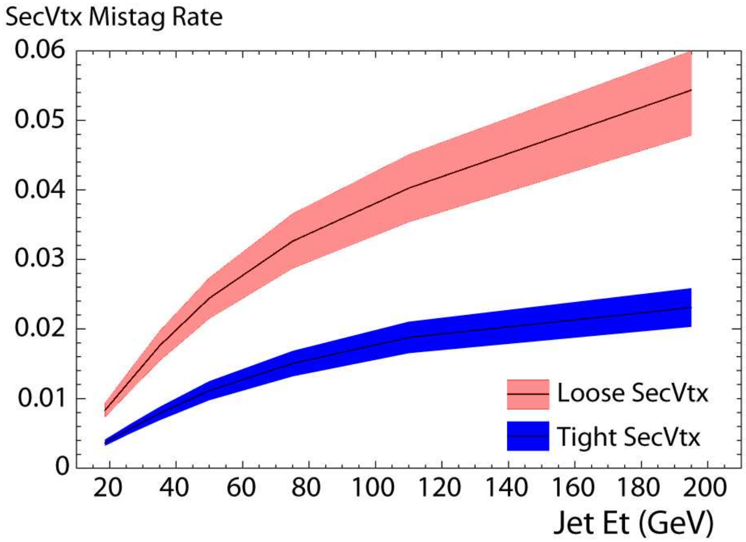


Figure 3.5: Rate of tagging light jets as b quark jets in $t\bar{t}$ events. The "loose" SecVtx tagging algorithm relaxes cuts applied on the secondary vertex in the "tight" tagging algorithm. The "tight" tagging is used in this analysis.

component of neutrino momenta. In the transverse plane the sum of momenta of all particles leaving the collision is 0. Negative vector sum of transverse energies in all towers will be equal to the total transverse momentum vector of escaping neutrinos. The transverse energies are calculated with respect to the $p\bar{p}$ interaction point.

We need to account for the fact that the muons escaping the collision leave very little energy in the calorimeter. If the raw energy of the calorimeter tower which the muon traverses were to be used in the missing energy calculation we would obtain spurious missing energy collinear with the muon direction. Instead we add back the energy deposited in the tower corresponding to the muon and add the negative p_T of the muon to the sum.

The resolution of jet energy measurement is expected to be better than the resolution of the measurement of energy resulting from addition of single towers. To better reconstruct the missing energy we remove the towers from the sum corresponding to

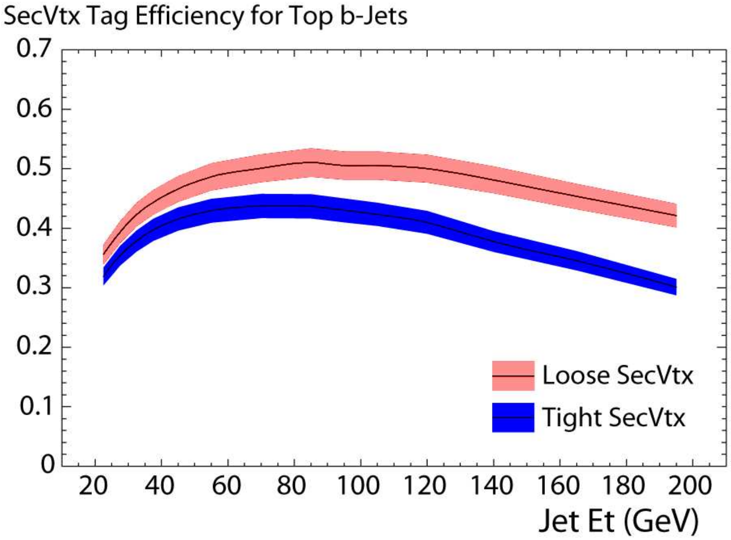


Figure 3.6: Efficiency of tagging b-quark jets in $t\bar{t}$ events.

jets. Next we construct jets which are corrected using the relative and absolute energy corrections. We then add negative vector sum of the transverse energies of these jets to the missing energy. The jets entering the calculations are not corrected for multiple interactions. Minimum bias interactions are expected to create no missing transverse energy, therefore not adding the energy deposited by multiple interaction within the jet cone to the missing energy sum would result in missing energy mismeasurement.

CHAPTER 4

EVENT SELECTION

4.1 Trigger Selection

Both decay channels under consideration are characterized by presence of at least one high transverse momentum lepton from the decay of the W boson. We use this property of $t\bar{t}$ events to design lepton triggers which will suppress presence of the multijet final states resulting from the high-rate QCD processes.

4.1.1 Muon Triggers

Data was collected using two muon trigger paths: the CMUP trigger path ($|\eta_{det}| < 0.6$) and the CMX trigger path ($0.6 < |\eta| < 1.0$).

The CMUP path requires at Level 1 presence of hits on both wires of a CMU trigger wire pair within a window of 124 ns of one another. A pattern of hits in 3 out of 4 CMP layers must be found. A track extrapolating (in $r - \phi$ plane) to both CMU and CMP hits must be found by XFT with transverse momentum of at least 4.09 GeV/c. At Level 2 the requirements have been changing through the data taking period as follows:

- Before October 2002: Level 2 requirements replicate the Level 1 requirements.
- October 2002 - April 2004: A track with transverse momentum of at least 8.34 GeV/c must be found by XFT.

- April 2004 - June 2005: A 8.34 GeV/c XFT track that extrapolates to the CMU and CMP hits must be found.
- June 2005 - March 2007: A 14.77 GeV/c XFT track that extrapolates to the CMU and CMP hits must be found.
- Since March 2007 the requirements have been the same as above but a combinations of prescales[1] and instantaneous luminosity thresholds has been employed.

The Level 3 requirements include a fully reconstructed muon with $p_T > 18$ GeV/c COT track. The track, when extrapolated out to the muon chambers must match the CMU stub within 10 cm and CMP stub within 20 cm in the $r - \phi$ plane.

The CMX path requires at Level 1 a hit on each wire of the CMX wire pair. Hits must occur within 124 ns of one another. An XFT track that extrapolates to the hits with transverse momentum of 8.34 GeV/c is required. After October 2002 a scintillator hit in the CSX system is also required. Similarly to the CMUP trigger path the Level 2 requirements have also been changing over time as follows:

- Before April 2004: Level 2 requirements replicate Level 1 requirements.
- April 2004 - June 2005: XFT track extrapolating to the CMX hits is required. Transverse momentum of the track must be at least 10.10 GeV/c.
- June 2005 - September 2006: The transverse momentum of the track was increased to 14.77 GeV/c
- September 2006 - October 2006: The Level 2 requirements were the same as above but the trigger was enabled at instantaneous luminosities below 200×10^{30} cm^{-2}s^{-1}

---

[1] When only one in n of the events that fired the trigger is actually accepted, that trigger is said to be “prescaled” by a factor n

- October 2006 - February 2007: XFT track with transverse momentum of 14.77 GeV/c is required. Stereo Layers of COT are required to match the axial layers in construction of the track. The track must match the CMX hits. In addition a Calorimeter cluster with transverse energy of at least 10 GeV is required in pseudorapidity range of $|\eta_{det}| < 3.6$
- Since January 2007: Trigger was enabled at instantaneous luminosities below 270×10^{30} cm^{-2}s^{-1}

At Level 3 a reconstructed muon with transverse momentum of 18 GeV/c with a CMX stub is required. The extrapolated track must match the stub within 10 cm in the $r - \phi$ plane.

4.1.2 Electron Trigger

The electron trigger requires at Level 1 a trigger tower with total transverse energy of at least 8 GeV and a track with transverse momentum of at least 8.34 GeV/c that extrapolates to the face of this tower. The trigger tower must lie in the pseudorapidity region of $|\eta_{det}| < 1.3$. To ensure that the selected object is in fact an electron a the ratio of energies in hadronic and electromagnetic compartments must be smaller than 1/8.

Electron trigger path at Level 2 requires that a cluster of trigger towers with transverse energy of 16 GeV is found within pseudorapidity region of $|\eta| < 1.3$. The hadronic and electromagnetic energy deposition ratio must be smaller than 1/8. Matching XFT track wit $p_T < 8$ GeV/c is required.

Level 3 requires a fully reconstructed central ($|\eta| < 1.0$) electron with transverse momentum of 18 GeV/c.

4.2 Offline Selection for the Dilepton Channel

4.2.1 Selection Criteria

The dilepton selection requires two leptons (electrons or muons). The transverse energy (momentum) of an electron (muon) is required to be greater than 20 GeV (GeV/c). One of the leptons has to be isolated and be a CEM electron or CMUP or CMX muon. The second lepton can be non-isolated CEM electron or CMUP, CMX, CMU or CMP muon. It can also be an isolated CEM or PHX electron or CMUP, CMX, CMU, CMP or CMIO muon. One of the leptons (isolated or non-isolated) has to correspond to the trigger fired in the event. We require that leptons have opposite signs.

We require two jets corrected to the particle level with $E_T > 15$ GeV and falling in the detector region $|\eta_{det}| < 2.5$. The same jets are used in missing transverse energy calculation but for this purpose the correction for multiple interactions is not applied (*cf.* 3.7).

Since there are two neutrinos escaping $t\bar{t}$ dilepton decays we expect large missing transverse energy and apply a cut $\not{E}_T > 25$ GeV. The final decay products of the $t\bar{t}$ pair have masses much smaller than the top quark mass so the sum of energies of all the decay products must be comparable to double of the top quark mass. We require $H_T > 200$ GeV, where H_T is a scalar sum of transverse momenta of leptons, transverse energies of jets and $\not{E}_T$. The H_T uses jets corrected to particle level in the sum.

In order to reject events originating from Z boson decay we require that if a dimuon or dielectron pair is present, the invariant mass of the pair cannot lie within the window of $76-106$ GeV/c^2. This requirement is applied only if missing transverse energy significance is smaller than $4.0\sqrt{\text{GeV}}$. We define $\not\!\!E_T$ significance as:

$$s_{\not E_T} = \frac{\not\!\!E_T}{\sqrt{E_T^{sum}}} \tag{4.2.1}$$

E_T^{sum} is defined as a scalar sum of the transverse energies of all towers. In a similar fashion to the calculation of $\not\!\!E_T$ we subtract energies of towers associated with muons and add muon transverse momenta. We also remove energies of towers associated with jets and add energies of the jets corrected to the particle level.

Missing transverse energy measurement may result from an experimental effect such as a jet or lepton depositing energy in an uninstrumented region. In this case the missing transverse energy vector will get a spurious contribution collinear with that jet or lepton. In addition some backgrounds such as Drell-Yan $\tau^+\tau^-$ production have real missing transverse energy collinear with jet or lepton associated with the τ decay. To reduce contamination from both of those effects we apply a cut $\not\!\!E_T > 50$ GeV if the angle in $r - \phi$ plane between the $\not\!\!E_T$ vector and any jet or lepton is less than 20^{o}

We divide the dilepton sample into tagged and non-tagged subsamples. We count the b tags only on the two highest energy jets in the event.

Accepted events are reconstructed using a Neutrino Weighting Algorithm described later (Sec. 5.1.1). We require that the reconstruction is successful and impose additional cuts on the reconstructed observables (*cf.* Sec. 6.1.1). The effect of these additional cuts is very small but it is taken into account in the calculation of expected number of signal and background events below.

4.2.2 Estimate of Signal and Background Events

The template method relies on correct background model. Therefore a relative normalization of all backgrounds is needed. Overall normalization of background is not required but improves the precision of the measurement. The calculation of the signal and background contributions has been performed by the CDF dilepton cross section group [43] and is only outlined here.

One of the main backgrounds for the dilepton channel is Drell-Yan process with two leptons in the final state. A Drell-Yan event can pass the selection requirements due to jet production from initial state radiation. The missing energy appears due to inaccurate measurement of jet or lepton energy. In the case of states with τ lepton pair in the final state the missing energy appears in the τ decay.. To estimate contribution of Drell-Yan events with dielectron or dimuon we employ a data-Monte Carlo technique. We select, in the data, events where the dilepton invariant mass lies within the range $76 - 106$ GeV$/c^2$. We subtract contribution form other backgrounds from the number of selected events. We estimate the ratio of the number of events outside the invariant mass window to the number of events inside using Monte Carlo samples. This ratio is then used to estimate Drell-Yan content outside the Z peak window. Similarly we obtain the ratio of events passing and failing the $s_{\not{E}_T}$ cut within the Z mass window in the Monte Carlo samples. This fraction combined with number of observed events in the data gives the estimate for number of Drell-Yan events inside the Z window and passing the significance cut.

Another important background for the dilepton channel comes from the production of WW, WZ and ZZ pairs. In such events dilepton pair may be produced in the final state together with real missing transverse energy from the W boson decay.

Contributions of the Diboson backgrounds, the Drell-Yan background where $\tau^+\tau^-$ pair is produced as well as expected $t\bar{t}$ contribution are estimated using Monte Carlo simulations. We calculate the acceptance in each dilepton category. This is necessary in order to convolute the trigger efficiencies and data Monte Carlo scale factors which are derived based on the lepton type as described in sections 3.2 and 3.3. Using the theoretical cross sections and the luminosity measurement we obtain the number of events expected for each of the processes. The expected yield of $t\bar{t}$ events is calculated using the theoretical cross section at $\mathrm{M_{top}}$ =175 GeV/c^2 of 6.7 pb. The estimate for $t\bar{t}$ contribution is not used in the likelihood fit, but it is used to check the method for bias (*cf.* Sec. 7.1).

A hadron collider produces QCD multijet events at high rate. Those events can fake a dilepton signature if we have one real lepton and a jet "faking" lepton signature. We measure the fake rate for each of the lepton categories (separately for isolated and non-isolated leptons) using samples collected with triggers that require a jet with transverse energy of 20 ("JET20"), 50 ("JET50"), 70 ("JET70") and 100 GeV ("JET100"). Estimating this process directly from Monte Carlo samples is considered not reliable due to difficulty in modelling jet hadronization and fragmentation. The calculation of fake rates requires defining "fakeable" leptons - objects which are likely to have come from quarks or gluons but resemble closely electrons or muons.

Fakeable electrons are required to have transverse energy of 20 GeV and ratio of hadronic to electromagnetic energy less than 0.125. For the central electrons we relax the quality requirements on the COT track. At least one of the cuts on the lateral shower development, track-shower maximum matching or the hadronic to electromagnetic energy ratio is then required to fail.

Fakeable muons have to have transverse momentum of 20 GeV/c and ratio of energy to track momentum less than 1.0. The track quality requirements are the same as for regular muon selection. We then require at least one of remaining criteria to fail. These are: presence of the stub (for CMUP, CMU, CMP and CMX muons), track-stub match, hadronic energy threshold and electromagnetic energy threshold.

In the JET50 sample we count number of leptons as well as number of fakeable leptons. The count is performed in 6 E_T bins. To remove trigger bias we remove events where the denominator lepton is within cone of R=0.4 from the highest energetic jet. We estimate using Monte Carlo samples how many leptons and fakeable leptons are expected to have come from W boson decay and subtract those counts. Therefore we have a count of reconstructed leptons and fakeable leptons originated by quarks. The ratio of the two numbers is interpreted as a probability of jet faking a lepton. We perform the same calculation using other jet samples and assign 30% systematic uncertainty based on the observed differences.

Using the probabilities calculated above we estimate the number of events with a fake lepton in our sample. In the data sample triggered with central electrons and muons (the same sample as the one used to make final measurement) we select events with one lepton and at least one fakeable. We substitute the fakeable for the second lepton in an event. If the fakeable is an electron and it is within cone of R=0.4 from a jet we remove this jet from jet collection. We count events which after such substitution pass the dilepton selection requirements. We assign a weight to each event corresponding to the fake rate for the considered fakeable. If there are multiple denominator leptons in the event the event can enter the count several times. The sum gives number of expected events with a fake lepton. Sums are performed separately in non-tagged and tagged samples

The background estimate for the Drell-Yan and diboson backgrounds as well as signal obtained as described above is divided into an estimate for the untagged and tagged subsamples using Monte Carlo samples. A probability for each of the two most energetic jets to be tagged is evaluated. If we find that a given jet is originated by a light flavor quark then probability of the jet to receive a b tag is equal to mistag

Table 4.1: Expected event yield for the dilepton selection after all cuts.

	non-tagged	tagged
Diboson	9.1 ±2.2	0.3 ±0.1
Drell-Yan	16.0 ±2.5	0.9 ±0.1
Fakes	19.3 ±5.6	2.7 ±1.0
Total Background	44.3 ±7.0	3.9 ±1.0
$t\bar{t}$ (6.7 pb)	40.1 ±3.1	55.8 ±4.2
Observed	83	61

probability evaluated using the mistag matrix. However if jet originates from a heavy flavor quark and if it is tagged then the probability is assigned as a b tag scale factor (described in section 3.6). If a heavy flavor jet is not tagged we assign probability of 0 for this jet to be tagged. Once the probabilities of each jet to be tagged is known we evaluate the probability for the event to have no tags or at least one tag. Sum of probabilities to obtain no tags gives the estimate on the 0-tag sample and sum of probabilities to find tags gives the estimate in the 1-tag sample.

The results of the calculations described above are summarized in Table 4.1. Effects of reconstruction efficiency and boundary cut are included.

4.3 Offline Selection for the Lepton + Jets Channel

4.3.1 Selection Criteria

To select events resulting from a $t\bar{t}$ decay into lepton + jets channel we impose somewhat stricter requirements than in the dilepton channel in terms of types of leptons and jet thresholds used. This is possible due to higher statistics in this channel. We require one electron or muon with transverse energy (or momentum) of 20 GeV. The only allowed electron category is CEM. Muon categories allowed are CMUP and CMX. Strictly one lepton is allowed to remove overlap with the dilepton sample. In this dilepton veto we look for the leptons of other categories also including stubless muons and plug electrons.

At least one b tag is required. We divide the events based on numbers of b tags present into 1-tag and 2-tag subsamples. In 1-tag sample we require four jets with transverse energy greater than 20 GeV and falling in detector region of $|\eta_{det}| < 2.0$. Events with a fifth jet are disallowed if that jet has transverse energy greater than 20 GeV. Since the background contamination of 2-tag sample is expected to be very small we relax the jet requirements to increase the event count of this sample. We require presence of 3 jets with $E_T > 20$ GeV and $|\eta_{det}| < 2.0$. The fourth jet has to have $E_T > 12$ GeV and $|\eta_{det}| < 2.4$. No requirements on extra jets are made. Only the four most energetic jets are considered for b tagging. All jet threshold requirements are made as in the dilepton sample on jets corrected to particle level.

In addition to the lepton and jet requirements we require missing transverse energy of 20 GeV. Missing transverse energy is corrected for jets and muons as described in Sec. 3.7. Jets with $E_T > 12$ GeV and $|\eta_{det}| < 2.4$ are used to calculate the correction. This is slightly different than in the dilepton sample so that the requirements are consistent between this analysis and the respective $t\bar{t}$ cross section measurements.

As described in Sec. 5.2.1 accepted events are reconstructed using a χ^2 fitter. A cut of $\chi^2 < 9.0$ is imposed to reject poorly reconstructed events. In addition as described in Sec. 6.1.1 boundaries are imposed on the reconstructed observables. Efficiency for passing the χ^2 cut is 66% and 87% in doubly tagged and single tagged $t\bar{t}$ events. Boundary cut is about 99% efficient for all signal events. Those efficiencies diminish to as low as ∼40% for the boundary cut and ∼50% for the χ^2 cut in some backgrounds.

4.3.2 Estimate of Signal and Background Events

Estimation of background and signal contribution to the lepton + jets sample relies on a mixed Monte Carlo/data approach. Estimating background contribution form data is generally preferable since it is not exposed to uncertainties of Monte Carlo modelling. However some background contributions are too small and no convenient ways of extrapolating their contribution from some control region to a signal region exist. This is also true for the signal contribution. The method is sequential: we successively estimate background components and the estimate of the next background component relies on the estimate derived previously. We obtain the background contribution estimates in the sample before application of b tagging requirements ("pre-tag" sample) and when the tagging requirements are applied ("tagged" samples). The machinery used to obtain the estimates, so called so called "Method II" [40], is provided by the CDF top group. An overview of the method follows.

The event selection employed by Method II is slightly different from the selection employed in this analysis. One small difference is that Method II allows b tags on 5^{th} and lower jets whereas we allow b tags on only 4 leading jets. In addition we apply the χ^2 and boundary cuts mentioned earlier. We account for those differences by scaling Method II estimates [18].

The contributions estimated first are WW, WZ, ZZ, single top production is s and t channels and $t\bar{t}$ production. Each of the processes contributes to the final sample because it can produce a lepton and a neutrino. Contribution of those processes is estimated directly from Monte Carlo samples both in pre-tag, 1-tag and 2-tag samples. The estimate $N^i_{process}$ is:

$$N^i_{process} = \sigma^{theory} \times \epsilon^{MC,i} \times L, \tag{4.3.1}$$

where i signifies pretagged, 1-tag or 2-tag sample, σ^{theory} is the theoretical cross section for this process, L is the integrated luminosity and $\epsilon^{MC,i}$ is selection efficiency obtained from Monte Carlo sample. The efficiency term includes trigger efficiencies and data - Monte Carlo lepton identification scale factors (*cf.* Sections 3.2, 3.3). In case of tagged samples probabilities to tag light flavor jets are evaluated using the mistag matrix and b tagging scale factor is applied to tagged heavy flavor jets (*cf.* Sec. 3.6).

A potential source of leptons other than W and Z boson is QCD multijet production, where the lepton is a fake (a jet mimicking a lepton) or comes from semileptonic decay of a heavy quark. The $\not\!\!E_T$ spectrum of events coming from QCD production is well separated from $\not\!\!E_T$ spectrum of events with W bosons. This is because in QCD events a non-zero $\not\!\!E_T$ is due to purely experimental effects such as jets entering an uninstrumented region of the calorimeter or fake muon's momentum being added to the missing energy sum. The QCD background is modeled in the data by events passing all cuts where instead of a lepton an electron-like (so called "antielectron") object is required. Antielectrons pass all the kinematic cuts that are imposed on the electrons, but fail shower development or track quality cuts. In an alternative, high statistics model for the QCD background, data events are used where the isolation cut on the charged lepton is inverted, requiring isolation greater than 0.2. In the tagged sample only the antielectron template is used for estimate of QCD contribution, but we use the non-isolated lepton model in the later stages of the analysis. To obtain the QCD estimate in the pretagged as well as in the tagged sample the QCD $\not\!\!E_T$ shape and $\not\!\!E_T$ shape from the W+jets Monte Carlo samples are fit to the $\not\!\!E_T$ shape in the data obtaining the relative normalization of the two. The QCD estimate is then the fitted fraction times the total number of events in the sample.

Knowing the QCD contribution as well as the diboson and top contributions we attribute the remainder of the pretag sample to W+jets production: we simply subtract the contributions from the other sources from the total number of observed events.

To estimate the number of W+heavy flavor sample in the tagged samples we start with the total W+jets estimate in the pretagged sample. Using the Monte Carlo W+jets models we estimate the fraction of events with reconstructed jets matched to heavy flavor quarks. We then estimate the tagging probability for those events. This is done by first evaluating probability of tagging each jet. This probability is 0 if the jet is not tagged and equal to the b tagging scale factor if jet is tagged. Probability of having one tag or two or more tags in an event is then calculated using the jet tagging probabilities. Therefore now we have fraction of W+jets events that contain heavy flavor jets and b tags. Multiplying this fraction by the W+jets estimate in the pretagged sample gives the W+ heavy flavor estimate in the tagged samples.

The remaining piece is estimating W+light flavor component in the tagged samples. We again start with the W+jets estimate in the pretagged sample. We find the fraction of events with light jets only and evaluate probabilities of tagging those jets using the mistag matrix. Based on the jet tagging probabilities we calculate probability of each event being tagged and overall tagging probability for this sample. using this probability and W+jets estimate in the pretagged sample we estimate W+light flavor contribution to the tagged samples.

The expected composition of the lepton + jets sample is shown in Table 4.2. As in the dilepton channel estimate we use 6.7 pb for the cross section of $t\bar{t}$ production.

Table 4.2: Expected event yield for the lepton + jets selection after all cuts.

	1-tag	2-tag
$Wb\bar{b}$	9.1±3.7	2.1±0.9
$Wc\bar{c}$, Wc	8.3±3.4	0.5±0.3
W (mistags)	10.4±2.3	0.2±0.1
Single top	2.0±0.1	0.7±0.1
Diboson	2.4±0.2	0.21±0.02
QCD	10.4±8.7	0.3±1.6
Total Background	42.7 ± 12.5	4.2 ± 1.9
$t\bar{t}$ (6.7 pb)	156.7±21.1	76.6±12.0
Observed	233	99

CHAPTER 5

EVENT RECONSTRUCTION

Each of the selected events presents us with an overwhelming amount of information. We have access to energies of towers in the calorimeter, positions of hits in the tracking, muon and shower maximum detectors. This information is synthesized to form higher level objects: electrons, muons and jets as well as the missing transverse energy. Information contained in these objects has to be related to the mass of the top quark. In the dilepton channel we have 18 numbers and in the lepton + jets channel we have 22 numbers describing the event. Such large number of variables would still be impossible to analyze statistically and must be somehow further reduced. We want to construct a variable whose distribution varies substantially as a function of M_{top}. Naturally the invariant mass of the top quark is expected to be the best choice. We estimate it in each event differently depending on which decay channel the event belongs to. In the following sections we describe the procedures employed in each of the two channels. We also describe other event variables used to improve the measurement precision and to estimate *in situ* the jet energy scale.

5.1 Event Reconstruction in the Dilepton Channel

5.1.1 Neutrino Weighting Algorithm

In the $t\bar{t}$ →dilepton system we do not have enough information to reconstruct the masses of the decaying top quarks. In the detector we measure four-momenta of jets and leptons and an overall imbalance in transverse energy $\not{E}_T$. However in the dilepton channel we have 2 escaping neutrinos. This means that even when we use all the knowledge of the event such as masses of the particles in the decay cascade and we assume that masses of the two top quarks are the same, we still lack one constraint to reconstruct the four-vectors of final state partons. We need to integrate over the unknown quantities taking the probability density functions from the Monte Carlo simulations. In this method we scan over the top quark mass m_t and integrate over pseudorapidities η_1 and η_2 of the two neutrinos at each point in the scan. As inputs we use jets corrected up to parton level and lepton momenta. Knowing the top quark mass and neutrino pseudorapidities we can solve for the neutrino transverse momenta. This allows us to weight the m_t hypothesis by comparing the solution to the measured $\not{E}_T$. The hypothesis that yields the highest weight will be taken as the reconstructed top quark mass m_t^{NWA}. The algorithm has been used previously in [44, 45, 46], and is adapted for use in this analysis with minor modifications.

Since jet charge is not reconstructed there are two ways to form b quark jet - lepton pairs. Assuming that we know the correct pairing and given the value of the top quark and the neutrino pseudorapidities we write the invariant mass of the top quark and W boson:

$$m_t^2 = (\boldsymbol{l} + \boldsymbol{\nu} + \boldsymbol{b})^2 \tag{5.1.1}$$

$$m_W^2 = (\boldsymbol{l} + \boldsymbol{\nu})^2 \tag{5.1.2}$$

Where $\boldsymbol{l}$, $\boldsymbol{\nu}$, and $\boldsymbol{b}$ are the four-vectors of the lepton, neutrino and the b quark respectively. We can rearrange the two equations above and form two relativistic invariants:

$$B \equiv 2\boldsymbol{b\nu} = m_t^2 - m_W^2 - m_b^2 - 2\boldsymbol{bl} \tag{5.1.3}$$

$$L \equiv 2\boldsymbol{l\nu} = m_W^2 - m_l^2 \tag{5.1.4}$$

We assumed $m_\nu = 0$ in the last equation. We calculate the numerical values of these invariants using world average values for the masses of the particles and the four-vectors measured in the detector where jets have been corrected to the particle level. We express the four-vector of the neutrino in terms of its transverse momentum p_T, pseudorapidity η and the azimuthal angle ϕ as:

$$\boldsymbol{\nu} = (p_T \cos(\phi), p_T \sin(\phi), p_T \sinh(\eta), p_T \cosh(\eta)) \tag{5.1.5}$$

We expand the products of b quark and lepton four-momenta with the neutrino momentum in Equations 5.1.3 and 5.1.4 to obtain:

$$B = 2p_T(E^b cosh(\eta) - p_x^b cos(\phi) - p_y^b sin(\phi) - p_z^b sinh(\eta)) \tag{5.1.6}$$

$$L = 2p_T(E^l cosh(\eta) - p_x^l cos(\phi) - p_y^l sin(\phi) - p_z^l sinh(\eta)) \tag{5.1.7}$$

Now we isolate p_T from Equation 5.1.7 and insert into Equation 5.1.6. Upon rearranging the terms we obtain:

$$\begin{aligned}(LE^b - BE^l)cosh(\eta) - (Lp_z^b - Bp_z^l)sinh(\eta) = \\ (Lp_x^b - Bp_x^l)cos(\phi) + (Lp_y^b - Bp_y^l)sin(\phi)\end{aligned} \tag{5.1.8}$$

Now define:

$$\beta \equiv \frac{(LE^b - BE^l)cosh(\eta) - (Lp_z^b - Bp_z^l)sinh(\eta)}{\sqrt{(Lp_x^b - Bp_x^l)^2 + (Lp_y^b - Bp_y^l)^2}} \tag{5.1.9}$$

We can also define:

$$\cos(\alpha) \equiv \frac{(Lp_x^b - Bp_x^l)}{\sqrt{(Lp_x^b - Bp_x^l)^2 + (Lp_y^b - Bp_y^l)^2}} \tag{5.1.10}$$

$$\sin(\alpha) \equiv \frac{(Lp_y^b - Bp_y^l)}{\sqrt{(Lp_x^b - Bp_x^l)^2 + (Lp_y^b - Bp_y^l)^2}} \tag{5.1.11}$$

We divide Equation 5.1.8 by the factor $\sqrt{(Lp_x^b - Bp_x^l)^2 + (Lp_y^b - Bp_y^l)^2}$ and using the definitions above we obtain:

$$\beta = \cos(\alpha)cos(\phi) + \sin(\alpha)sin(\phi) = \cos(\phi - \alpha) \tag{5.1.12}$$

If $|\beta| < 1$ Equation 5.1.12 has two solutions for ϕ:

$$\phi = \alpha \pm \arccos(\beta) \tag{5.1.13}$$

Where the $(0, \pi)$ branch of arccos was chosen. In practice we only need the values $sin(\phi)$ and $cos(\phi)$:

$$sin(\phi) = \sin(\alpha)\beta \pm \cos(\alpha)\sqrt{1 - \beta^2} \tag{5.1.14}$$

$$cos(\phi) = \cos(\alpha)\beta \mp \sin(\alpha)\sqrt{1 - \beta^2} \tag{5.1.15}$$

Transverse momentum of the neutrino is obtained from Equation 5.1.7 and we multiply by $cos(\phi)$ and $sin(\phi)$ to obtain the x and y components of neutrino momentum. The same procedure is repeated to solve for the x and y components of momentum of the other neutrino in the event. Thus we have four possible solutions.

For each jet-to-quark assignment i and each solution j we form a weight $w_{i,j}(m_t, \eta_1, \eta_2)$ by comparing the solution for the transverse momenta of the neutrinos with the measured $\not\!\!E_T$ using a Gaussian weight:

$$w_{i,j}(m_t, \eta_1, \eta_2) = exp(-\frac{\not\!\!E_{Tx} - p_x^{\nu} - p_x^{\bar{\nu}}}{2\sigma^2_{\not E_T}})exp(-\frac{\not\!\!E_{Ty} - p_y^{\nu} - p_y^{\bar{\nu}}}{2\sigma^2_{\not E_T}}) \tag{5.1.16}$$

The resolution $\sigma_{\not E_T} = 19.2$ GeV is extracted from a $t\bar{t}$ Monte Carlo sample generated at $\mathrm{M_{top}}$ =175 GeV/c^2. The value of the x and y components of the sum of neutrino momenta obtained from the Monte Carlo generator information are compared to the x and y component of $\not\!\!E_T$.

To obtain the event weight that depends only on the top quark mass we sum over the jet-to-quark assignments i, solutions j and integrate over the neutrino momenta η_1, η_2 as follows:

$$w(m_t) = \sum_{i=1}^{2} \iint d\eta_1\, d\eta_2\, g(\eta_1) g(\eta_2) \sum_{j=1}^{4} w_{i,j}(m_t, \eta_1, \eta_2) \tag{5.1.17}$$

The function g is a Gaussian with 0 mean and width of 1 and describes the neutrino pseudorapidity distribution in $t\bar{t}$ events. We extract the width of the Gaussian from the generator information in the $\mathrm{M_{top}}$ =175 GeV/c^2 Monte Carlo sample. The integration over the neutrino pseudorapidities is numerically approximated as a sum over a grid with spacing of 0.15 and bounded by $(-3.0, 3.0)$.

Scan in m_t is performed from m_t = 80 GeV/c^2 to m_t = 380 GeV/c^2 in steps of 3 GeV/c^2. A maximum is found in the total weight and in weights corresponding to each of the two jet-to-quark assignments. Scan is then iteratively repeated near the maxima with successively smaller steps until the step size reaches 0.03 GeV/c^2. The

m_t value maximizing the weight is taken as the reconstructed top quark mass m_t^{NWA} in this event.

Distributions of m_t^{NWA} for non-tagged and tagged fully simulated Monte Carlo $t\bar{t}$ events with input top quark masses of 160, 170 and 180 GeV/c^2 passing the dilepton selection are shown in Figure 5.1. The event reconstruction in the dilepton channel is imperfect but there is a clear dependence of the shape of the m_t^{NWA} distribution on the input top quark mass. The mean of the distribution rises linearly by 0.6 GeV/c^2 for a 1.0 GeV/c^2 rise in the input top quark mass. The RMS of the distribution is 19% of it's mean for the non-tagged sample at the input top quark mass of 170 GeV/c^2. The distribution is about 10% narrower in the tagged sample. Relative resolution of reconstructed top quark mass improves slightly with increase in the input mass.

5.1.2 H_T

The momenta of the final state particles in a $t\bar{t}$ decay event are directly linked to the invariant mass of the top quarks in an event, therefore H_T- a linear sum of the jet and lepton transverse momenta and $\not\!\!E_T$ has a strong dependence on the top quark mass as seen in Fig. 5.2 where distributions of H_T in $t\bar{t}$ Monte Carlo samples with three generated top quark masses are shown. To calculate the H_T variable we use jets corrected to the parton level. The same cuts are made on jets used for the H_T calculation as during the selection. Jet has to have pseudorapidity $|\eta_{det}| < 2.5$ and $E_T > 15$ GeV at particle level. At the input top quark mass of 170 GeV/c^2 the mean of the distribution in approximately 330 GeV and rises linearly with the input top quark mass with constant of proportionally of 1.2. For the input top quark mass of 170 GeV/c^2 the RMS of the distribution is 84 GeV for the non-tagged sample and 80 GeV and rises slightly with increase in the input top quark mass.

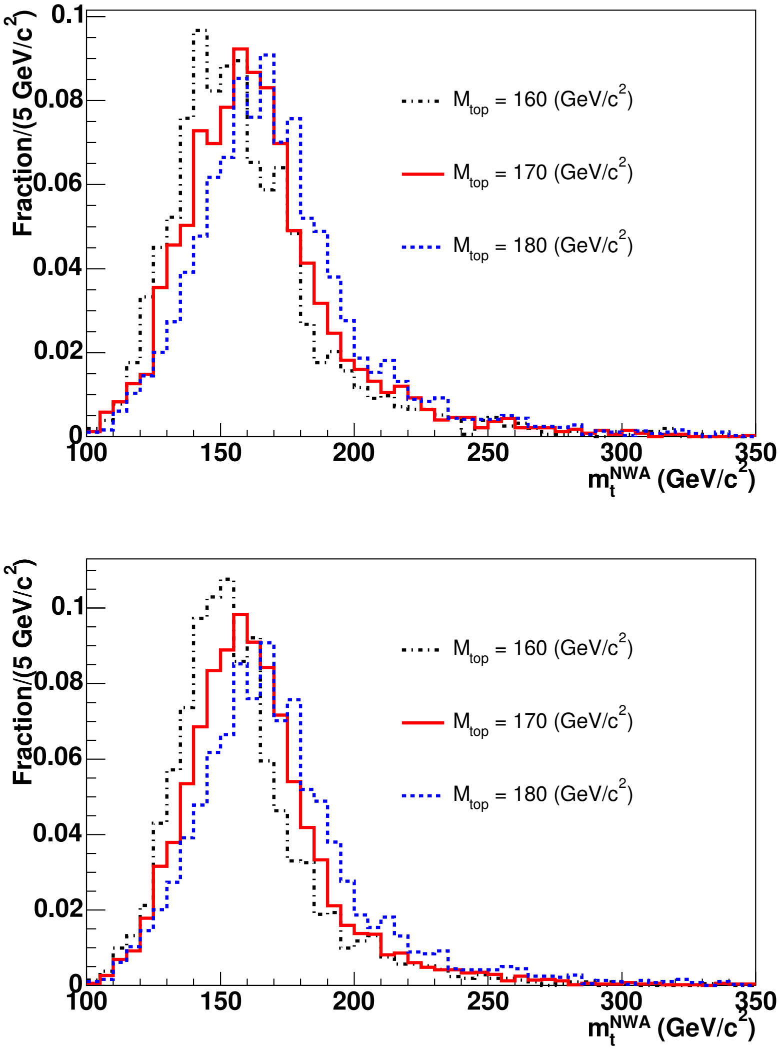


Figure 5.1: Output of the NWA algorithm for Monte Carlo events with three input values of M_{top} passing the 0-tag (top) and tagged selection (bottom).

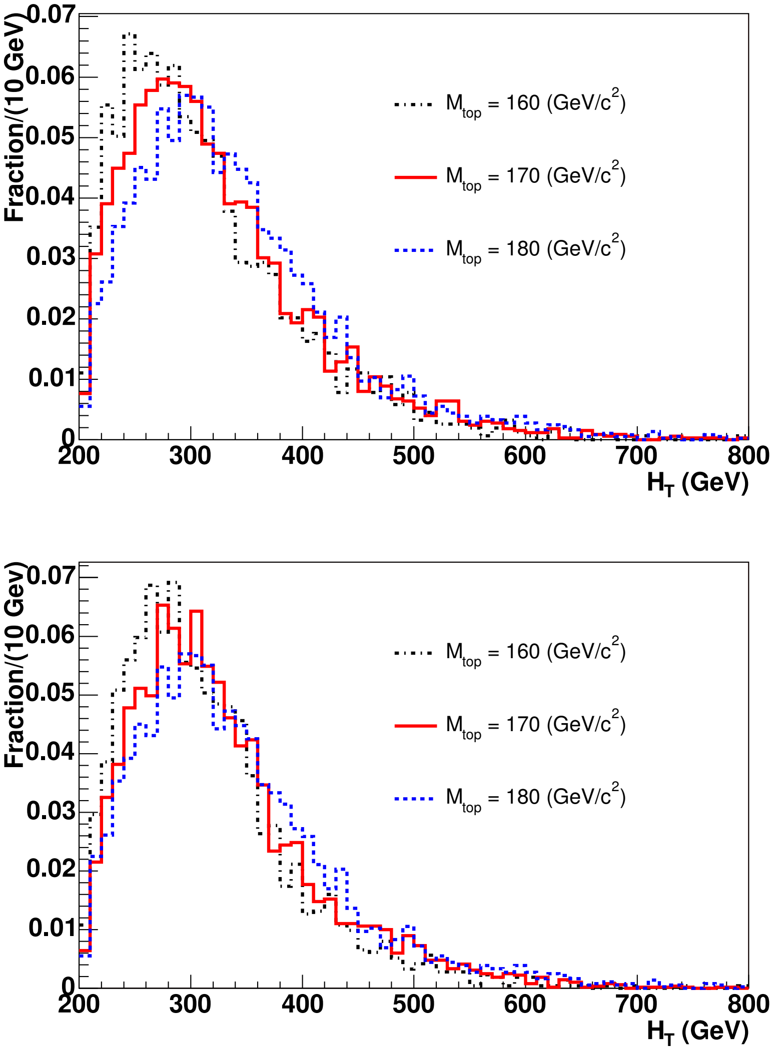


Figure 5.2: H_T for Monte Carlo events with three values of input $\mathrm{M_{top}}$ passing the 0-tag (top) and tagged selection (bottom).

5.2 Event Reconstruction in the Lepton + Jets Channel

5.2.1 Top Mass Reconstruction: the χ^2 Fitter

In contrast to the dilepton channel in the lepton + jets channel we measure enough quantities to reconstruct four-momenta of all particles escaping the collision. We get the four-momenta of the quarks directly from the measurement of jets in the calorimeter and knowledge of the primary vertex location. Energy of an electron is measured by the calorimeter. Since we know it's polar angle and we can assume that it's massless we obtain the three components of it's momentum. Similarly for a muon it's transverse momentum is measured in the tracker and z component of momentum we get from the polar angle. Assuming zero mass we get muon's energy. Since the $p\bar{p}$ system before the collision has no momentum in the transverse plane we can assign the missing transverse imbalance in x and y directions as the components of the neutrino momentum. The only component missing is the z component of neutrino momentum. However we have additional information: sum of the two light jets' four-vectors must have the invariant mass of a W; the same is true for sum of lepton and neutrino four-vectors. We can also assume that the masses of t and $\bar{t}$ are the same (but unknown). We find the best hypothesis for the four momenta of the particles, and therefore the top quark mass in the event (m_t^{reco}) using a χ^2 fit. In subsequent sections we describe the χ^2 expression and the inputs of the fitter and relations between fitted quantities.

The χ^2 Expression

The expression minimized using MINUIT package [47] is:

$$\begin{aligned}\chi^2 &= \sum_{i=\ell,4\,\text{jets}} \frac{(p_T^{i,\text{fit}} - p_T^{i,\text{meas}})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(U_j^{\text{fit}} - U_j^{\text{meas}})^2}{\sigma_j^2} \\ &+ \frac{(M_{jj} - M_W)^2}{\Gamma_W^2} + \frac{(M_{\ell\nu} - M_W)^2}{\Gamma_W^2} \\ &+ \frac{(M_{bjj} - m_t^{\text{reco}})^2}{\Gamma_t^2} + \frac{(M_{b\ell\nu} - m_t^{\text{reco}})^2}{\Gamma_t^2} \qquad (5.2.1)\end{aligned}$$

The parameters of the fit are:

- $p_T^{i,\text{fit}}$ - transverse momenta of the lepton and jets
- U_x^{fit}, U_y^{fit} - unclustered energy in x and y directions. This is energy that is not associated with jets or leptons from $t\bar{t}$ decay.
- m_t^{reco} - the reconstructed top quark mass
- $p_z^{\nu,\text{fit}}$ - the z component of the neutrino momentum.

The first term in the χ^2 expression constrains the lepton transverse momentum and jet transverse momenta to their measured values within experimental resolution σ_i. The angles of jets and leptons are assumed to be measured perfectly, therefore the full four-vector is obtained straightforwardly from p_T.

The second term constrains the calorimeter energy not associated with the $t\bar{t}$ decay products to it's measured value. In conjunction with the lepton and jet momenta unclustered energy gives the neutrino transverse momentum.

Third term constrains the four-vectors of the light jet quarks to the W boson mass. Similarly the fourth term constrains the lepton and neutrino four-vectors. The

Breit-Wigner W boson width is used in the constraint even though the constraint is Gaussian.

The last two terms constrain the products of leptonically decaying top quark to form the same invariant mass (within the Breit-Wigner width) as the products of the hadronically decaying top. The m_t^{reco} variable floats in the fit and is taken as the reconstructed top quark mass.

An immediate problem is that it is unknown which quarks from the $t\bar{t}$ decay cause which jets observed in the detector. In the 2-tag sample we know which jets are from b quark decay. We don't know however which b jet to pair with the dijet system of the hadronically decaying W. In 1-tag sample the situation is even more complicated giving us possible 6 jet to quark assignments. Another difficulty comes from the z component of the neutrino momentum. We can solve for it knowing $\not{E}_T$, lepton momentum and W mass, but two solutions exist. We run the fit for each possible jet-to-parton assignment twice - initializing the z component of neutrino momentum to each of the two solutions. The m_t^{reco} for the combination resulting in the lowest χ^2 is kept. We reject events with $\chi^2 > 9.0$.

Lepton and Jets

The lepton transverse momenta supplied as input to the χ^2 fitter are simply the reconstructed p_T or E_T of muon or electron as described in Chapter 3. If the lepton is an electron the resolution on it's transverse momentum is equal to the calorimeter resolution (*cf.* Sec. 2.2.2)

$$\sigma_e = p_T^{e,\text{meas}} \times \sqrt{\left(\frac{0.135}{p_T^{e,\text{meas}}}\right)^2 + 0.02^2} \tag{5.2.2}$$

If the lepton is a muon it's resolution is set to:

$$\sigma_{\mu} = 0.0011 \times (p_T^{\mu,\mathrm{meas}})^2, \tag{5.2.3}$$

given by the tracking p_T resolution.

Differences between light flavor jets and b jets are not taken into account by the generic jet corrections described in Sec. 3.5. Such differences may arise due to different fragmentation and semileptonic decays. Explicitly treating b jets and light flavor jets differently may result in better resolution of the reconstructed top quark mass. We therefore derive "top-specific" corrections using PYTHIA $t\bar{t}$ samples generated with top quark mass in the range $160 - 180$ GeV/c^2. Jets are initially corrected to particle level as described in Sec. 3.5, and a response:

$$S = \frac{p_T^{parton}}{p_T^{jet}} - 1 \tag{5.2.4}$$

is calculated for light flavor jets and b jets separately. Jets are required to match the originating partons within R of 0.2. Most probable value (MPV) of the response is found in 6 bins of $|\eta|$ and 10 bins of p_T. The response is then fit to an exponential in p_T where the parameters of the exponential are allowed to vary as a function of η. The correction is then found by rearranging Equation 5.2.4:

$$p_T^{parton} = (S_{MPV} + 1) \times p_T^{jet} \tag{5.2.5}$$

The correction for light jets are about 30% for central jet with jet momentum of 20 GeV/c and lower with increasing $|\eta|$. The correction decays rapidly as a function of momentum and for 60 GeV/c jets it's of the order of several %. Correction for b

jets are few percent higher and have the same decaying behavior.

Resolution terms are derived using the same approach but instead of the MPV width of the response is fit as a function of p_T and η. The width of the response is essentially independent of flavor and pseudorapidity of the quark. For jets at 20 GeV$/c$ the response width is around 0.2 and decays asymptotically to about 0.1. A detailed description of the top-specific corrections is given in [18].

Unclustered Energy

An assumption made by the fitter is that the four leading jets come from the $t\bar{t}$ decay. Additional jets, allowed in 2-tag sample, are assumed to have come from initial state radiation. We ignore the possibility of jet originating from initial state radiation or other effects having higher energy. It is possible that extra jets in an event originates from final state radiation and ought to be included in the reconstruction. For practical reasons this effect is also ignored. The energy contained in those additional jets and energy not belonging to any jets or the lepton is relevant in calculation of the neutrino transverse momentum in the fitter. Let us define "Additional jets" as the jets that fall in the pseudorapidity region of $|\eta_{det}| < 2.4$ and have $E_T > 12$ GeV after correction to particle level. We exclude the four leading jets ("$t\bar{t}$ jets") from this set. Now define "raw" unclustered energy $\vec{U}_T^{raw}$ as the (2-d vector) sum of transverse energies of all towers in the detector excluding the towers belonging to $t\bar{t}$ jets, additional jets and the lepton. We scale the raw unclustered energy by a factor of 1.4 which is approximately the particle level correction applied to a jet with $E_T = 8$ GeV. We then add back the transverse energies of the additional jets. The unclustered transverse energy in the fit is then defined as:

$$\vec{U}_T^{meas} = 1.4\vec{U}_T^{raw} + \sum_{\text{additional jets}} \vec{E}_T^{\text{jet}} + \sum_{t\bar{t}\text{ jets}} \vec{E}_T^{\text{jet, UE}} + \sum_{t\bar{t}\text{ jets}} \vec{E}_T^{\text{jet,OOC}} \tag{5.2.6}$$

Where $\vec{E}_T^{jet,UE}$ is the transverse energy of a jet corrected for all effects up to Underlying Energy flow and $\vec{E}_T^{jet,OOC}$ is the transverse energy of a jet corrected for all effects. As described in Sec. 3.5 the corrections are applied sequentially. Therefore the last two terms have the effect of subtracting from the unclustered energy the out-of-cone energy flow of the $t\bar{t}$ jets, to avoid counting this energy twice in p_T^{jet} and $\vec{U}_T^{meas}$.

The transverse momentum of the neutrino is then calculated at each step in the minimization:

$$p_{x,y}^{\nu} = -\left(\sum_{t\bar{t}\text{ jets}} p_{x,y}^{\text{jet, fit}} + p_{x,y}^{\text{lepton, fit}} + U_{x,y}^{\text{fit}} \right) \tag{5.2.7}$$

The z momentum of the neutrino is constrained only through the constraints on the lepton - neutrino system in the fit.

The resolution on the unclustered energy is set to $0.4 \times \sqrt{E_T^{unclustered}}$, where $E_T^{unclustered}$ is a scalar sum of transverse energies of towers in the calorimeter excluding the towers associated with $t\bar{t}$ jets. That sum is also multiplied by a factor of 1.4 to roughly mimic the particle level correction. This term is obtained from studies of minimum bias events. Other approaches in calculating the unclustered energy and it's resolution were tried and had negligible effect on the reconstructed top quark mass resolution. [48].

Results of the Top Mass Reconstruction

Distributions of $m_t^{\rm reco}$ for the 1-tag and 2-tag subsamples extracted from Monte Carlo samples are shown in Fig. 5.3. Plotted are distributions for three values of the top quark mass. The distribution in 2-tag sample is about 10% narrower than in the 1-tag sample. This is because the correct jet-quark assignment is aided by presence of additional b tag. In the 1-tag sample the mean $m_t^{\rm reco}$ grows linearly by about 0.6 GeV/c^2 for 1 GeV/c^2 increase in the input top quark mass, and approximately 10% faster in the 2-tag sample.

5.2.2 The Jet Energy Scale Parameter $\Delta_{\rm JES}$

As we have seen in Sec. 3.5.6 the jet energy scale is known within ~ 3 to $\sim 10\%$. This systematic uncertainty on the momenta of the jets is derived in data and Monte Carlo samples of various physical processes such as dijet production or γ+jet production. However information about all the jet energy scale effects is also present in the $t\bar{t}$ events. To reduce the overall jet energy scale systematic uncertainty we use the hadronically decaying W boson in lepton + jets events. In order to do so we must construct a nuisance parameter describing the possible difference in the jet energy scale calibration between Monte Carlo samples and data. We define the parameter $\Delta_{\rm JES}$ to be the offset of the JES calibration from the nominal in σ_c - a unit of the total systematic uncertainty on the JES calibration. Our *a priori* estimate of this parameter is $\Delta_{\rm JES} = 0 \pm 1\sigma_c$ since the Monte Carlo simulation is tuned to mimick the physics processes and instrumental effects with no bias. The possible remaining unknown offset between Monte Carlo and data is quantified as the systematic uncertainty. A positive $\Delta_{\rm JES}$ parameter indicates that the corrected jet momenta in Monte Carlo samples need to be increased.

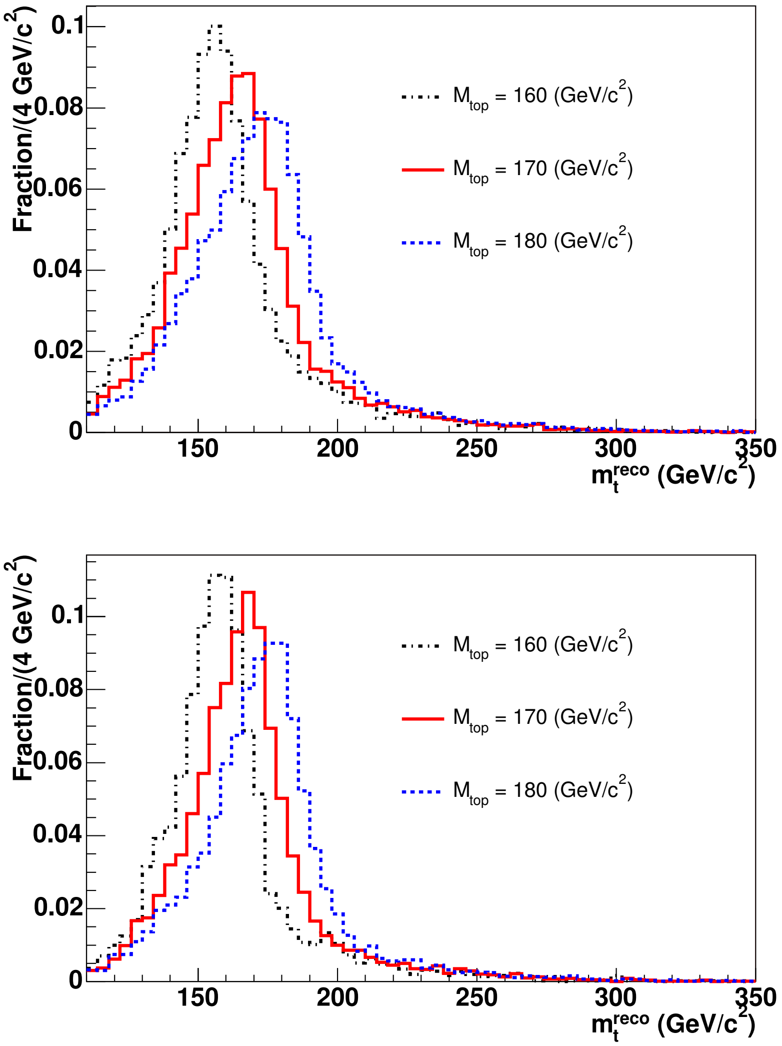


Figure 5.3: Output of the χ^2 fitter for Monte Carlo events with three input values of M_{top} passing the 1-tag (top) and 2-tag selection (bottom).

5.2.3 Dijet Mass Reconstruction

The variable sensitive to JES is the invariant mass of hadronically decaying W. We reconstruct it using the jets corrected using the top specific corrections. As in the reconstruction of the top quark mass, we assume the simplest scenario for the $t\bar{t}$ decay and search among the leading 4 jets for the two coming from W boson decay. In the

2-tag sample we take the invariant mass of the sum of the two jets which do not have b tags. In the 1-tag sample we have 3 possible pairs of jets that could come from the W decay. We use the pair that gives the invariant mass closest to the world average W mass. This variable was found to have the strongest sensitivity to Δ_{JES} shifts among several possibilities. The dijet mass distribution in 1-tag and 2-tag sample is shown in Fig. 5.4. The $t\bar{t}$ Monte Carlo sample generated with $\mathrm{M}_{\mathrm{top}} = 170$ GeV/c^2 was used with three Δ_{JES} values. The distribution in 1-tag sample is much narrower than the 2-tag distribution (about 10 GeV/c^2 in 1-tag sample and 15 GeV/c^2 in 2-tag sample). The change is twice as large as a function of Δ_{JES} in the 2-tag sample as in the 1-tag sample. This difference in behavior between the two samples is caused by the incorrect combinations being chosen forcing events into the peak in the 1-tag sample.

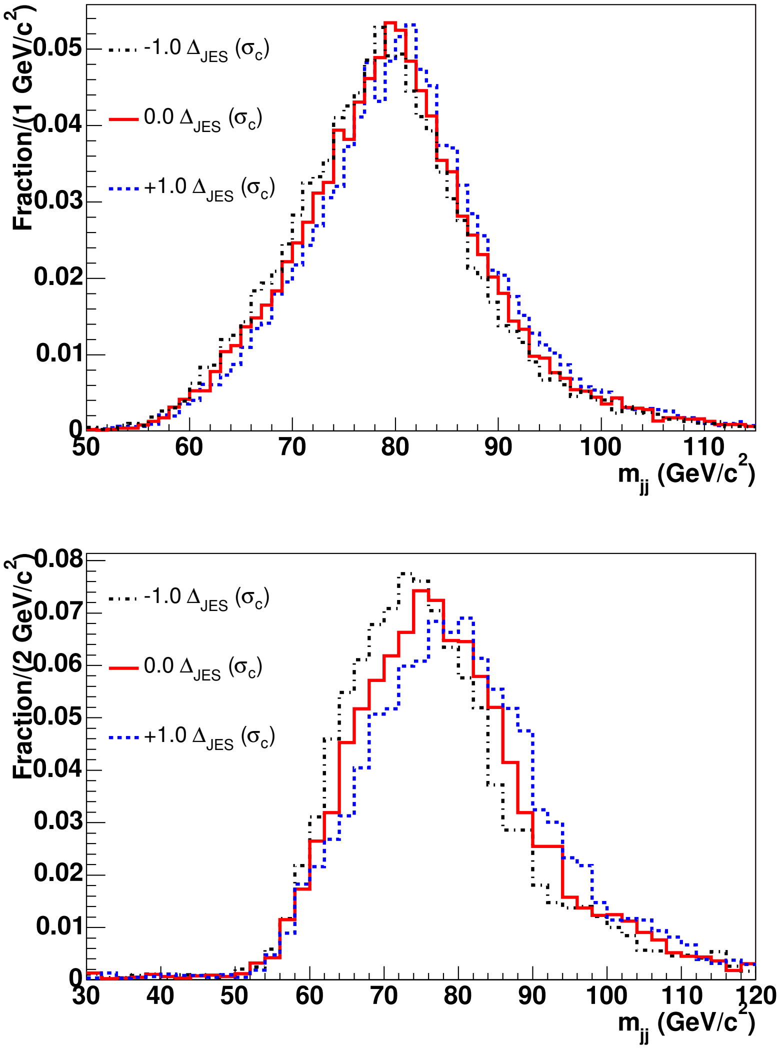


Figure 5.4: Dijet invariant mass for three values of Δ_{JES} for Monte Carlo events passing the lepton + jets 1-tag (top) and 2-tag selection (bottom).

CHAPTER 6

THE TOP QUARK MASS FIT

To extract the mass of the top quark $\mathrm{M_{top}}$ from the data we employ a likelihood fit method. We also include the jet energy scale offset Δ_{JES} and the signal and background cross sections (σ_s, σ_b) as parameters of the likelihood. By doing so we turn the systematic uncertainty on the top quark mass associated with these additional parameters into a statistical uncertainty. The overall uncertainty will be reduced since information about these parameters is present in the data. Instead of considering the values of cross sections we convert them to the expected number of signal and background events n_s and n_b which are more straightforward to treat. We first derive a general form for the likelihood which compares the distributions of observables (*cf.* Chapter 5) found in data to the probability density functions for those observables. Next we discuss how these probability density functions are extracted from Monte Carlo samples using Kernel Density Estimation and Local Polynomial Smoothing.

6.1 The Likelihood

We want to obtain the probability for values of the parameters given the observation of data $p(\boldsymbol{\alpha}, n_s, n_b|D)$ where D is a set of N vectors of observables $\boldsymbol{v}_i$ from each event in the data. We grouped the $(\mathrm{M_{top}}, \Delta_{\mathrm{JES}})$ pair into a vector $\boldsymbol{\alpha}$. We start by evaluating the probability of observing the data D given values of the parameters:

$p(D|\boldsymbol{\alpha}, n_s, n_b)$ i.e. the likelihood function. The discussion below follows closely the treatment in [49].

In the dataset D we have some unknown number of signal N_s and background N_b events. We can successively decompose the probability over all possible values of those two random variables:

$$\begin{aligned} p(D|\boldsymbol{\alpha}, n_s, n_b) &= \sum_{N_s=0}^{N} \Big(p(D|N_s, \boldsymbol{\alpha}, n_s, n_b) p(N_s|\boldsymbol{\alpha}, n_s, n_b) \Big) && (6.1.1) \\ &= \sum_{N_s=0}^{N} \sum_{N_b=0}^{N} \Big(p(D|N_s, N_b, \boldsymbol{\alpha}, n_s, n_b) \\ &\times\ p(N_s|\boldsymbol{\alpha}, n_s, n_b) p(N_b|\boldsymbol{\alpha}, n_s, n_b) \Big) && (6.1.2) \end{aligned}$$

Once we know the numbers of signal and background events in a dataset the probability of observing this dataset no longer depends on values of n_s and n_b so

$$p(D|N_s, N_b, \boldsymbol{\alpha}, n_s, n_b) = p(D|N_s, N_b, \boldsymbol{\alpha}). \qquad (6.1.3)$$

Moreover the sum of numbers of signal and background events must be equal to N, this allows us to collapse the second sum:

$$\begin{aligned} p(D|\boldsymbol{\alpha}, n_s, n_b) &= \sum_{N_s=0}^{N} \Big(p(D|N_s, N_b = N - N_s, \boldsymbol{\alpha}) \\ &\times\ p(N_s|\boldsymbol{\alpha}, n_s, n_b) p(N_b|\boldsymbol{\alpha}, n_s, n_b) \Big) && (6.1.4) \end{aligned}$$

N_s and N_b are independent, Poisson-distributed random variables with means n_s and

n_b respectively, therefore the last two terms in the product can be expressed:

$$p(N_s|\boldsymbol{\alpha}, n_s, n_b)p(N_b|\boldsymbol{\alpha}, n_s, n_b) = \frac{e^{-n_s}{n_s}^{N_s}}{N_s!}\frac{e^{-n_b}{n_b}^{N_b}}{N_b!} \tag{6.1.5}$$

In the dataset of N events there are 2^N combinations for assigning each event to signal or background. We will label each combination by a number t^N varying from 0 to $2^N - 1$. It is convenient to interpret the combination label t^N in binary notation. We will adopt a convention that if the i^{th} digit of $t^N_{(2)}$ (or the i^{th} 'bit' of t^N) is 1, then the i^{th} event comes from signal in combination t^N. Let us also define $t^N_{N_s}$ as a label for a combination with exactly N_s signal events. $t^N_{N_s}$ varies from 0 to $\binom{N}{N_s}$ and $p(t^N_{N_s})$ is a probability of obtaining the combination labeled $t^N_{N_s}$.

With these definitions we tackle the first term in Equation 6.1.4. We expand the term $p(D|N_s, N_b = N - N_s, \boldsymbol{\alpha})$ by summing over all combinations with N_s signal events:

$$\begin{aligned} p(D|N_s, N_b = N_s - N, \boldsymbol{\alpha}) &= \sum_{t^N_{N_s}} \Big(p(D|t^N_{N_s}, N_s, N_b = N - N_s, \boldsymbol{\alpha}) \\ &\times \; p(t^N_{N_s}|N_s, N_b = N - N_s, \boldsymbol{\alpha}) \Big) \\ &= \sum_{t^N_{N_s}} p(D|t^N_{N_s}, \boldsymbol{\alpha}) p(t^N_{N_s}|N_s, N_b = N - N_s, \boldsymbol{\alpha}) \\ &= \sum_{t^N_{N_s}} p(D|t^N_{N_s}, \boldsymbol{\alpha}) \frac{1}{\binom{N}{N_s}} \end{aligned} \tag{6.1.6}$$

In the last step we used the fact that probabilities of observing all combinations with N_s signal events are equal so $p(t^N_{N_s}|N_s, N_b = N - N_s, \boldsymbol{\alpha}) = 1/\binom{N}{N_s}$. We insert

Equations 6.1.6 and 6.1.5 into Equation 6.1.4 to obtain:

$$\begin{aligned} p(D|n_s,n_b,\boldsymbol{\alpha}) &= \sum_{N_s=0}^{N}\sum_{t^N_{N_s}}\left(p(D|t^N_{N_s},\boldsymbol{\alpha})\frac{1}{\binom{N}{N_s}}\frac{e^{-n_s}{n_s}^{N_s}}{N_s!}\frac{e^{-n_b}{n_b}^{N_b}}{N_b!}\right) && (6.1.7)\\ &= \frac{e^{-(n_s+n_b)}}{N!}\sum_{N_s=0}^{N}\sum_{t^N_{N_s}}\left({n_s}^{N_s}{n_b}^{N_b}p(D|t^N_{N_s},\boldsymbol{\alpha})\right) && (6.1.8)\end{aligned}$$

The nested sum in the last equation is over all values of N_s and all combinations with N_s signal events therefore it is equivalent to a single sum over all combinations (with order of the terms in the expansion rearranged):

$$p(D|n_s,n_b,\boldsymbol{\alpha}) = \frac{e^{-(n_s+n_b)}}{N!}\sum_{t^N=0}^{2^N-1}\left({n_s}^{N_s(t^N)}{n_b}^{N_b(t^N)}p(D|t^N,\boldsymbol{\alpha})\right) \qquad (6.1.9)$$

Where $N_s(t^N)$ and $N_b(t^N)$ are numbers of respectively signal and background events in combination t^N. Expanding the last term using the probability density functions for the observables $\boldsymbol{v}$ we obtain:

$$\begin{aligned} p(D|n_s,n_b,\boldsymbol{\alpha}) &= \frac{e^{-(n_s+n_b)}}{N!}\\ &\times \sum_{t^N=0}^{2^N-1}\left({n_s}^{N_s(t^N)}{n_b}^{N_b(t^N)}\prod_{i\in S(t^N)}P_s(\boldsymbol{v_i},\boldsymbol{\alpha})\prod_{i\in B(t^N)}P_b(\boldsymbol{v_i},\boldsymbol{\alpha})\right)\\ &= \frac{e^{-(n_s+n_b)}}{N!}\\ &\times \sum_{t^N=0}^{2^N-1}\left(\prod_{i\in S(t^N)}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^N)}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\right) && (6.1.10)\end{aligned}$$

Where $S(t^N)$ $\left(B(t^N)\right)$ is a set of indices of signal $\left(\text{background}\right)$ events in the combination t^N. For example if $t^4_{(2)} = 1010$ $S(t^4) = \{1,3\}$ and $B(t^4) = \{0,2\}$. We assumed that the probability density function P_b is a constant function of $\mathrm{M}_{\mathrm{top}}$, but may be a non-trivial function of other parameters in $\boldsymbol{\alpha}$, hence we include $\boldsymbol{\alpha}$ in it's argument list.

Now we show by induction that the sum in Equation (6.1.10) can be written as a single product:

$$\sum_{t^N=0}^{2^N-1}\Big(\prod_{i\in S(t^N)}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^N)}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big) = \prod_{i=1}^{N}\Big(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha})+n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha})\Big) \tag{6.1.11}$$

Base Case: $N = 1$

For $N = 1$ the sum on the left hand side of Equation (6.1.11) runs over $t^N = \{0, 1\}$. The first term of the sum is simply $n_bP_b(\boldsymbol{v_1},\boldsymbol{\alpha})$ since $t^1 = 0$ denotes that the first and only event in the dataset is background, so that the first product has no terms and the second product has just one term. Similarly the second term is $n_sP_s(\boldsymbol{v_1},\boldsymbol{\alpha})$, so Equation (6.1.11) is satisfied for $N = 1$

Induction Step: $N > 1$

We assume that Equation (6.1.11) holds for dataset size $N - 1$. Now divide the sum on LHS into two sums:

$$
\begin{aligned}
&\sum_{t^N=0}^{2^N-1}\Big(\prod_{i\in S(t^N)}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^N)}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big)\\
=&\sum_{t^N=0}^{2^{N-1}-1}\Big(\prod_{i\in S(t^N)}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^N)}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big) \qquad (6.1.12)\\
+&\sum_{t^N=2^{N-1}}^{2^N-1}\Big(\prod_{i\in S(t^N)}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^N)}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big) \qquad (6.1.13)
\end{aligned}
$$

In the first sum on RHS the most significant digit in the base 2 representation of t^N is always 0 indicating that the N^{th} event is a background event therefore the first product in that sum reduces to a product over all combinations t^{N-1}. In the second product on line 6.1.12 the term $n_bP_b(\boldsymbol{v_N},\boldsymbol{\alpha})$ is present for all terms in the sum on that line. We explicitly pull out this term from the product, while readjusting the set of events that the product runs over to $i \in B(t^{N-1})$ since the sequence $(t^N : 1 \to 2^{N-1}-1)$ is the same as sequence $(t^{N-1} : 1 \to 2^{N-1}-1)$. Thus the line 6.1.12 becomes:

$$
\sum_{t^{N-1}=0}^{2^{N-1}-1}\Big(n_bP_b(\boldsymbol{v_N},\boldsymbol{\alpha})\prod_{i\in S(t^{N-1})}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^{N-1})}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big) \qquad (6.1.14)
$$

We treat the sum on line 6.1.13 similarly. We use the fact that in this sum the N^{th} bit of t^N is always set therefore we change limits of the summation: $\sum_{t^N=2^{N-1}}^{2^N-1} \to \sum_{t^{N-1}=0}^{2^{N-1}-1}$ while explicitly pulling out the term $n_sP_s(\boldsymbol{v_N},\boldsymbol{\alpha})$ from the first product.

Therefore the line 6.1.13 can be written as:

$$\sum_{t^{N-1}=0}^{2^{N-1}-1}\Big(n_sP_s(\boldsymbol{v_N},\boldsymbol{\alpha})\prod_{i\in S(t^{N-1})}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^{N-1})}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big) \tag{6.1.15}$$

Therefore we have:

$$\begin{aligned}\sum_{t^N=0}^{2^N-1}\Big(\prod_{i\in S(t^N)}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^N)}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big)\\ &= \Big(n_sP_s(\boldsymbol{v_N},\boldsymbol{\alpha})+n_bP_b(\boldsymbol{v_N},\boldsymbol{\alpha})\Big)\\ &\times\sum_{t^{N-1}=0}^{2^{N-1}-1}\Big(\prod_{i\in S(t^{N-1})}(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha}))\prod_{i\in B(t^{N-1})}(n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha}))\Big)\\ &= \prod_{i=1}^{N}\Big(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha})+n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha})\Big)\end{aligned} \tag{6.1.16}$$

In the last step we used the induction hypothesis. This completes the proof of Equation (6.1.11).

Using Equation (6.1.11) in Equation (6.1.10) we find the probability of observing the dataset D given the values of parameters n_s, n_b, $\boldsymbol{\alpha}$:

$$\begin{aligned}p(D|n_s,n_b,\boldsymbol{\alpha}) &= \frac{e^{-(n_s+n_b)}}{N!}\prod_{i=1}^{N}\Big(n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha})+n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha})\Big)\\ &= \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!}\\ &\prod_{i=1}^{N}\Big(\frac{n_sP_s(\boldsymbol{v_i},\boldsymbol{\alpha})+n_bP_b(\boldsymbol{v_i},\boldsymbol{\alpha})}{n_s+n_b}\Big)\end{aligned} \tag{6.1.17}$$

Since in each channel we use a different set of observables it is beneficial to separate the dataset into exclusive subsamples belonging to each channel. Furthermore in a particular channel categories of events based on additional criteria, such as number of b tagged jets, have different purities and different distributions of a given observable. Thus in general the dataset can be logically divided into a number of subsamples M. We write the probability of observing the dataset D given $\boldsymbol{\alpha}$ and expected number of signal and background events n_{sq}, n_{bq} in subsample q:

$$p(D|(n_{s1}, n_{b1}), \ldots, (n_{sM}, n_{bM})) = \prod_{q=1}^{M} \Big(p(D_q|n_{sq}, n_{bq}, \boldsymbol{\alpha}) \Big) \tag{6.1.18}$$

In the equation above each term in the product is given by Equation (6.1.17). We have assumed here that the signal and background expectation values n_{sq}, n_{bq} are independent for different subsamples. This is equivalent to stating that we do not assume knowledge of how the cross section and acceptance for one subsample influence these of another subsample.

We now have a form for the likelihood function of observing the dataset D given values of parameters $\boldsymbol{\alpha}$ and n_{sq}, n_{bq}. The purpose of this analysis is to update the knowledge of the parameters $\boldsymbol{\alpha}$, n_{sq}, n_{bq} based on the observation D, therefore we need posterior probability $p((n_{s1}, n_{b1}), \ldots, (n_{sM}, n_{bM}), \boldsymbol{\alpha}|D)$. We use Bayes' Law to obtain:

$$\begin{gathered} p((n_{s1}, n_{b1}), \ldots, (n_{sM}, n_{bM}), \boldsymbol{\alpha}|D) = \\ \frac{p((n_{s1},n_{b1}),\ldots,(n_{sM},n_{bM}),\boldsymbol{\alpha})p(D|(n_{s1},n_{b1}),\ldots,(n_{sM},n_{bM}),\boldsymbol{\alpha})}{p(D)} \end{gathered} \tag{6.1.19}$$

We can drop the constant term $p(D)$ from the denominator. The term $p((n_{s1}, n_{b1}), \ldots, (n_{sM}, n_{bM}), \boldsymbol{\alpha})$ is the prior probability which can be factorized on assumption that parameters are independent:

$$\begin{aligned} &p((n_{s1}, n_{b1}), \ldots, (n_{sM}, n_{bM}), \boldsymbol{\alpha}) \\ &= p(\mathrm{M_{top}}) p(\Delta_{\mathrm{JES}}) \prod_{q=1}^{M} \left(p(n_{sq}) p(n_{bq}) \right) \end{aligned} \tag{6.1.20}$$

We express our ignorance about the number of expected signal events and $\mathrm{M_{top}}$ by applying uniform priors for these parameters (and dropping the normalization constants). We have defined the jet energy scale shift Δ_{JES} in units of the uncertainty on JES so the prior for Δ_{JES} is just a Gaussian with zero mean and unit width:

$$p(\Delta_{\mathrm{JES}}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\Delta_{\mathrm{JES}}^2}{2}} \tag{6.1.21}$$

The prior probabilities for the background expectation values n_{bq} are Gaussians with means of the *a priori* estimates n_{b0q} obtained as described in Sections 4.2 and 4.3 and errors on these estimates are $\sigma_{n_{b0q}}$.

$$p(n_{bq}) = \frac{1}{\sigma_{n_{b0q}} \sqrt{2\pi}} e^{-\frac{(n_{bq} - n_{b0q})^2}{2\sigma_{n_{b0q}}}} \tag{6.1.22}$$

Finally we can write the form of the posterior probability that will be maximized in the fit:

$$p((n_{s1}, n_{b1}), \ldots, (n_{sM}, n_{bM}), \boldsymbol{\alpha} | D) = p(\Delta_{\mathrm{JES}}) \prod_{q=1}^{M} \left(p(n_{bq}) p(D_q | n_{sq}, n_{bq}, \boldsymbol{\alpha}) \right) \tag{6.1.23}$$

It is more practical to take a negative logarithm of Equation (6.1.23) since the product on the right side of the equation is converted to a sum and the product over all events in the term $p(D_q | n_{sq}, n_{bq}, \boldsymbol{\alpha})$ is also converted to a sum. Minimizing the negative log likelihood is more practical from a computational standpoint since the range of values of the probability function is diminished leading to less numerical artifacts. The values of the parameters are then found by numerically searching for the minimum of $-\log(p((n_{s1}, n_{b1}), \ldots, (n_{sM}, n_{bM}), \boldsymbol{\alpha}|D))$. The uncertainty on any given parameter is obtained by searching for the values of the parameter in question for which log likelihood deviates by 0.5 from the minimum while minimizing the log likelihood with respect to all other parameters. This error estimate will provide a 68.3% coverage in the case where likelihood form is exactly Gaussian. We will correct for any deviation in the coverage as described in Section 7.1. The uncertainty on $\mathrm{M_{top}}$ obtained in this fashion includes the statistical error as well as the uncertainty due to the jet energy scale calibration.

Performing top quark mass fits in separate channels or single subsamples is accomplished by including only the appropriate terms in the product in Equation (6.1.23). These fits will be performed as a cross check to the main result and to check for consistency between the channels and samples. Note that in the the dilepton channel we do not have enough information to measure the top quark mass and Δ_{JES} simultaneously. Such two-dimensional fit is possible to perform because the information about the jet energy scale offset will be supplied by the prior $p(\Delta_{\mathrm{JES}})$ with no additional information provided by the distributions of observables. As a part of studies of systematic uncertainties we break the assumption that the processes leading to jet energy scale uncertainty can be modeled by just one parameter (Δ_{JES}) as described in Section 8.1. The adopted treatment of the systematic effect due to this assumption would be impractical in a channel with no intrinsic power to resolve Δ_{JES}, therefore in a dilepton-only fit we fix the Δ_{JES} parameter to it's nominal value and fit for $\mathrm{M_{top}}$ only.

Additional cross checks can be performed by removing the priors $p(\Delta_{\text{JES}})$ and $p(n_{bq})$, which is equivalent to claiming ignorance about the value of the jet energy calibration and expected numbers of background events.

6.1.1 Kernel Density Estimation

To obtain the probability density functions (PDFs) P_s and P_b we employ a two step process. First we estimate the probability density functions for the signal samples and background samples using the Kernel Density Estimation (KDE) approach described in this section. KDE gives us the probability density functions at discrete values of the parameters M_{top}, Δ_{JES} (or in case of background just Δ_{JES}) where the MC samples were generated. We obtain the PDFs as continuous functions of the parameters using the Local Polynomial Smoothing (LPS) [50] described in the next section. Reference [18] gives a through overview of the KDE technique.

Given a signal model generated with some value of parameters M_{top}, Δ_{JES} or a background model we are faced with a problem of constructing a PDF for the observables $\boldsymbol{v}_i$ in that sample. Several alternatives exist. One is to use histograms. A product of two one-dimensional histograms may suffice if the correlations between the observables are very small. This is not the case here especially for the Dilepton sample where the correlation between H_T and m_t^{NWA} are approximately 40% in the signal samples and as much as 60% in the background samples. We can also use a two-dimensional histogram. This is also not optimal as the bins of the histograms will have to be chosen separately for different signal and background samples with very different numbers of events. We can choose to parametrize the distributions of observables. This approach was used in the previous lepton + jets and dilepton template top quark mass measurements [24, 45] where the PDFs were parametric in the values of observables and values of parameters $\mathrm{M_{top}}$, Δ_{JES} (or just $\mathrm{M_{top}}$ in the case of the dilepton measurement). An argument from first principles that would predict a functional form of the PDF including the instrumental effects is hard to make therefore an *ad-hoc* form was chosen in those previous measurements. Moreover it is formidable to account for correlations between observables since a function of two variables needs to be parametrized.

One approach that avoids the problems mentioned above is the Kernel Density Estimation. The technique gives an estimate $\hat{f}$ for the value of the PDF for a given sample as a function of the observables $\boldsymbol{v}_i$. The number of events in the sample is taken into account and correlation between observables are treated intrinsically. We limit the discussion here to PDFs which are two dimensional in observables, however the same approach can be applied to PDFs of any dimension.

We start by computing a value of a "pilot" density estimates for all events of the model samples *i.e.* all Monte Carlo signal samples and all background model samples. For the event j in a given sample S_{model} for a particular model ($t\bar{t}$, QCD etc.) the value of pilot density estimate is given by Equation (6.1.24).

$$\hat{f}_{S_{model},pilot}(\boldsymbol{v_j}) = \frac{1}{nh_1h_2\sum w_i}\sum_{i=1}^{n}\left(w_iK\left(\frac{\boldsymbol{v}_{\mathbf{j},1}-\boldsymbol{v}_{\mathbf{i},1}}{h_1}\right)\times K\left(\frac{\boldsymbol{v}_{\mathbf{j},2}-\boldsymbol{v}_{\mathbf{i},2}}{h_2}\right)\right). \tag{6.1.24}$$

In the equation above n is the number of events in the sample. Since in general events in the sample have unequal weights (due to *e.g.* different mistag probabilities or fake lepton probabilities) we apply the weighting factor w_i to each event. Function K is given by Equation (6.1.25).

$$K(t) = \frac{3}{4}(1-t^2) \text{ if } |t| < 1 \text{ and } K(t) = 0 \text{ otherwise} \tag{6.1.25}$$

The product of the two functions K in the sum in Equation (6.1.24) is a two-dimensional "kernel". The sample or a set of observables $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}\}$ can be itself considered as an estimate for a probability density function consisting of a sum of two-dimensional delta functions. This is of course not a very useful PDF unless our sample has infinite statistics. Computing the sum of finite width kernels gives us a well behaved PDF even for samples with relatively few events. The particular choice of the kernel function is motivated in [51] as giving the best overall estimate for the PDF given a fixed number of data points.

This smoothing in first and second dimension is controlled by the bandwidth parameters h_1 and h_2 with larger bandwidths generating a wider kernel. If the underlying PDF is known it is possible to find the optimal value of the smoothing parameter [51]. However since we naturally do not know the true underlying distribution we use the "oversmoothed bandwidth" - the upper bound on the optimal bandwidth. This can be calculated knowing only the variance and event count of the sample. For our particular choice of kernel oversmoothed h is [18]:

$$h_x = \left(\frac{9^{d-1} \left(d+8\right)^{\left(\frac{d+6}{2}\right)} \pi^{\frac{d}{2}}}{16 \cdot 15^{d-2} n \Gamma\left(\frac{d+8}{2}\right) \left(d+2\right)} \right)^{\frac{1}{d+4}} \sigma_x, \tag{6.1.26}$$

where d is dimensionality of the observable space (here 2) and we use root mean square width for σ_x.

The observables considered have natural cut-offs, for example in the lepton + jets channel a boundary occurs at m_t^{reco} below the W boson mass. Similarly in the dilepton channel H_T can not be lower than 200 GeV. An event near such a boundary will contribute to probability outside of physical region and cause the PDF inside the physical region to be not normalized. We impose a boundary cut rejecting events outside. If an event is near the edge of the allowed region it's kernel will be truncated at the boundary and scaled to maintain proper normalization. The allowed region is set at 100 GeV/c^2 $< m_t^{\mathrm{NWA}} <$ 350 GeV/c^2 and 200 GeV $< H_T <$ 800 for the dilepton channel. In the lepton + jets samples we allow 110 GeV/c^2 $< m_t^{\mathrm{reco}} <$ 350 GeV/c^2 and 50 GeV/c^2 $< m_{jj} <$ 110 GeV/c^2 in the 1-tag sample and 50 GeV/c^2 $< m_{jj} <$ 115 GeV/c^2 in the 2-tag sample. The efficiency of passing the boundary cut is approximately 98% for dilepton and lepton + jets signal samples and is indepentent of $\mathrm{M_{top}}$ and Δ_{JES}.

Once the pilot density estimate has been computed for all signal and background samples we are ready to estimate the probability density in sample S_{model} for an event k from a dataset D as follows:

$$\hat{f}_{S_{model}}(\boldsymbol{v_k}) = \frac{1}{n\sum w_i} \sum_{i=1}^{n_{S_{model}}} \left(\frac{w_i}{h_{1_i} h_{2_i}} K\left(\frac{\boldsymbol{v}_{\mathbf{k},1} - \boldsymbol{v}_{\mathbf{i},1}}{h_{1_i}} \right) \times K\left(\frac{\boldsymbol{v}_{\mathbf{k},2} - \boldsymbol{v}_{\mathbf{i},2}}{h_{2_i}} \right) \right). \tag{6.1.27}$$

Both sums in the equation above run over the $n_{S_{model}}$ events in the model sample, not in the dataset D. Note that the bandwidth of the kernel is now separate for each event i in the model sample. The parameter h_{x_i} designates the width of the kernel in dimension x (x is 1 or 2) in the event i. This bandwidth is calculated using the pilot density estimate for that event:

$$h_{x_i} = h_x \sqrt{\frac{G}{\hat{f}_{S_{model},pilot,clipped}(\boldsymbol{v_i})}}. \tag{6.1.28}$$

The "clipped" pilot estimate for the i^{th} event $\hat{f}_{S_{model},pilot,clipped}(\boldsymbol{v_i})$ is the greater of the two quantities: 0.1 of the maximal pilot density estimate in the model sample

$\hat{f}_{S_{model},pilot,max}$ and the pilot density estimate for that event:

$$\hat{f}_{S_{model},pilot,clipped}(\boldsymbol{v_i}) = max(0.1 \times \hat{f}_{S_{model},pilot,max}, \hat{f}_{S_{model},pilot}(\boldsymbol{v_i})). \quad (6.1.29)$$

The quantity G is a weighted geometric mean of the pilot estimates for all the events in the model sample:

$$G = \left(\prod_{i=1}^{n_{S_{model}}} \left(\hat{f}_{S_{model},pilot}(\boldsymbol{v_i}) \right)^{w_i} \right)^{(1/\sum w_i)} . \quad (6.1.30)$$

The estimate $\hat{f}_{S_{model}}(\boldsymbol{v_k})$ given by Equation (6.1.27) is then an “adaptive” estimate. This means that the applied smoothing varies for different regions of the PDF, with smaller smoothing applied near the peak than in the tails. This allows us to extract information about the location of the peak of the PDF and at the same time not be exposed to effects of small statistics in tails.

The smoothing is proportional to the inverse square root of the PDF in a given region, and since the actual PDF is naturally not known the pilot density estimate is taken as a proxy. The choice of the inverse square root dependence on the PDF is motivated in [52] as minimizing the bias of the estimate where the expected squared difference of the true PDF and the estimate has been taken as a figure of merit.

The smoothing parameter is varied in relation to the constant h_x, therefore we need to make sure that scaling of the smoothing constant is of order 1. This is achieved by using a geometric mean in the ratio in Equation (6.1.28).

Application of the adaptive procedure may lead to undesirable effects where the value of PDF at some point $\boldsymbol{v}$ depends on events in the tails of the distribution more than it depends on the events near to point $\boldsymbol{v}$. This occurs when the value of pilot estimate changes rapidly. As an example in one dimension consider two events with observable values v_1 and v_2 with $|v - v_1| < |v - v_2|$. If $|v - v_1|/|v - v_2| > \sqrt{\hat{f}_{pilot}(v_2)/\hat{f}_{pilot}(v_1)}$ the second event contributes more probability at v than the first, even though it lies further away from v. To control such effects we impose a limit on how large the smoothing parameter can get. Using the clipped pilot density estimate in calculation of h_{x_i} (Equation (6.1.28)) ensures that the maximum bandwidth for a sample is at most factor of $\sqrt{10}$ larger than the minimum bandwidth. The particular choice of factor $\sqrt{10}$ is suggested in [52] and was not optimized for. For a typical signal Monte Carlo sample most events obtain smoothing bandwidth of $10 - 20$ GeV/c^2 in m_t^{NWA} and $30 - 45$ GeV in H_T in the dilepton sample and $6 - 10$ GeV/c^2 for the m_t^{reco} and $2 - 4$ GeV/c^2 for m_{jj} in the lepton + jets samples.

The density estimates for a signal sample generated at $\mathrm{M_{top}} = 170$ GeV/c^2 and $\Delta_{\mathrm{JES}} = 0\sigma_c$ in the dilepton and lepton + jets channels are shown in Figures 6.1 and 6.2. While no substantial correlation between observables is present in the lepton + jets samples the observables are correlated in the dilepton samples. The correlation coefficient between observables in the dilepton signal samples approximately 0.4 and approximately 0.6 in the background dilepton samples[1]. As mentioned above no special treatment is needed to account for this. The background KDE estimates are shown in Figures 6.3 and 6.4. The model for the fake lepton background in the dilepton channel has very low statistics particularly in the 1-tag subsample. This is why an irregular shape appears in the bottom panel of Fig. 6.3.

---

[1] The correlation coefficient for two random variables x and y is defined as $\rho_{xy} = \frac{Cov(x,y)}{\sigma_x \sigma_y}$, where σ_x and σ_y are the standard deviations of the random variables

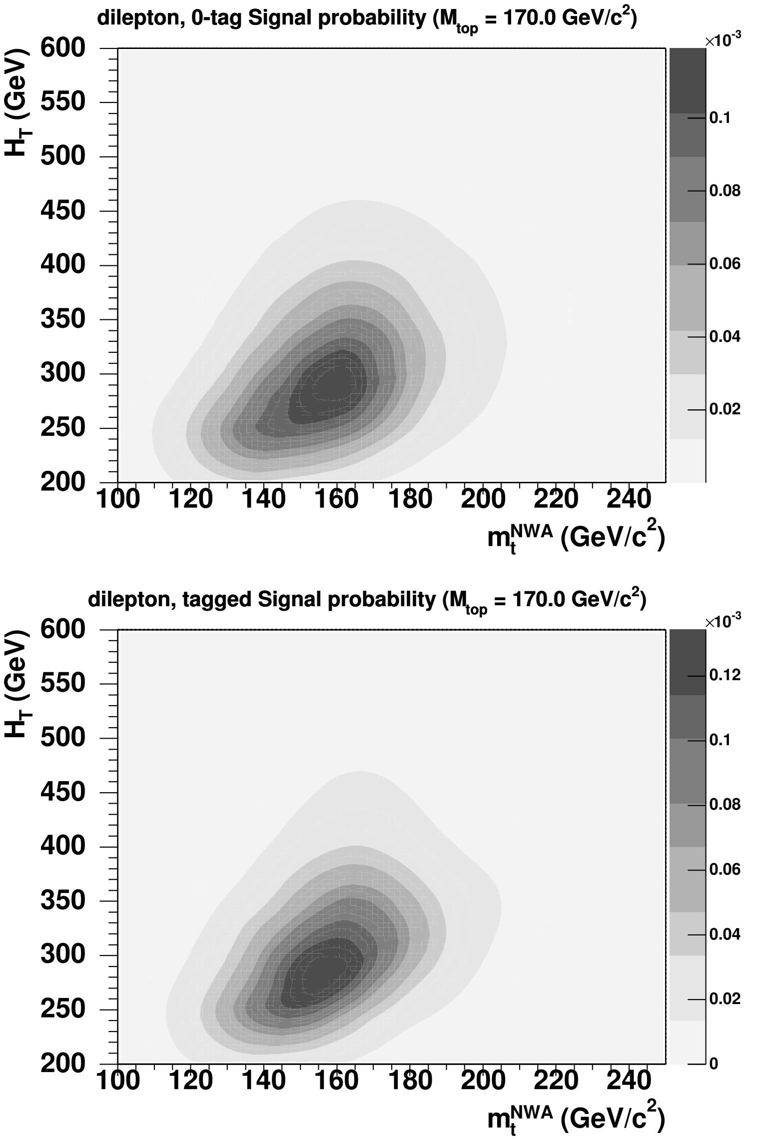


Figure 6.1: Kernel density estimate for a signal sample $\mathrm{M_{top}} = 170$ GeV/c^2 and $\Delta_{\mathrm{JES}} = 0\sigma_c$ in the dilepton samples: 0-tag (top) and 1-tag (bottom).

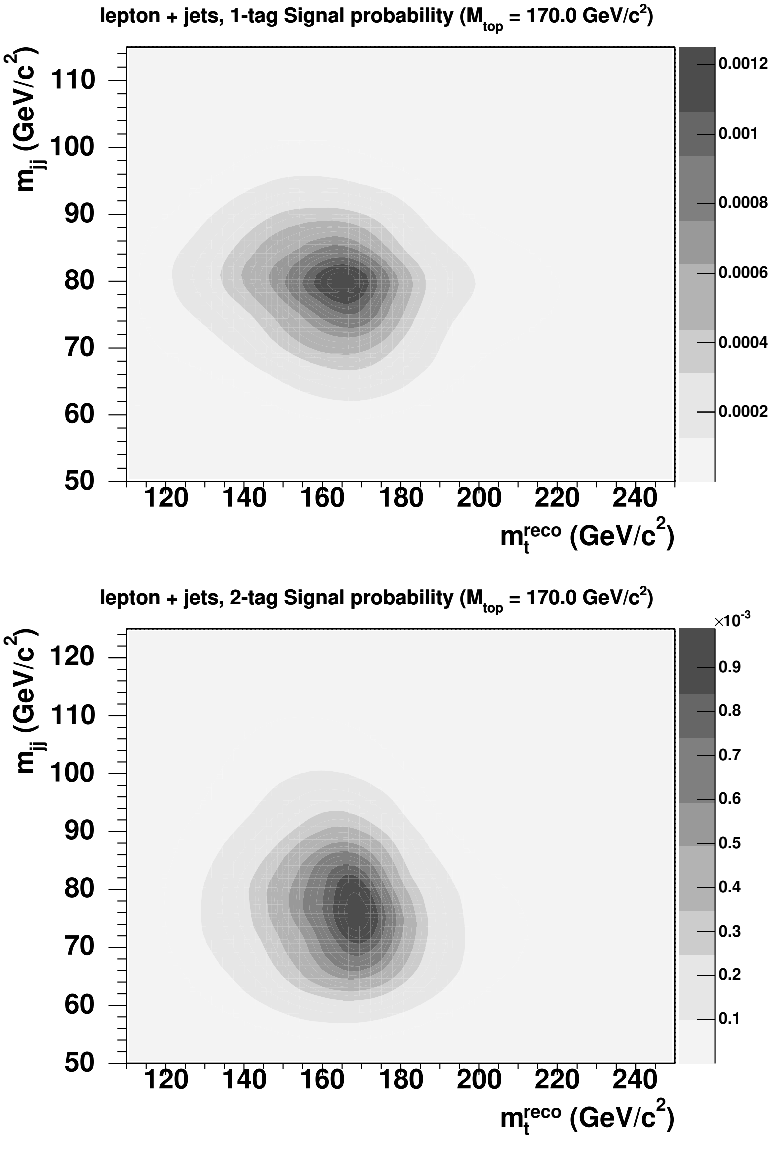


Figure 6.2: Kernel density estimate for a signal sample M_{top} = 170 GeV/c^2 and $\Delta_{JES} = 0\sigma_c$ in the lepton + jets samples: 1-tag (top) and 2-tag (bottom).

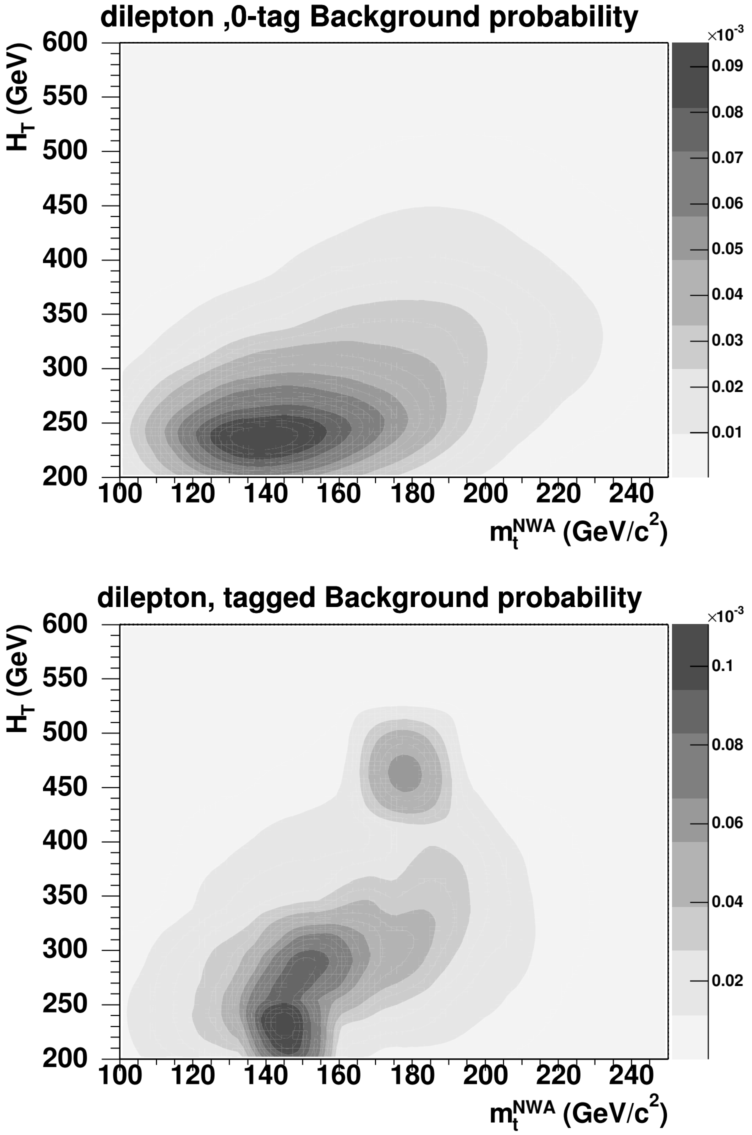
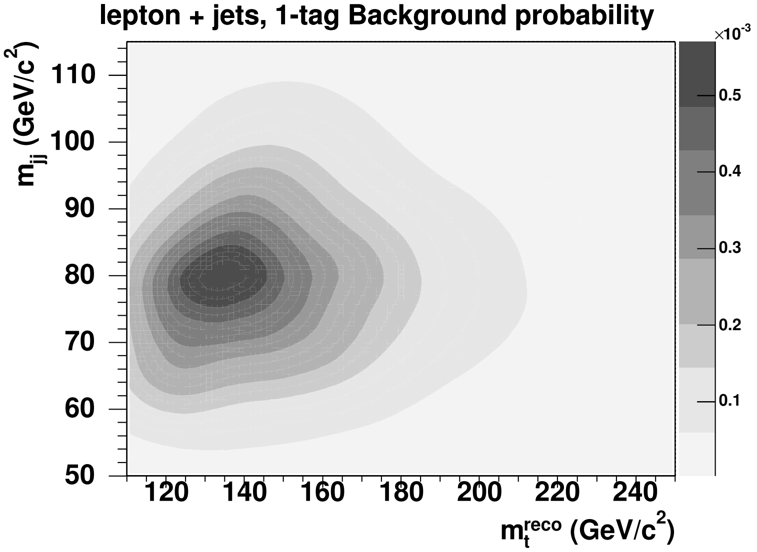


Figure 6.3: Kernel density estimate for a background sample at and $\Delta_{\mathrm{JES}} = 0\sigma_c$ in the dilepton samples: 0-tag (top) and 1-tag (bottom).



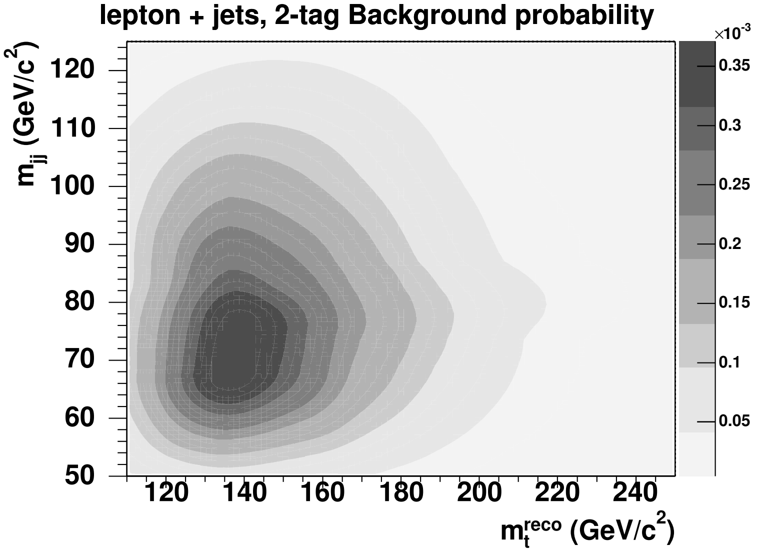


Figure 6.4: Kernel density estimate for a background sample at and $\Delta_{\mathrm{JES}} = 0\sigma_c$ in the lepton + jets samples: 1-tag (top) and 2-tag (bottom).

6.1.2 Local Polynomial Smoothing

As mentioned above the KDE technique gives us the estimate of PDFs at discrete values of parameters $\boldsymbol{\alpha}$, where in the case of signal PDFs $\boldsymbol{\alpha}$ is a two-dimensional vector $\boldsymbol{\alpha} = (\mathrm{M_{top}}, \Delta_{\mathrm{JES}})$. In the case of background PDFs $\boldsymbol{\alpha}$ reduces to the scalar Δ_{JES}. Using Local Polynomial Smoothing (LPS) [50] we will obtain an estimate $\hat{P}(\boldsymbol{\alpha})$ for the true value of the function $P(\boldsymbol{\alpha})$ where we have omitted the labels of the type of the probability density functions (signal or background) as well as arguments designating the observable values $\boldsymbol{v}$. From the kernel density estimation we obtain a vector of N_e estimates $\boldsymbol{Y}_k$ for the values of $P(\boldsymbol{x_k})$ at N_e points $\boldsymbol{\alpha_k}$. We assume that the uncertainty on the estimate is common for all the points so that:

$$Y_k = P(\boldsymbol{\alpha_k}) + \epsilon_k \tag{6.1.31}$$

where the ϵ_k are identically distributed random variables with finite variance and vanishing mean. A second order expansion of the function P in the neighborhood of $\boldsymbol{\alpha}$ can be written as:

$$P(\boldsymbol{t}) = \langle \boldsymbol{a}, \boldsymbol{A}(\boldsymbol{t} - \boldsymbol{\alpha}) \rangle \tag{6.1.32}$$

In the equation above the angle bracket denotes inner product. The quantity $\boldsymbol{A}$ is a vector of basis functions for space of second order polynomials. We will designate the dimensionality of this function space by N_f. For a two-dimensional $\boldsymbol{u}$ $\boldsymbol{A}(\boldsymbol{u})$ is:

$$\boldsymbol{A}(\boldsymbol{u}) = \begin{pmatrix} 1 \\ \boldsymbol{u}_0 \\ \boldsymbol{u}_1 \\ \boldsymbol{u}_0\boldsymbol{u}_1 \\ \frac{1}{2}\boldsymbol{u}_0^2 \\ \frac{1}{2}\boldsymbol{u}_1^2 \end{pmatrix} \tag{6.1.33}$$

If $\boldsymbol{u}$ is a scalar $\boldsymbol{A}(\boldsymbol{u})$ naturally reduces to a 3-component vector. Coefficients of the expansion 6.1.32 are given by the components of the vector $\boldsymbol{a}$.

To evaluate value of $\hat{P}$ at point $\boldsymbol{\alpha}$ we find a vector of coefficients $\hat{\boldsymbol{a}}$ for which the criterion given by Equation 6.1.34 is minimized, that is we find a second order expansion around $\boldsymbol{\alpha}$ of function P which matches best the estimates Y_k at points $\boldsymbol{\alpha}_{\boldsymbol{k}}$.

$$C = \sum_k w_k(\boldsymbol{\alpha})(Y_k - \langle \boldsymbol{a}, \boldsymbol{A}(\boldsymbol{\alpha}_{\boldsymbol{k}} - \boldsymbol{\alpha})\rangle)^2 \tag{6.1.34}$$

The weight of each estimate is given by the factor:

$$w_k(\boldsymbol{\alpha}) = \mathcal{W}\left(\sqrt{\sum_{d=1}^{N_d}(\frac{\boldsymbol{\alpha}_d - \boldsymbol{\alpha}_{\boldsymbol{k},d}}{h_d})^2}\right) \tag{6.1.35}$$

The sum in the equation above runs over the components of the vectors and N_d is the dimensionality of the parameter space. Function $\mathcal{W}$ is $\mathcal{W}(u) = (1 - |u|^3)^3$ for $|u| < 1$ and 0 otherwise. This gives a smoothly decreasing weight to the estimates $\boldsymbol{Y}_k$ obtained at points far away from the evaluation point $\boldsymbol{\alpha}$. The constants h_d control the amount of smoothing, larger values of h_d giving more weight to the estimates further away from the point $\boldsymbol{\alpha}$.

To find the minimum of the criterion C with respect to $\boldsymbol{a}$ we require that all partials $\frac{\partial C}{\partial \boldsymbol{a}_q}$ vanish. This is sufficient to establish the minimum since C is a quadratic function of all $\boldsymbol{a}_q$ with the terms entering with positive coefficients. First write Equation (6.1.34) in terms of components:

$$C = \sum_{k=1}^{N_e} \left(w_k \left(\boldsymbol{Y}_k - \sum_{j=1}^{N_f} \left(\boldsymbol{a}_j \boldsymbol{X}_{jk} \right)^2 \right) \right) \tag{6.1.36}$$

In the equation above the matrix $\boldsymbol{X}$ consists of the vectors $\boldsymbol{A}(\boldsymbol{\alpha} - \boldsymbol{\alpha_k})$ arranged in columns:

$$\boldsymbol{X} = \left\{ \begin{array}{ccc} | & & | \\ \boldsymbol{A}(\boldsymbol{\alpha} - \boldsymbol{\alpha_1}) & \dots & \boldsymbol{A}(\boldsymbol{\alpha} - \boldsymbol{\alpha_{N_e}}) \\ | & & | \end{array} \right\} \tag{6.1.37}$$

Now taking partials of C with respect to all coefficients $\boldsymbol{a}_q$ we obtain N_f equations:

$$\frac{\partial C}{\partial \boldsymbol{a}_q} = 0 \Rightarrow \sum_{k=1}^{N_e} \left(-2 w_k \boldsymbol{Y}_k \boldsymbol{X}_{qk} + 2 w_k \left(\sum_{j=1} N_f \boldsymbol{a}_j \boldsymbol{X}_{jk} \right) \boldsymbol{X}_{qk} \right) = 0 \tag{6.1.38}$$

$$\Rightarrow \boldsymbol{X} \boldsymbol{W} \boldsymbol{X}^T \boldsymbol{a} = \boldsymbol{X} \boldsymbol{W} \boldsymbol{Y} \tag{6.1.39}$$

In the last step we rewrote the system of equations in a matrix form, where the matrix W is a diagonal matrix with entries being the weights w_k:

$$W \equiv \left\{ \begin{array}{ccc} w_1 & & \\ & \ddots & \\ & & w_{N_e} \end{array} \right\} \tag{6.1.40}$$

Solving Equation (6.1.39) for $\boldsymbol{a}$ gives the desired vector $\hat{\boldsymbol{a}}$:

$$\hat{\boldsymbol{a}} = (\boldsymbol{X}\boldsymbol{W}\boldsymbol{X}^T)^{-1}(\boldsymbol{X}\boldsymbol{W}\boldsymbol{Y}) \tag{6.1.41}$$

The value of the PDF at point $\boldsymbol{\alpha}$ is then readily obtained by taking the first component of the vector $\hat{\boldsymbol{a}}$ as can be seen from Equation (6.1.32). Also the first and second partial derivatives can be obtained from the successive components of $\hat{\boldsymbol{a}}$, however these are not used in this analysis.

6.1.3 Implementation and Optimization of the Likelihood Fit

The first step in performing a top quark mass fit is to compute the the values of the KDE estimates for every event i in the dataset. Values of the estimates for signal PDF for event's observable pair $\boldsymbol{v_i}$ are stored for every point in the grid of $(\mathrm{M_{top}}, \Delta_{\mathrm{JES}})$ parameters where the $t\bar{t}$ samples have been generated. Similarly we store the value of background PDF estimates for event's observable pair $\boldsymbol{v_i}$ for every value of the Δ_{JES} parameter where background samples have been generated. Thus for every event i in the dataset we have two sets of numbers: $\left\{P_{s_{KDE}}(\boldsymbol{v_i}; (\mathrm{M_{top}}_k, \Delta_{\mathrm{JES}k})), k \in [1, N_{e_{t\bar{t}}}]\right\}$ $\left\{P_{b_{KDE}}(\boldsymbol{v_i}, \Delta_{\mathrm{JES}k}); k \in [1, N_{e_{background}}]\right\}$ (These are the estimates $\boldsymbol{Y}_k$ in the nomenclature of the previous section). The quantities $N_{e_{t\bar{t}}}$ and $N_{e_{background}}$ are the numbers of respectively signal and background samples generated. Note that the events are divided into subsamples based on the decay channel and number of b-tags and the calculation above is performed separately for each subsample.

We then employ MINUIT [47] package to minimize the value of the negative logarithm of the posterior probability given by Equation (6.1.23). During the minimization process MINUIT package is scanning the parameter space and evaluating the value of the posterior probability at each point in the scan. Therefore at each step for a given subsample the values of $P_s(\boldsymbol{v_i}; (\mathrm{M_{top}}, \Delta_{\mathrm{JES}}))$ and $P_b(\boldsymbol{v_i}; \Delta_{\mathrm{JES}})$ are needed for each of the N events in the subsample. This means that at each step in the minimization for each subsample we need to perform $2N$ LPS fits (N two-dimensional fits for the signal PDF and N one-dimensional fits for the background PDF). Let us focus on obtaining the estimate for the signal PDF. Notice that the LPS fits needed are all for the same pair $(\mathrm{M_{top}}, \Delta_{\mathrm{JES}})$. The grid of points $(\mathrm{M_{top}}\,_k, \Delta_{\mathrm{JES}}\,_k)$, where we have the KDE estimates $\boldsymbol{Y}_k$, is also the same for all events in the subsample. This leads to optimization in terms of the computing time required. Referring to Equation (6.1.41) we find that the matrices $(\boldsymbol{XWX}^T)^{-1}$ and $(\boldsymbol{XW})$ depend only on the point where the estimate for PDF is needed and on the grid of points where KDE estimates are available. At each minimization step we will start by pre-computing these two matrices. Obtaining the LPS estimate for each event amounts then to taking a product of these two matrices with the vector $\boldsymbol{Y}$ which is specific to each event. The computation of LPS estimates for the background PDF is performed in the same fashion.

Computation of the matrix $(\boldsymbol{XWX}^T)$ can be optimized once we notice that most entries in the diagonal matrix W are zero. Recall that the entries w_k are the weights of estimates $\boldsymbol{Y}_k$ that depend on distance of the evaluation point $\boldsymbol{\alpha}$ to the points $\boldsymbol{\alpha}_k$ where the KDE estimates are available. The weight w_k is 0 if the argument of the function $\mathcal{W}$ in Equation (6.1.35) is more than 1, that is if the distance to the given point $\boldsymbol{\alpha}_k$ is large as compared to the smoothing parameters h_d. Since we keep the parameters h_d relatively small, as described in the next section, most of the weights w_k vanish. The matrix $(\boldsymbol{XWX}^T)$ is has size $N_f \times N_f$ (N_f as mentioned above is the number of basis functions for space of second order polynomials). To evaluate $(\boldsymbol{XWX}^T)$ naively we need at least $N_e^2 N_f$ multiplications - a high cost considering that number of $t\bar{t}$ samples is $N_e \sim 3000$. Instead we start with a $\mathbf{0}$ $N_f \times N_f$ matrix. Then for each non-zero weight w_k we increment all entries (i, j) of the matrix by $w_k \times X_{i,k} \times X_{j,k}$. This involves only $N_f^2 N_{e\emptyset}$ multiplication operations, where $N_{e\emptyset}$ is a number of non-zero weights. We can also simplify calculation of the matrix $(\boldsymbol{XW})$ by multiplying every column of this matrix by the corresponding weight w_k. We remove columns corresponding to zero weights, while removing also the corresponding entries from the vector $\boldsymbol{Y}$

6.1.4 Smoothing Parameters for Local Polynomial Smoothing

We have freedom to choose the parameters $h_{\mathrm{M_{top}}}$, $h_{\Delta_{\mathrm{JES}}}$ for the two-dimensional signal PDF smoothing and $h_{\Delta_{\mathrm{JES}},background}$ for the one-dimensional background PDF smoothing. The choice of parameters h_d in Equation (6.1.35) is based on the performance of the method in terms of expected performance and bias.

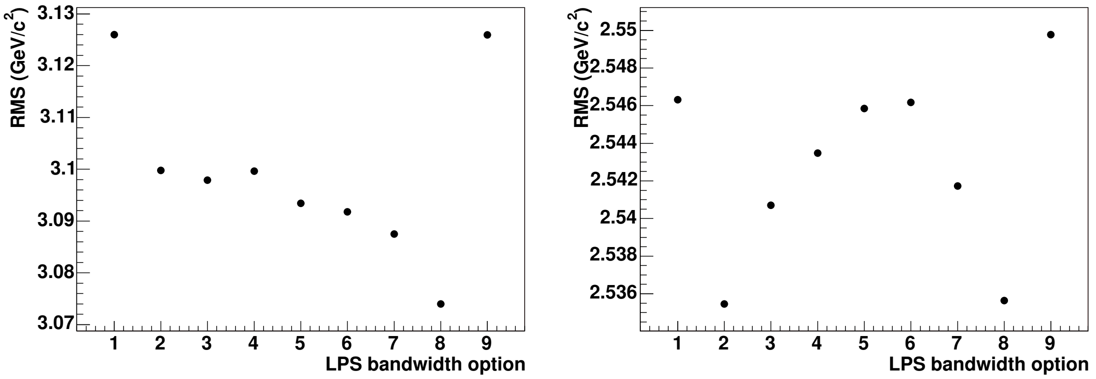
To perform this study we construct pseudoexperiment ensembles using $t\bar{t}$ Monte Carlo sample with the generated top quark mass of 172.5 GeV/c^2. In each pseudoexperiment a number of signal and background events (representative of the expectation in the data) will be drawn from Monte Carlo samples. A top quark mass fit is performed for each simulated dataset. The procedure of pseudoexperiment construction is described in more detail in Section 7.1.

We investigate the options for the smoothing parameters given in Table 6.1. For each of these options a set of 3000 pseudoexperiments is run where we perform two-dimensional fits for $\mathrm{M_{top}}$ and Δ_{JES}. Four fits are performed using the four subsamples separately. A good proxy for the expected uncertainty is the RMS width of the distribution of fitted mass from the pseudoexperiments. This is shown in Figure 6.5. The relative uncertainty on the points is expected to be very small as for each value

Table 6.1: Smoothing parameter options.

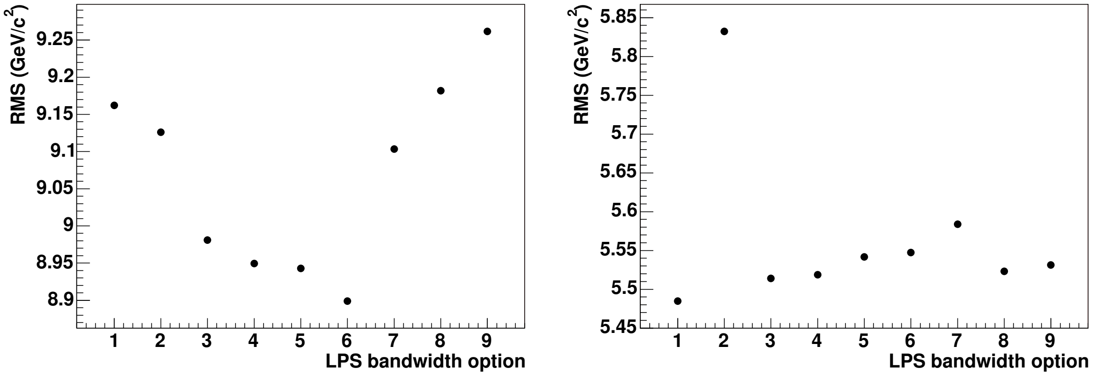
Option	$h_{\mathrm{M_{top}}}$ (GeV/c^2)	$h_{\Delta_{\mathrm{JES}}}$ (σ_c)	$h_{\Delta_{\mathrm{JES}},background}$ (σ_c)
1	8.0	0.8	3.0
2	10.0	0.4	3.0
3	10.0	0.6	3.0
4	10.0	0.8	3.0
5	10.0	1.2	3.0
6	10.0	1.4	3.0
7	12.0	0.8	3.0
8	15.0	0.8	3.0
9	20.0	0.8	3.0

of the option same set of pseudoexperiments was performed. As the $h_{\mathrm{M_{top}}}$ and $h_{\Delta_{\mathrm{JES}}}$ parameters are varied the width of the fitted top quark mass distribution varies by at most 2-3% for $h_{\mathrm{M_{top}}}$ within $10-15$ GeV/c^2 range and $h_{\Delta_{\mathrm{JES}}}$ around $0.8\sigma_c$, therefore exact tuning of the smoothing parameters is not important for the precision of the method. We expect smaller values of the smoothing parameters to perform better in terms of method bias with large smoothing parameter locally forcing the shape of PDF to a parabola. However for small values of smoothing parameters the smoothed PDF becomes sensitive to statistical fluctuations of the KDE estimates caused by limited statistics of the underlying Monte Carlo data sets. We choose $h_{\mathrm{M_{top}}}$ of 10 GeV/c^2 for both lepton + jets samples and 15 GeV/c^2 for both dilepton samples. We set $h_{\Delta_{\mathrm{JES}}} = 0.8\sigma_c$ for all samples. An example of Local Polynomial Smoothing result for a Monte Carlo event in the 0-tag dilepton sample is shown in Figure 6.6. The open points in the graphs indicate the estimates $\boldsymbol{Y}_k$ obtained from KDE. the solid line in each of the panels is the result of LPS. In panels (a) and (b) projections of the two-dimensional signal PDF are shown for a fixed value of Δ_{JES} ($0.0\sigma_c$) and fixed value of $\mathrm{M_{top}}$ (172 GeV/c^2) respectively. The variation of the KDE estimates is larger



(a) Lepton + Jets 1-tag

(b) Lepton + Jets 2-tag



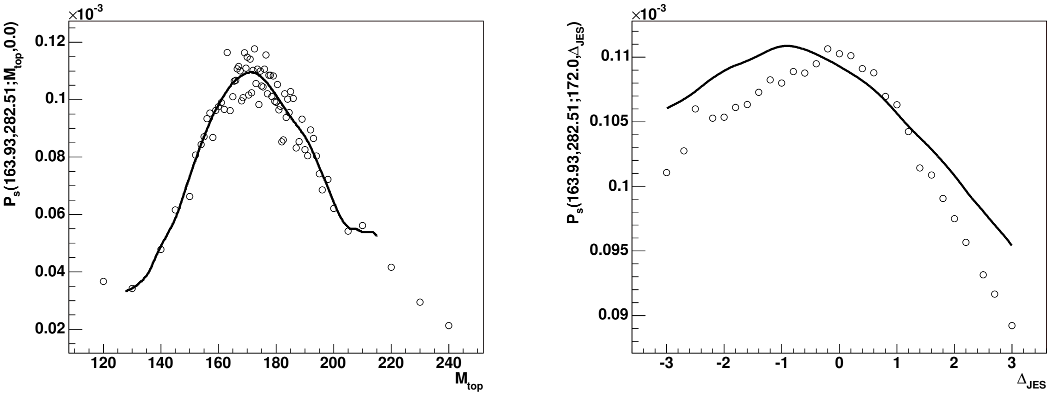
(c) Dilepton non-tagged

(d) Dilepton tagged

Figure 6.5: Width of fitted M_{top} distributions from 3000 pseudoexperiments as a function of the smoothing parameters chosen. The option index is defined in Table 6.1.

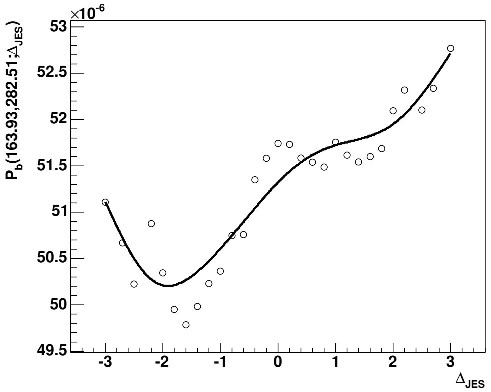
as a function of M_{top} in panel (a) than as a function of Δ_{JES} in panel (b) since as mentioned in Section 2.2.6 the signal samples for a given value of M_{top} but different values of Δ_{JES} are generated form the same underlying Monte Carlo sample. Any fluctuations in the KDE estimates will be correlated for a fixed value of M_{top} and therefore a bias can be seen in panel panel (b). Panel (c) shows the background PDF. Relatively large variation is present in the KDE estimates due to selection effects.

The background model consists of multiple samples with small statistics. Events in these samples are accepted near the jet energy cuts and therefore varying Δ_{JES} can change sample composition resulting in the observed jitter. A large value smoothing parameter $h_{\Delta_{\mathrm{JES}},background}$ is required. We choose $h_{\Delta_{\mathrm{JES}},background} = 3.0\sigma_c$ for all samples. Figures 6.7, 6.8 and 6.9 show projections of signal density functions for selected events from other subsamples: 1-tag dilepton and 1-tag and 2-tag lepton + jets respectively. The fluctuations of KDE estimates for the signal and background PDFs will be correlated among events in a given subsample with similar values of observables. As we see in the figures below the LPS technique largely removes these fluctuations. Since LPS is applied on a per-event basis the fluctuations in KDE estimates will not translate into jitter in the likelihood function for the whole dataset. An alternative approach is to obtain the likelihood function on a grid $(\mathrm{M}_{\mathrm{top}}, \Delta_{\mathrm{JES}})$ where MC samples are available and then fit the log likelihood to a paraboloid. This would involve multiplying the the estimates $\boldsymbol{Y}_k$. In this approach the correlated fluctuations in $\boldsymbol{Y}_k$ would be amplified. For this reason this approach is not viable for the low statistics dilepton samples. In addition the adopted approach with per-event Local Polynomial Smoothing does not force the likelihood form to a Gaussian.



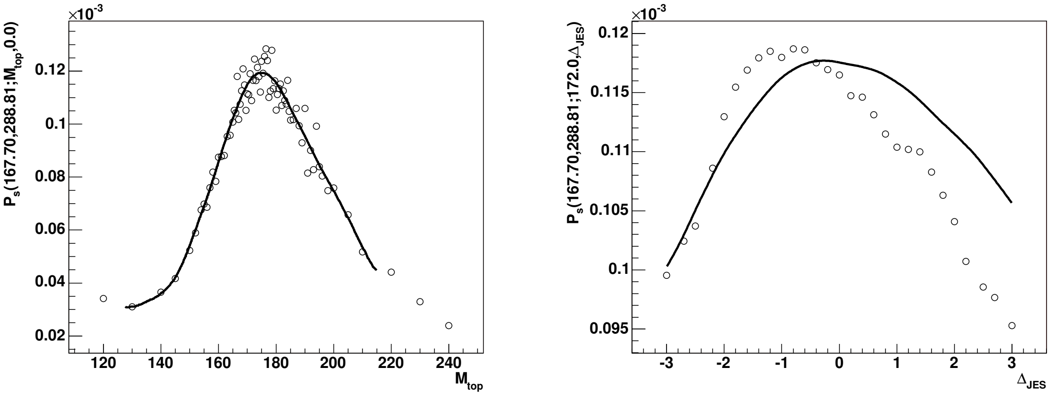
(a) Signal PDF at $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ as a function of $\mathrm{M}_{\mathrm{top}}$

(b) Signal PDF at $\mathrm{M}_{\mathrm{top}} = 172\ \mathrm{GeV}/c^2$ as a function of Δ_{JES}



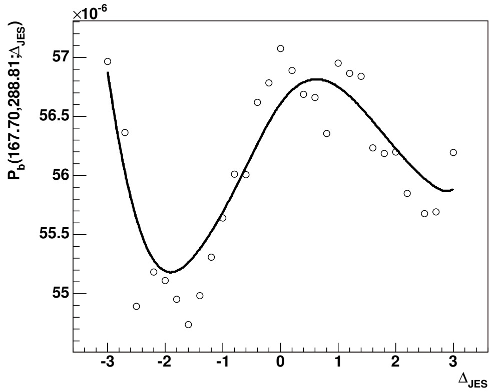
(c) Background PDF as a function of Δ_{JES}

Figure 6.6: Probability density functions for a Monte Carlo 0-tag dilepton event reconstructed at $m_t^{\mathrm{NWA}} = 163.9\ \mathrm{GeV}/c^2$ and $H_T = 282.5$ GeV. Signal PDF projections for (a) $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ and (b) $\mathrm{M}_{\mathrm{top}} = 172\ \mathrm{GeV}/c^2$ are shown. Background PDF is shown in panel (c). See explanation in text.



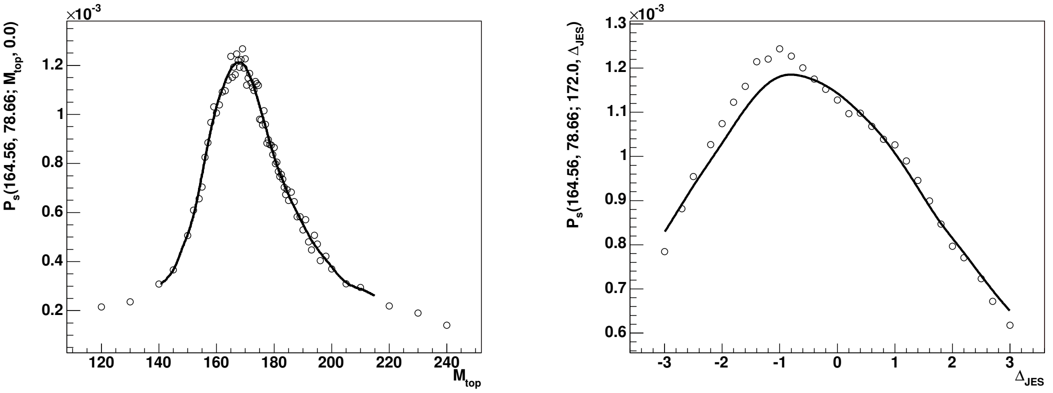
(a) Signal PDF at $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ as a function of $\mathrm{M_{top}}$

(b) Signal PDF at $\mathrm{M_{top}} = 172$ GeV$/c^2$ as a function of Δ_{JES}



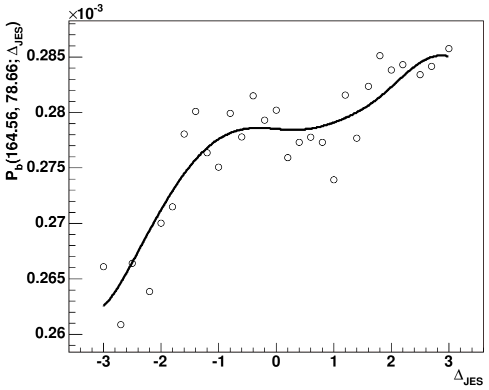
(c) Background PDF as a function of Δ_{JES}

Figure 6.7: Probability density functions for a Monte Carlo tagged dilepton event reconstructed at $m_t^{\mathrm{NWA}} = 167.7$ GeV$/c^2$ and $H_T = 288.8$ GeV. Signal PDF projections for (a) $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ and (b) $\mathrm{M_{top}} = 172$ GeV$/c^2$ are shown. Background PDF is shown in panel (c). See explanation in text.



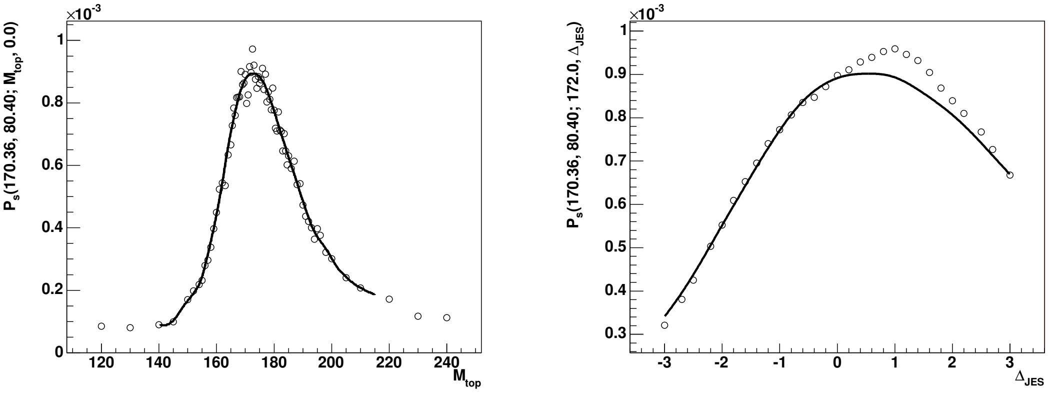
(a) Signal PDF at $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ as a function of $\mathrm{M_{top}}$

(b) Signal PDF at $\mathrm{M_{top}} = 172\ \mathrm{GeV}/c^2$ as a function of Δ_{JES}



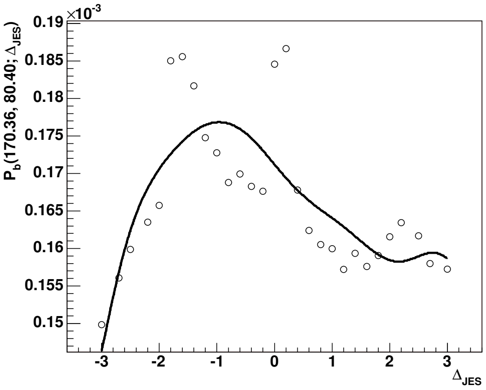
(c) Background PDF as a function of Δ_{JES}

Figure 6.8: Probability density functions for a Monte Carlo 1-tag Lepton + Jets event reconstructed at $m_t^{\mathrm{reco}} = 164.6\ \mathrm{GeV}/c^2$ and $m_{jj} = 78.7\ \mathrm{GeV}/c^2$. Signal PDF projections for (a) $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ and (b) $\mathrm{M_{top}} = 172\ \mathrm{GeV}/c^2$ are shown. Background PDF is shown in panel (c). See explanation in text.



(a) Signal PDF at $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ as a function of $\mathrm{M_{top}}$

(b) Signal PDF at $\mathrm{M_{top}} = 172$ GeV/c^2 as a function of Δ_{JES}



(c) Background PDF as a function of Δ_{JES}

Figure 6.9: Probability density functions for a Monte Carlo 2-tag Lepton + Jets event reconstructed at $m_t^{\mathrm{reco}} = 170.4$ GeV/c^2 and $m_{jj} = 80.4$ GeV/c^2. Signal PDF projections for (a) $\Delta_{\mathrm{JES}} = 0.0\sigma_c$ and (b) $\mathrm{M_{top}} = 172$ GeV/c^2 are shown. Background PDF is shown in panel (c). See explanation in text.

CHAPTER 7

BIAS CHECKS

7.1 Check for Fit Bias

We test the fitting procedure using large number of “pseudoexperiments” where we construct the samples of pseudodata with known values of $\mathrm{M_{top}}$ and Δ_{JES}. The ensembles of pseudodata reflect possible statistical variations of the data sample. In a given pseudoexperiment we select a number of signal (background) events from a Poisson distribution with mean of the expected number of signal (background) in the sample. The $t\bar{t}$ pair production cross section used in constructing the pseudo-experiments is 6.7 pb. The signal events are then drawn at random from a MC sample generated at given mass and jet energy scale values. Background events are drawn from a collection of background models with probabilities corresponding to the contribution of each background model to the total background and the weight of the event within given background model. Note that the drawing of signal as well as background events is done with replacement so that separate pseudoexperiments share subsets of events in their pseudodata. It is also possible for a given event to enter the same pseudodata set more than once. After the pseudodata is constructed we perform a fit as described in the previous chapter. We fluctuate the jet energy prior and the expected number of background events priors for each subsample to reflect the possible deviation of those parameters form their *a priori* values in data.

The jet energy scale constraint in Equation 6.1.23 is replaced by:

$$e^{-\frac{(\Delta_{\mathrm{JES}}-\Delta_{\mathrm{JES}_f})^2}{2}} \tag{7.1.1}$$

where Δ_{JES_f} has been selected from a Gaussian with zero mean and unit width. Similarly each n_{b0q} in Equation (6.1.22) is replaced by a value drawn from a Gaussian with mean of n_{b0q} and width of $\sigma_{n_{b0q}}$.

The likelihood fit should on average return the value of the top quark mass used to generate the pseudoexperiments. Figure 7.2 shows average residuals from 3000 pseudoexperiments run using a range of input top quark masses and values of Δ_{JES}. The results of pseudoexperiments run using pseudodata with $\Delta_{\mathrm{JES}} = 0\ \sigma_{\mathrm{c}}$ are shown in black. For other values of Δ_{JES} marker shapes presented in Figure 7.1 is used. Residuals from the combined, lepton + jets and dilepton fits are shown. The fit to a constant shows no bias for the combined and lepton + jets only fits and a small positive bias for the dilepton only fit. This bias does not warrant a correction since it is small in comparison to the expected uncertainty with probability of 9% for a statistical fluctuation generating the observed shift. The $t\bar{t}$ Monte Carlo samples generated at different values of Δ_{JES} but the same value of $\mathrm{M_{top}}$ share events while the Δ_{JES} value is shifted at reconstruction level as described in section 2.2.6. This leads to large correlations in the pseudoexperiment results therefore we use only the pseudoexperiments generated at $\Delta_{\mathrm{JES}} = 0$ in the fit to a constant in Figure 7.2. The uncertainty on this fit is interpreted as a systematic uncertainty due to limited statistics of the signal Monte Carlo samples.

Value of the fitted jet energy scale shift is compared the value at which the pseudoexperiments were generated at in Figure 7.3. In both the combined fit and the lepton + jets only fit we observe a small negative bias. We do not correct for this bias since it is less than 5% of the observed uncertainty on Δ_{JES} and only slightly incompatible with no bias.

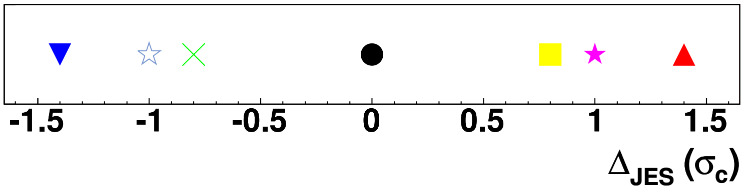


Figure 7.1: Shapes and colors of markers for pseudoexperiment results at different Δ_{JES} value for Figures 7.2, 7.3, 7.4, 7.5 and 7.6

The statistical uncertainty on the measurement is extracted from data. To test whether the error estimate is sound, for each pseudoexperiment we calculate the pull defined as a ratio of the residual to the uncertainty. If the residual is positive (negative) we use the negative (positive) error in the ratio. If the error uncertainty on the measurement is correct the pull distribution will have vanishing mean and unit RMS width. The width of the pull distribution for fitted top quark mass is shown in Figure 7.4. A fit to a constant is performed using ensembles of pseudoexperiments generated at a range of input top quark masses and Δ_{JES} value of 0. The pull width is larger than 1 for the combined and lepton + jets only measurements due to small event statistics used for those two-dimensional fits. It is also slightly larger than 1 in the dilepton only measurement. Derived correction is 3% for the combined and lepton + jets only measurements and 1% for the dilepton measurement. Graphs of the pull width for the jet energy scale are shown in Figure 7.5. Correction on the Δ_{JES} uncertainty is 4% for both the combined fit and lepton + jets only measurement.

The width of the fitted top quark mass distribution from an ensemble of pseudoexperiments is a measure of the expected uncertainty. The RMS width of the fitted top quark mass form pseudoexperiments run at a range of input top quark mass is shown in Figure 7.6. The expected statistical and jet energy scale uncertainty in the

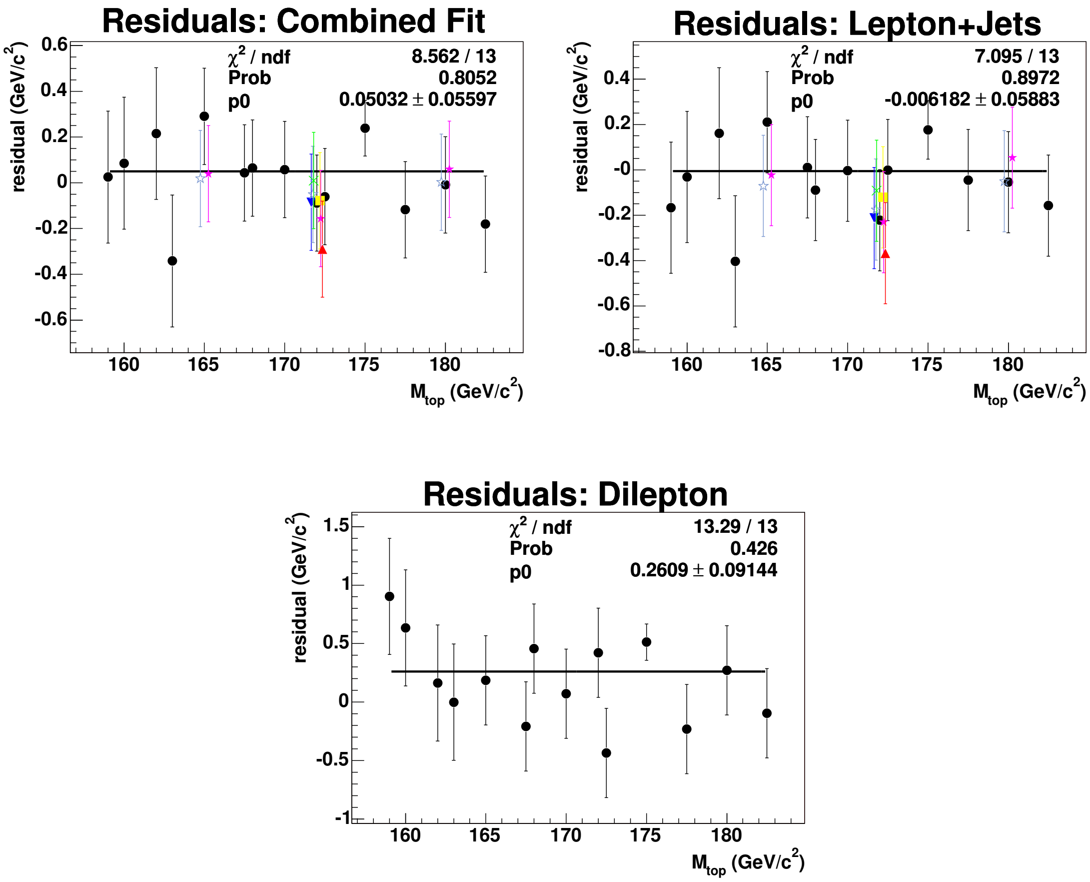


Figure 7.2: Check for bias in the fitted top quark mass for the combined fit, lepton + jets only fit and dilepton only fit.

combined fit is approximately 2.1 GeV$/c^2$ at the input top quark mass of 170 GeV$/c^2$. For the lepton + jets only measurement and for the dilepton only measurement the expected uncertainties are approximately 2.2 GeV$/c^2$ and 4.0 GeV$/c^2$ respectively. Note that the jet energy uncertainty is not included in the dilepton-only fit. The expected uncertainty raises by several percent for all the measurements in the range of input top quark masses considered. Another measure is the median returned error shown in Figure 7.7. Positive reported error and negative reported error are shown. The median reported uncertainty is lower than the RMS of the fitted top quark mass

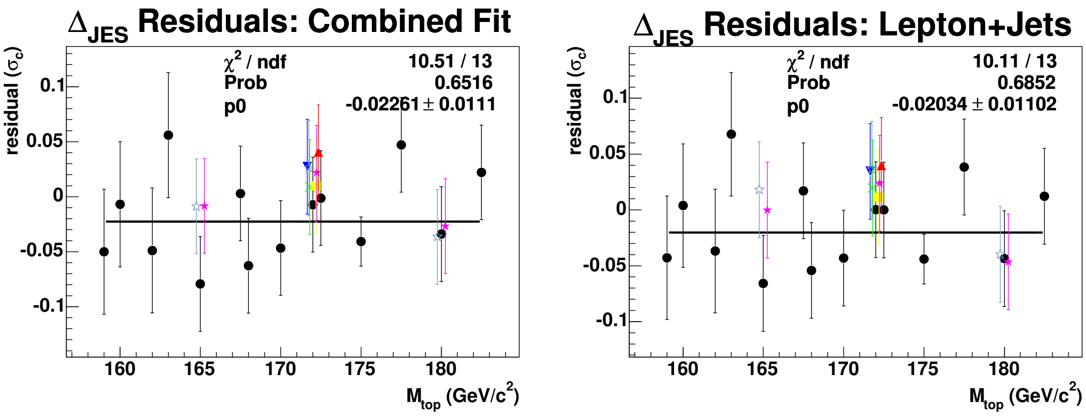


Figure 7.3: Check for bias in the fitted jet energy scale shift for lepton + jets only fit, and combined fit.

distribution by about 0.1 GeV/c^2 for all fits and the median negative error is lower by approximately $1-2\%$ than the median positive error due to non-Gaussian nature of the likelihood function caused by long positive tails in the probability density functions (*cf.* Figures 5.1 and 5.3).

As noted above when ensembles of pseudodata are constructed the events are drawn with replacement from the MC samples. If we were to draw the events from MC samples without replacement such that no two pseudoexperiments were to share events we would only be able to perform approximately 100 pseudoexperiments for given input values of $\mathrm{M}_{\mathrm{top}}$, Δ_{JES} due to limited statistics in the samples. Drawing events with replacement allows us to perform an arbitrarily large number of pseudoexperiments, fully exploring the possible combinations of events in order to identify possible bias. However this resampling does not allow us to reduce the uncertainty on the results of pseudoexperiments such as the value of residual or value of the pull width. To evaluate the uncertainties on those quantities we use bootstrap technique. In each bootstrap run we draw events from signal Monte Carlo sample with replacement until we reach the same number of events as in the original sample. We then

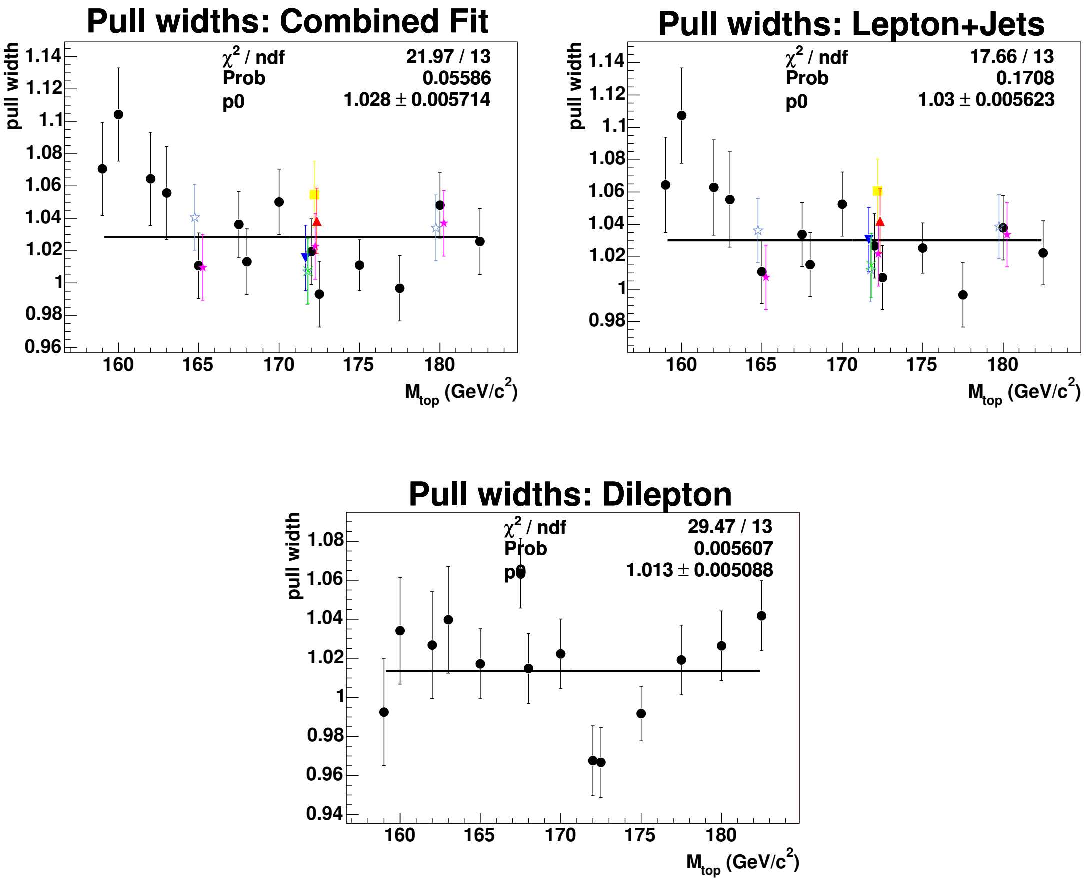


Figure 7.4: Width of the pull distribution for the fitted $\mathrm{M_{top}}$ for the lepton + jets only fit, dilepton only fit and combined fit.

run 3000 pseudoexperiments using this "bootstrapped" sample. We repeat the above procedure multiple times. In each of the bootstrap runs we calculate the desired pseudoexperiment statistic such as the mean residual or pull width. The RMS width of the distribution of that statistic from the bootstrap runs is used as the uncertainty on the statistic. We used the samples generated at $\mathrm{M_{top}} = 160$ GeV/c^2, $\mathrm{M_{top}} = 170$ GeV/c^2 and $\mathrm{M_{top}} = 175$ GeV/c^2 to perform this study since the number of events in those samples is different (between 10^6 and 4×10^6 before event selection). The sample generated at $\mathrm{M_{top}} = 170$ GeV/c^2 is typical and bootstrap results from this sample

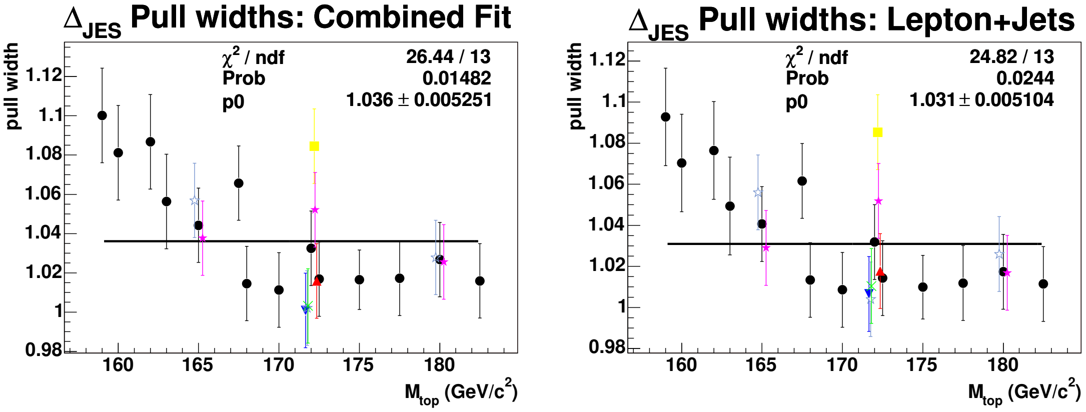


Figure 7.5: Width of the pull distribution for the fitted Δ_{JES} for the lepton + jets only fit, dilepton only fit and combined fit.

are used to estimate pseudoexperiment uncertainties in most other samples. Figures 7.8, 7.9 and 7.10 show the results of the bootstrap run for the fitted top quark mass bias, pull width and RMS uncertainty and Figures 7.11 and 7.12 show the results for the jet energy scale bias and pull width. Each entry in the histograms in these figures is extracted from 3000 pseudoexperiments run using one of the bootstrapped samples obtained from $\mathrm{M_{top}} = 170$ GeV/c^2 sample.

The bootstrap technique is also used to evaluate the systematic uncertainty of limited statistics in the background samples. Now instead of the resampling the underlying signal Monte Carlo sample we “bootstrap” the background model. Even though the events in the background samples have different weights each event is equally likely to enter the bootstrap set. The mean residual of fitted top quark mass is shown in Figure 7.13.

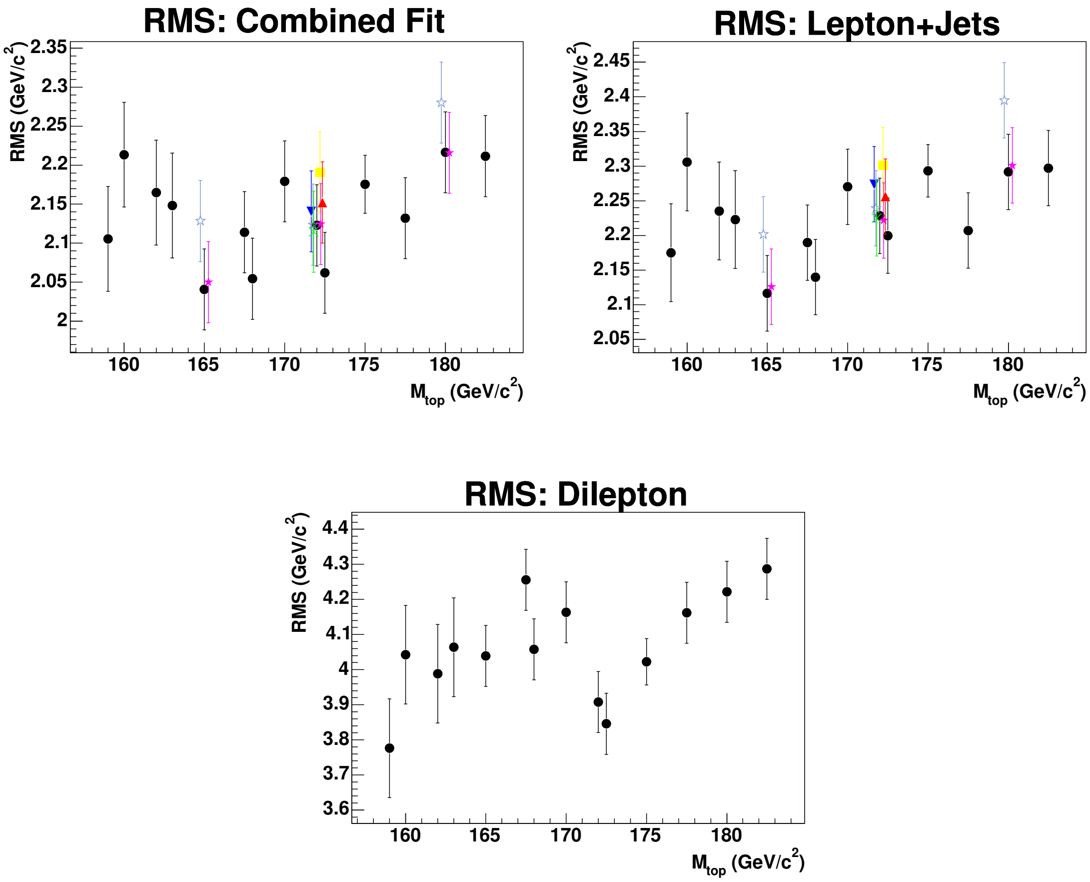


Figure 7.6: RMS width of the fitted mass distribution from the combined fit, lepton + jets only fit and dilepton only fit.

7.2 Bias Due to Imperfect Instantaneous Luminosity Profile

As noted in Section 2.2.6 the instantaneous luminosity profile used to generate Monte Carlo samples corresponds well to the instantaneous luminosity profile of data but only through the first 1.2 fb^{-1}. The instantaneous luminosities were increasing through the data taking and the second half of the dataset used for this measurement has been taken at higher average luminosities. We use number of z-vertices to study this effect. A z-vertex is an intersection of tracks in the $r - z$ plane. It is a good proxy for number of $p\bar{p}$ interactions present in a given event even for high

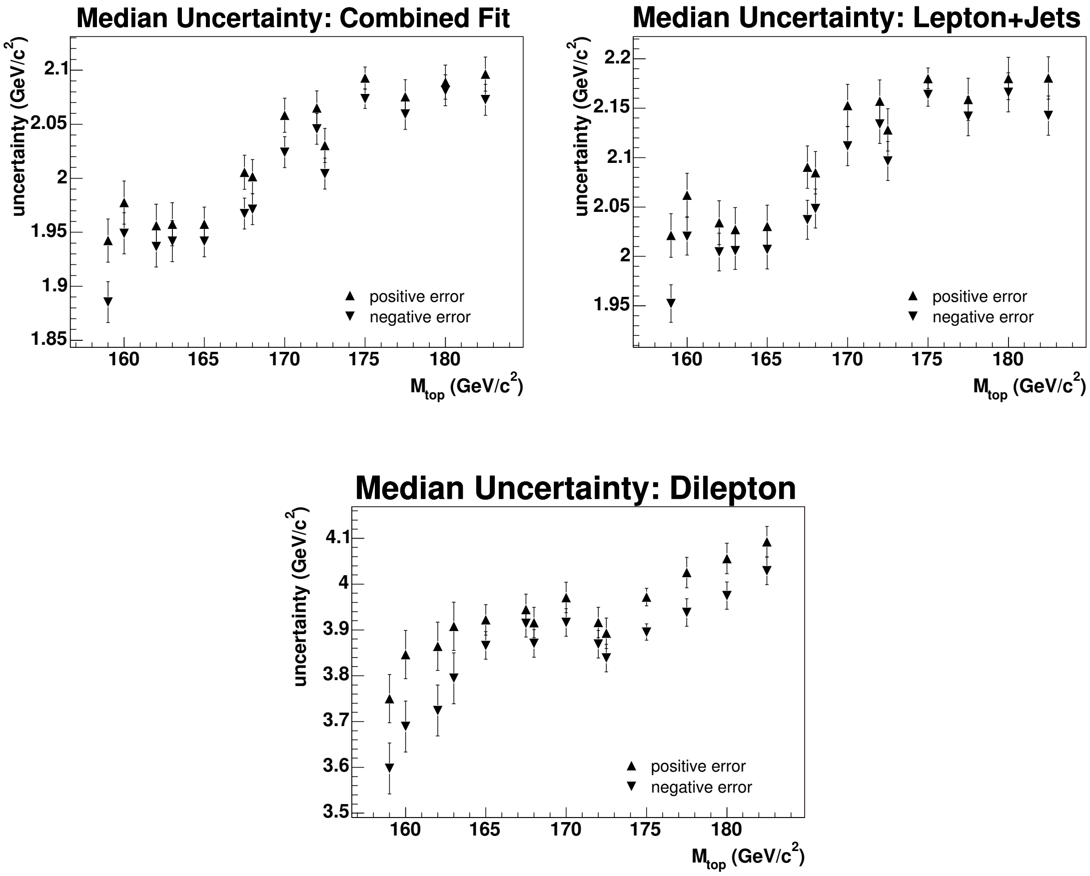


Figure 7.7: Median of the reported statistical uncertainty distribution from the combined fit, lepton + jets only fit and dilepton only fit. Only results with input $\Delta_{\mathrm{JES}} = 0\ \sigma_{\mathrm{c}}$ are shown.

instantaneous luminosities as $p\bar{p}$ interactions will generally be spread out along the beamline. To study this effect we divide the signal Monte Carlo sample generated at $\mathrm{M_{top}}$ =175 GeV/c^2 by number of z-vertices. Number of events with more than 3 z-vertices in the sample is very small therefore another sample was generated with high number of minimum bias $p\bar{p}$ interactions added. We run pseudoexperiments where the pseudodata belongs to the samples separated by number of z-vertices. Figures 7.14 and 7.15 show the results. The sample with artificially enhanced number of interactions was used in the >=4 vertices bin. In that bin results are plotted at the

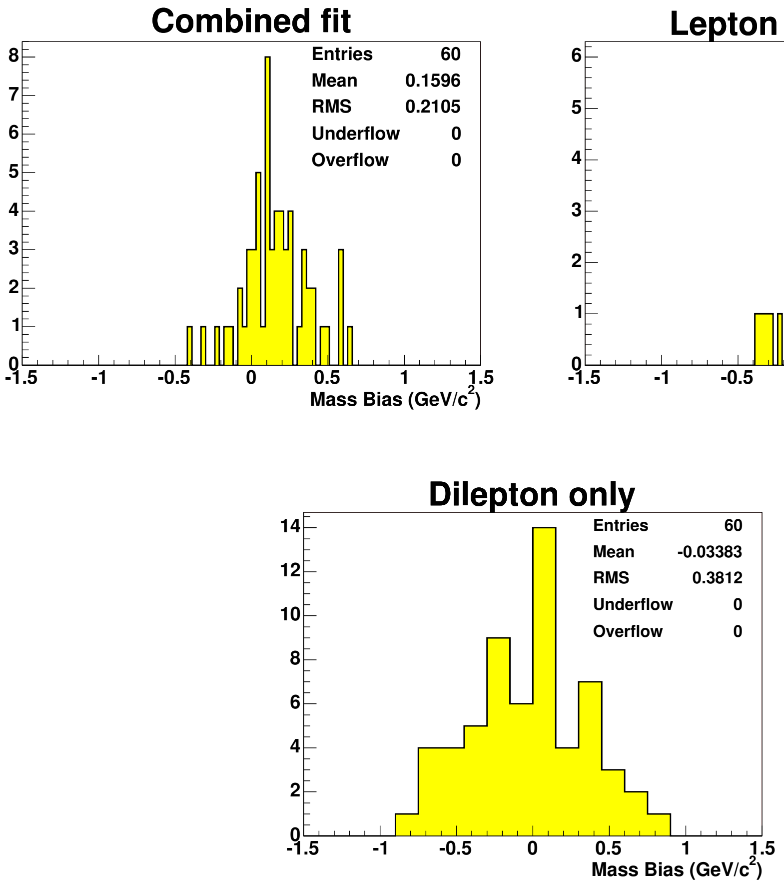
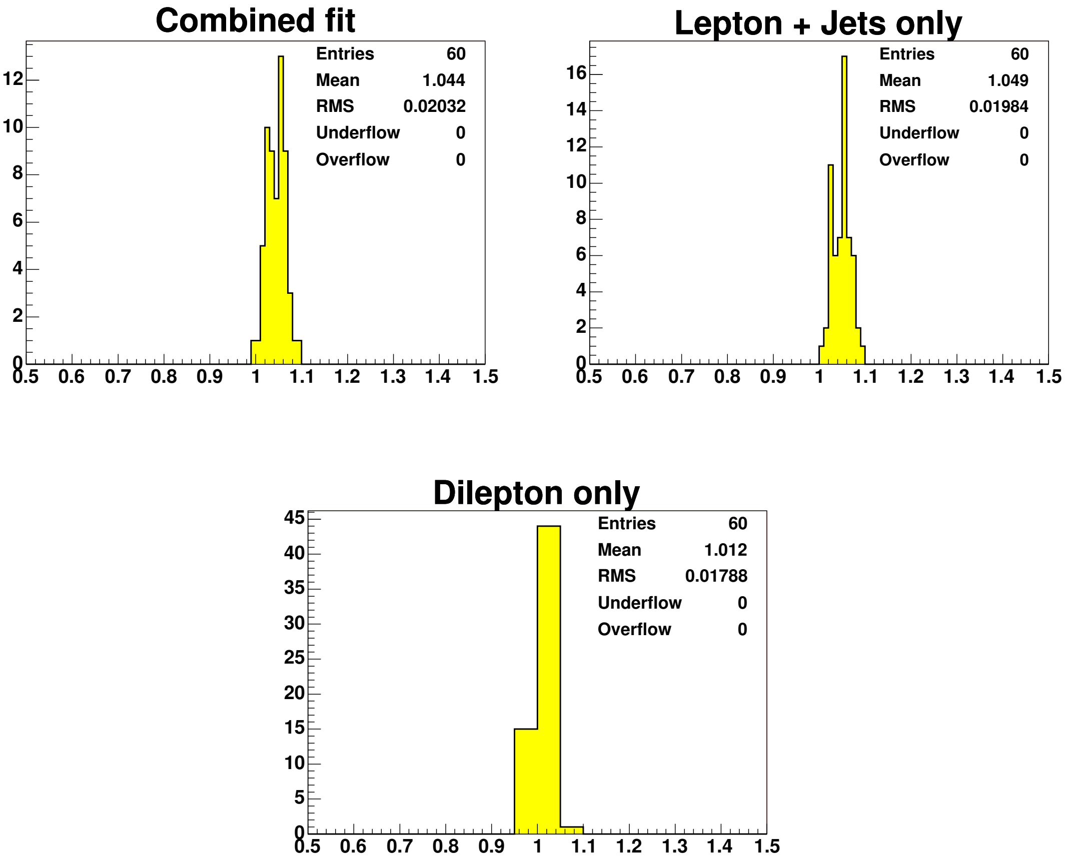


Figure 7.8: Results of 60 runs of the bootstrap technique with 3000 pseudoexperiments in each run. Mean mass residual from each run is plotted for the combined fit, lepton + jets only fit and dilepton only fit.

average number of vertices on the abscissa. Uncertainties on the points are obtained by scaling the relevant bootstrap results by a a square root of the fraction of events in a given bin. A clear trend is observed in the dilepton result. In the data we observe an average of 1.93 z-vertices per event and in the signal Monte Carlo sample we have an average of 1.50. Based on the difference between those two numbers we calculate that 0.40 GeV/c^2 needs to be subtracted from the dilepton only result obtained in data. The top mass result does not need to be corrected for the combined and lepton + jets only fit, however the quoted value of Δ_{JES} will be shifted low by 0.04 σ_c.


Combined fit

Lepton + Jets only

Dilepton only

Figure 7.9: Mass pull width from 60 runs of the bootstrap technique for the combined fit, lepton + jets only fit and dilepton only fit.

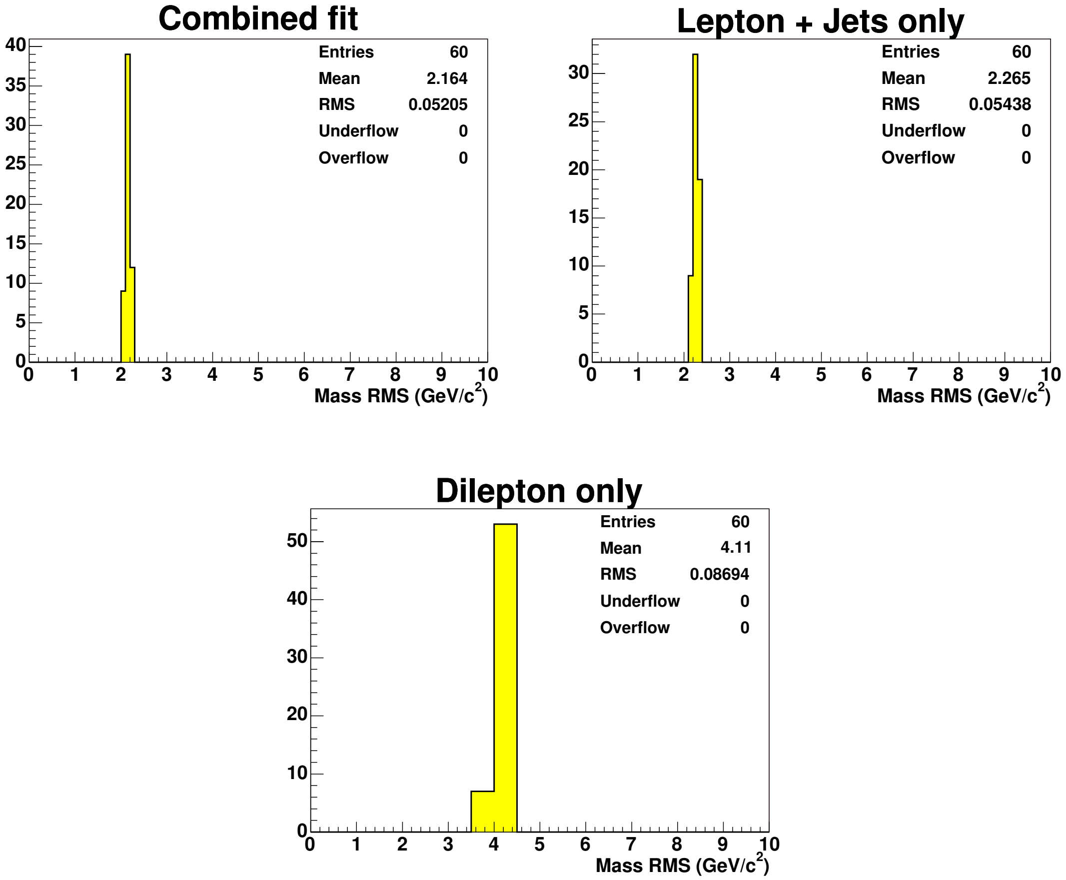


Figure 7.10: RMS of the fitted top quark mass distribution 60 runs of the bootstrap technique for the combined fit, lepton + jets only fit and dilepton only fit.

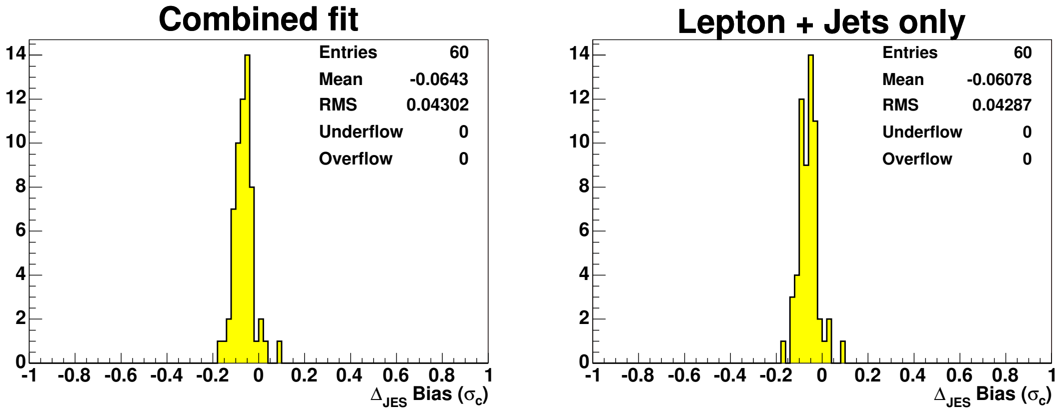


Figure 7.11: Mean Δ_{JES} residual from 60 runs of the bootstrap technique for the combined fit and lepton + jets only fit.

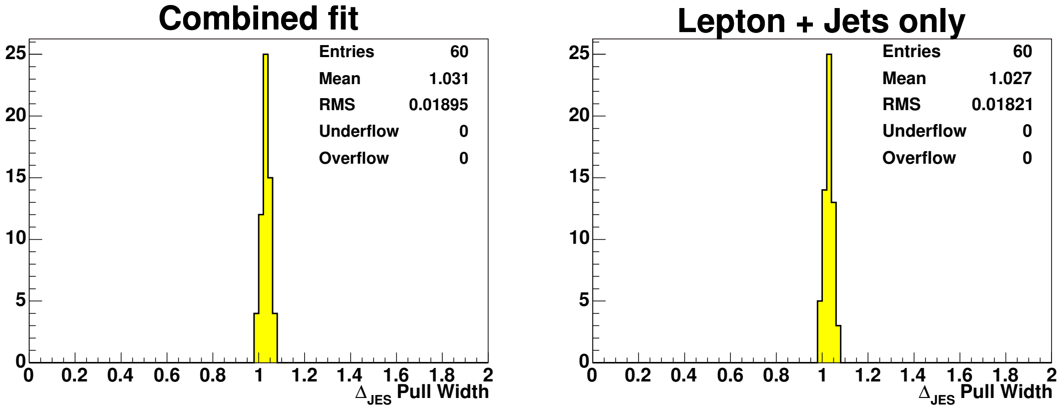


Figure 7.12: Δ_{JES} pull width from 60 runs of the bootstrap technique for the combined fit and lepton + jets only fit.

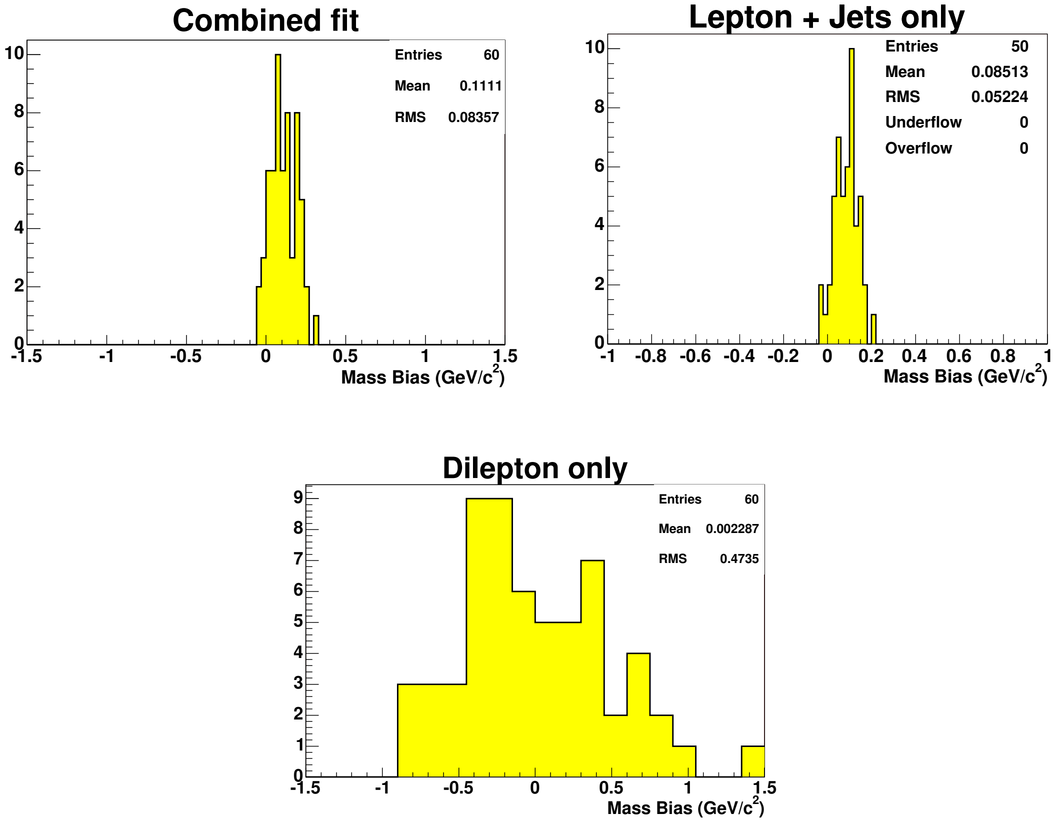


Figure 7.13: Mean mass residual from multiple runs of bootstrap procedure where the background was resampled for the combined fit, lepton + jets only fit and dilepton only fit.

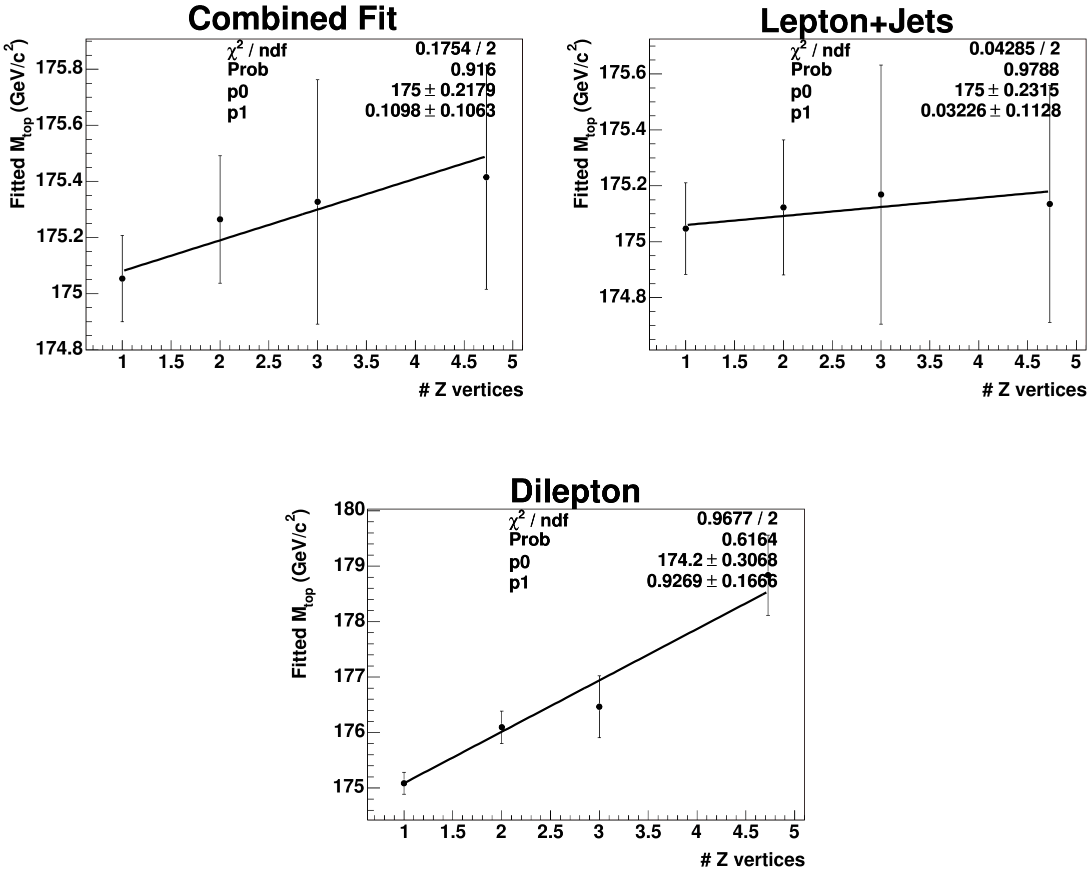


Figure 7.14: Dependence of the fitted top mass on the number of z-vertices. Combined, lepton + jets and dilepton samples.

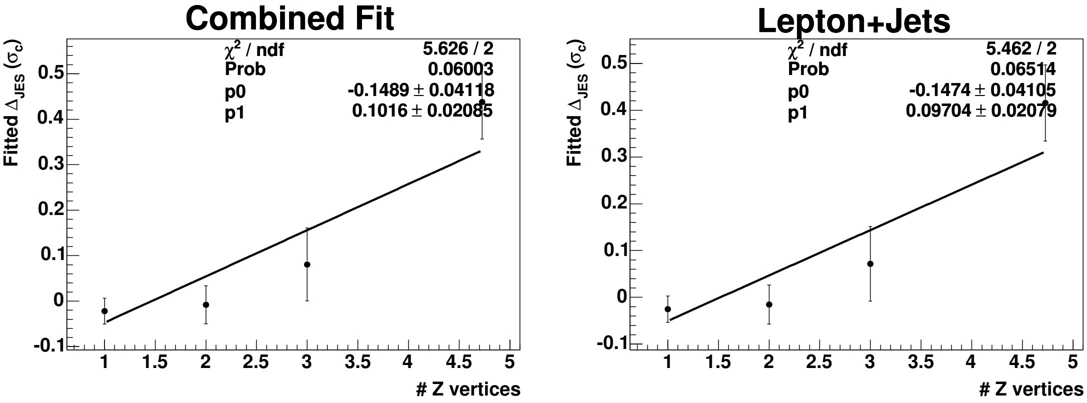


Figure 7.15: Dependence of the fitted Δ_{JES} mass on the number of z-vertices. Lepton + jets only and combined samples.

CHAPTER 8

SYSTEMATIC UNCERTAINTIES

We examine a series of effects that may bias our top quark mass measurement. A general procedure to obtain a systematic uncertainty associated with a particular effect is to create two samples with this effect increased and decreased by one unit of it's uncertainty. We then run two sets of pseudoexperiments, each set using the sample where the effect has been shifted up or down as pseudodata. We take the half difference between the mean fitted top quark mass from both sets as a systematic. For large systematics the effect on background is also modeled.

In many cases it is possible to construct the pseudodata by reweighting a sample. This means that the probabilities to draw events in pseudoexperiments are modified. In other cases we may apply a modified reconstruction to a sample. Both approaches are preferred since the resulting uncertainty (due to the statistics of the sample) on the calculated half difference will be small. When entirely new Monte Carlo sample has to be generated to model a particular effect statistical uncertainties may dominate the calculated half difference. In such cases we take the statistical uncertainty on the half difference as the systematic. In addition when the positive and negative shift result in a shift in the same direction in the mean fitted mass from the mean fitted mass when using a “nominal” sample, half the maximum difference from nominal is taken as a systematic. The Monte Carlo sample with the highest statistics (4.8 million events before selection) available to us was generated at $\mathrm{M_{top}} = 175$ GeV/c^2 and it will be used as the nominal sample. Additional samples needed were generated with this input top quark mass also.

8.1 Residual Jet Energy Scale

The method of using the W resonance to calibrate the jet energy scale depends on two assumptions. First we assume that different effects contribute to the unknown systematic shift in such a way that jets of different momenta and pseudorapidities are all shifted by common fraction of σ_c. We also assume that shifts are 100% correlated as a function of the jet p_T. That is we assume that all information about the systematic miscalibration of b quark jets (generaly high in p_T) is contained in the shift of the W daughter jets (at medium and low p_T). We test how the measurement behaves when these assumptions are broken and compute the residual JES uncertainty. We prepare two sets of of pseudodata. In the first set we increase and decrease the size of each jet energy scale effect up and down by 1 σ. This breaks the first assumption as the separate JES contributions will not have the same effect on jet energies as a function of p_T and η as their sum in quadrature.

The effect of systematic shifts being not fully correlated is expected to be the most important for the out-of-cone energy flow. The uncertainty on jet momentum due to this effect is of the order of 10% for low p_T jets, has a falling exponential shape until it becomes constant at p_T of 70 GeV/c (*cf.* Fig. 3.4). This imples that we can be artificially strongly constraining jets at high momenta using the jets at low momenta even if shifts at high and low momenta are de-correlated by only a small amount. In addition the uncertainty on the out-of-cone energy flow is derived using differences in data and Monte Carlo samples and not by estimating some set of parameters describing modeling of the jets, therefore there is no *a priori* reason to believe that the out-of-cone jet energy scale shifts are correlated. On the other hand the main effect impacting the out-of-cone energy flow is understood to be modelling of gluon radiation. This implies that the systematic shifts in jet energies due to out-of-cone energy flow are highly correlated between jets of different p_T in high momentum range. At high momenta jets are collimated and therefore the uncertainty is on the radiation at similar angle to the jet axis (or at similar momentum in the direction perpendicular to the jet axis). However in the low momentum region jets of different momenta will occupy different portion of the jet cone therefore in this region the shifts in jet energies are not neceserily correlated. Exact study of these correlations is not possible and we are forced to assume a reasonable form for them. Let us denote by $s_{ooc}(p_T)$ a relative systematic shift for a jet o momentum p_T (at parton level). We assume that the correlation between shifts at different momenta has the following form:

$$\rho(s_{ooc}(p_{T1}), s_{ooc}(p_{T2})) = k(p_{T1})^{(p_{T2}-p_{T1})}. \tag{8.1.1}$$

In the equation above p_{T1} is taken to be smaller of the two momenta. The coefficient k is a linear function of p_{T1} in the range $0 - 50$ GeV/c and becomes constant for $p_{T1} > 50$ GeV/c. We choose k so that $\rho(s_{ooc}(0), s_{ooc}(100)) = 0.5$ and $\rho(s_{ooc}(50), s_{ooc}(150)) = 0.8$. Since the jet momentum spectrum is continuous we have infinite number of random variables $s_{ooc}(p_T)$. The correlation is nearly 100% for separation of 1 GeV/c and varies little for separation of several GeV/c therefore we will consider a finite set of random variables - the shifts s_i at integer transverse momenta in the range $8 - 300$ GeV/c. Lowest energy jets used in this analysis have particle level momenta of 8 GeV/c and there will be essentially no jets above 300 GeV/c in p_T. Having this finite set of random variables, their standard deviations given by the size of out-of-cone uncertainty, and correlations given by Equation 8.1.1 we write a covariance matrix for them. Next we apply so called Princpal Component Analysis procedure. We find 293 eigenvalues λ_j of the covariance matrix and a set of 293 orthonormal eigenvectors v_j. We order the eigenvalues from largest to smallest and form a new set of random variables given by:

$$q_j = \sum_{i=1}^{293} v_{j,i} \times s_i \tag{8.1.2}$$

The covariance matrix for such defined random variables q is a diagonal matrix consisting of the eigenvalues λ, therefore the variables q are not correlated. This means that we can study the effect of jet shifts being not fully correlated at different momenta, by constructing pseudodata with variables q shifted up and down by unit of their uncertainty. Variable q_j taking on a value of it's uncertainty $\sqrt{\lambda_j}$ is equivalent to variables s_i taking on values $s_i = \sqrt{\lambda_j} v_{j,i}$. Figure 8.1 shows relative shift in jet momentum induced by 1 σ variation in the out-of-cone energy flow in thick black dashed curve. The thin solid curves show relative shifts induced by 1 σ shifts in successive q_j variables. Sum in quadrature of the shifts induced by first five variables q_j is drawn in a dotted curve showing that the first five q_j are sufficient to cover the out-of-cone systematic (the dotted curve almost exactly overlaps the thick dashed curve). Values of the shifts for non-integer momenta are obtained by linearly interpolating between the integer momenta. We therefore generate 10 additional pseudodata samples (both signal and background) where the first five variables q_j have been shifted up and down by unit of their uncertainty. Note that these pseudodata samples as well as the pseudodata samples where separate JES contributions were varied are constructed at selection level, taking into account possible threshold effects due to events with low energy jets entering and leaving the samples. Pseudoexperiments are run using both sets of pseudodata. In the lepton + jets fit and the combined fit the jet energy prior is turned off. If the prior remained it would force the variation in the fitted top mass to be artificially high.

The residual jet energy scale systematic for the combined and lepton + jets fits

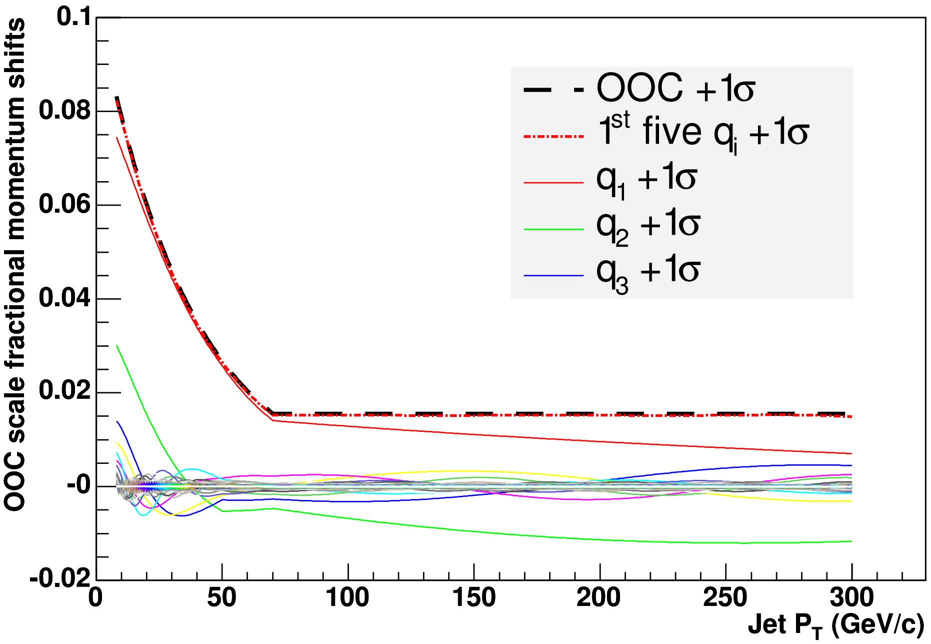


Figure 8.1: Relative shift on jet momentum induced by 1 σ variation in out-of-cone energy flow (Thick dashed black line); Relative shifts induced by 1 σ variations of variables q_j (thin solid lines); Sum in quadrature of the shifts induced by first five variables q_j (dotted line)

is constructed by adding in quadrature half differences of pseudoexperiment results where the absolute, relative, underlying energy and splash out energy scales as well as the q_j variables have been increased and decreased by 1 σ. For the dilepton measurement which has no *in situ* JES calibration we use half difference between pseudoexperiment results where the out-of-cone energy scale has been varied instead of the q_j variables. In all fits the pseudoexperiments generated with increased and decreased pileup energy scale fit slightly lower than nominal therefore the base pileup uncertainty is taken to be half of the largest shift from the nominal. The total residual JES uncertainty in the combined fit is 0.68 GeV/c^2, 0.69 GeV/c^2 in lepton + jets fit and 3.49 GeV/c^2 in the dilepton fit.

8.2 *b* Quark Energy Scale

We identify three sources of uncertainty related to heavy quark jet modeling: the b quark fragmentation, semileptonic decay branching fractions of b and c quarks differences in calorimeter response to b jets.

The Bowler fragmentation model [53] gives the probability density function of z in terms of the fragmentation function:

$$f(z) \propto \frac{1}{z^{1+r_Q b m_q^2}}(1-z)exp\left(\frac{-b(m_q^2+p_T^2)}{z}\right), \tag{8.2.1}$$

where z is the fraction of the b quark momentum carried by the B hadron. In the equation above m_Q is mass of the B hadron m_q is the mass of b quark, a, b, r_Q are free parameters and p_T is the momentum of the B hadron in direction transverse to the momentum of the b quark. The DØ collaboration has used the data from ALEPH, DELPHI and OPAL (ADO) as well as data from the SLD experiments to tune the

Table 8.1: Mean fitted top quark mass from pseudoexperiments where the jet energy scale effects were varied. Mean $\mathrm{M_{top}}$ results have units of GeV/c^2 and mean Δ_{JES} results have units of σ_c. The Abbreviations used are: MI - Multiple Interactions, UE - Underlying Event, OOC - Out-Of-Cone. Comb refers to the combined fit, LJ refers to the lepton + jets fit and DIL refers to the dilepton fit.

Sample:	LJ		DIL	Comb	
Mean fitted:	$\mathrm{M_{top}}$	Δ_{JES}	$\mathrm{M_{top}}$	$\mathrm{M_{top}}$	Δ_{JES}
Nominal $\mathrm{M_{top}} = 175.0$	175.18	-0.04	175.51	175.24	-0.04
Relative JES +1σ	175.18	0.16	176.05	175.21	0.17
Relative JES -1σ	175.16	-0.21	174.77	175.21	-0.21
MI JES +1σ	175.13	-0.01	175.48	175.16	0.00
MI JES -1σ	175.14	-0.05	175.46	175.21	-0.04
Absolute JES +1σ	175.54	0.51	177.77	175.56	0.53
Absolute JES -1σ	174.61	-0.52	173.26	174.68	-0.52
UE JES +1σ	175.13	0.05	175.79	175.18	0.06
UE JES -1σ	175.24	-0.13	175.1	175.28	-0.12
OOC JES +1σ	174.83	0.76	178.03	174.88	0.77
OOC JES -1σ	175.38	-0.78	172.95	175.41	-0.78
Splash-out JES +1σ	175.07	0.11	175.84	175.09	0.12
Splash-out JES -1σ	175.22	-0.16	175.16	175.29	-0.16
q_1+1σ	174.76	0.7	177.73	174.80	0.72
q_1-1σ	175.54	-0.75	173.33	175.60	-0.75
q_2+1σ	174.89	-0.09	174.89	174.93	-0.09
q_2-1σ	175.43	0.05	176.00	175.48	0.06
q_3+1σ	175.17	-0.14	175.08	175.22	-0.14
q_3-1σ	175.21	0.06	175.73	175.24	0.07
q_4+1σ	175.36	-0.12	175.35	175.38	-0.11
q_4-1σ	175.05	0.02	175.57	175.10	0.03
q_5+1σ	175.23	-0.01	175.64	175.29	0.00
q_5-1σ	175.08	-0.05	175.35	175.14	-0.05

Bowler model used in PYTHIA generator [54]. We reweight the signal sample to effectively replace the b fragmentation function used by default by PYTHIA generator with the one obtained from either the ADO data or SLD data (as calculated by the DØ collaboration). Pseudoexperiments using the ADO reweighted sample have largest shift from nominal for all fits and this difference is taken as a systematic due to b fragmentation: 0.13 GeV/c^2 for the lepton + jets fit, 0.08 GeV/c^2 for the dilepton fit and 0.14 GeV/c^2 for the combined fit.

In the weak decays of b and c quarks neutrinos can be produced which escape the collision undetected. Uncertainties on the fraction of these semileptonic decays will then affect the measured top quark mass. The value of the branching fraction of a b quark to a muon or electron is:

$$\mathrm{Br}(b \to l) = 0.1071 \pm 0.0022, \tag{8.2.2}$$

as measured in Reference [55]. The branching fraction of b quarks to τ leptons is [2]:

$$\mathrm{Br}(b \to \tau) = 0.0248 \pm 0.0026, \tag{8.2.3}$$

giving the total branching fraction of b quarks to leptons of 0.2390±0.007. Uncertainties have been added linearly. This is in agreement with the fraction of semileptonically decaying b quarks of 0.2396 in the PYTHIA $t\bar{t}$ Monte Carlo sample generated at $\mathrm{M_{top}}$ =175 GeV/c^2. We then reweight this Monte Carlo sample so that the $b \to e, \mu, \tau$ decay branching fraction varies by ±0.007.

The b quark can also produce leptons and neutrinos through cascade decays:

$b \to c \to l$. The branching fraction of this mode is:

$$\mathrm{Br}(b \to c \to l) = 0.0801 \pm 0.0018. \tag{8.2.4}$$

The branching fraction of charm quark to lepton is

$$\mathrm{Br}(c \to l) = 0.0969 \pm 0.0031, \tag{8.2.5}$$

as quoted in [55], where in both equations l is an electron or a muon. In the signal Monte Carlo sample we find on average 2.2 charm hadrons per event from the b decays. To estimate the expected number of charm hadrons from the W decay, we sum the expected number of c quarks from W decay in a given $t\bar{t}$ decay channel over all the $t\bar{t}$ decay channels. Branching ratio of $t\bar{t}$ system into all hadronic decay mode is 0.46; 0.43 into lepton + lets mode and 0.11 into dilepton mode. Note that τ leptons are counted in the decay channels. We expect 0.5 of W decays to involve a c quark therefore we expect average of 0.5 and 1.0 charmed hadrons from W decays in the lepton + jets mode and all hadronic mode respectively. Therefore we expect on average 0.5·0.43+1.0·0.46=0.68 charm hadrons from W decay in $t\bar{t}$ events. To obtain the average number of semileptonic charm decays from both W decay and cascade b quark decay per event, we use the semileptonic branching fractions 8.2.4 and 8.2.5 and obtain:

$$\begin{gathered}\text{average semileptonic charm decays per event} = \\ ((0.0969 \cdot 0.68) + (0.0801 \cdot 2.2))/(0.68 + 2.2) = 0.168.\end{gathered} \tag{8.2.6}$$

The value found in the signal Monte Carlo sample is 0.188 and therefore we reweight the signal sample so that the amount of semileptonic charm decays is increased or decreased by 2% (in absolute terms). This variation is performed coherently with the variation of the b semileptonic decay fractions resulting in 0.06 GeV/c^2 uncertainty for lepton + jets fit, 0.17 GeV/c^2 uncertainty for dilepton fit and 0.07 GeV/c^2 uncertainty for the combined fit.

A different calorimeter response to b jets and light flavor jets is in principle not a source of uncertainty as jets are treated identically in the data and Monte Carlo simulations. However the uncertainties on the absolute jet energy correction are derived by varying parameters of the light flavor jets. The same procedure as in [42] is applied to b jets in $t\bar{t}$ Monte Carlo samples to derive the absolute energy scale uncertainty. The maximum difference in absolute energy scale uncertainty between the light flavor jets and b jets is found to be 0.2% of jet p_T. To propagate this possible difference to the uncertainty on the top quark mass measurement we shift the energies of the jets matched to b quark up and down by 1% in the signal Monte Carlo sample. We run pseudoexperiments with pseudodata prepared this way and take 20% of the half-difference between the results as the systematic due to b jet calorimeter response. We obtain 0.18 GeV/c^2 for the lepton + jets fit, 0.24 GeV/c^2 for the dilepton fit and 0.19 GeV/c^2 for the combined fit.

8.3 Pileup

At the Tevatron accelerator multiple interactions can occur in a single bunch crossing. The interactions additional to the hard scattering process of interest will deposit energy inside the jet cones. As described in Section 3.5 this effect ought to be removed by the multiple interactions corrections. However we observe that the multiple interaction correction doesn't fully remove the effect of additional interactions in a signal Monte Carlo sample. We observe 250 MeV/jet/vertex slope for jets corrected for multiple interactions (this response should be 0) Uncertainty on the multiple interaction correction is 107 MeV/jet/vertex, which doesn't cover this effect - the multiple interactions correction uncertainty should be scaled by a factor of 2.3. The additional complication comes from the fact that the Monte Carlo samples do not match instantaneous luminosity profile of the data. We observe 1.93 z-vertices in data and 1.50 z-vertices in Monte Carlo samples. Therefore we need to scale the multiple interactions uncertainty further by (0.93/0.50). Such scaling gives resulting pileup systematic of 0.10 GeV/c^2, 0.11 GeV/c^2 and 0.07 GeV/c^2 for the lepton + jets, dilepton and combined fits respectively.

8.4 Background Shape

To study the effect of the dilepton background sample composition on the dilepton fit and on the combined fit we perform six sets of pseudoexperiments where we vary given type of background (DY, diboson and fakes) up or down by uncertainty on it's estimate while holding the total number of background constant. We take half the difference from each of the shifts and sum in quadrature. We obtain 0.11 GeV/c^2 uncertainty for the dilepton fit and 0.06 GeV/c^2 for the combined fit. The results of the pseudoexperiments are shown in Table 8.2.

The 30% uncerainty on the electron and muon fake rates are obtained by measuring the difference in the calculated fake ratio in a given E_T bin between different QCD samples. The origin of the large discrepancy in the fake ratios is not known, however the cause must be related to the energy available in the event and therefore will be correlated to the E_T of the fakable object. We examine the impact of this effect on the dilepton and combined fits by reweighting the dilepton fake template

Table 8.2: Results of pseudoexperiments when dilepton background composition is shifted. Masses are in units of GeV/$c^2$and Δ_{JES} are in units of σ_c. Comb refers to the combined fit and DIL refers to the dilepton fit.

Sample:	DIL	Comb	
Mean fitted:	M_{top}	M_{top}	Δ_{JES}
$+\sigma_{D-Y}$	175.59	175.24	-0.03
$-\sigma_{D-Y}$	175.41	175.19	-0.02
$+\sigma_{Diboson}$	175.43	175.22	-0.02
$-\sigma_{Diboson}$	175.54	175.15	-0.01
$+\sigma_{Fakes}$	175.51	175.26	-0.04
$-\sigma_{Fakes}$	175.45	175.17	-0.02

according to:

$$w_{\pm} = w_0 \pm \frac{1}{5}(2i_{E_T} - 7)\sigma_w \tag{8.4.1}$$

where w_0 is the unshifted event weight, σ_w is the uncertainty on the weight, and i_{E_T} denotes the E_T bin. Therefore we will have linear E_T-dependent shifts in the event weights. Results of pseudoexperiments are shown in table 8.3 Such scaling results in 0.12 GeV/c^2 effect for the dilepton fit. In the Combined fit we take the half difference to the maximum shift and obtain 0.04 GeV/c^2 systematic.

Table 8.3: Results of pseudoexperiments when dilepton fake template is reweighted. All units are GeV/c^2. Comb refers to the combined fit and DIL refers to the dilepton fit.

Sample:	DIL	Comb	
Mean fitted:	M_{top}	M_{top}	Δ_{JES}
+ve linear shift	175.67	175.17	-0.03
-ve linear shift	175.42	175.17	-0.03

Drell-Yan events appear in the dilepton channel signal region due to mismeasurement of $\not{E}_T$. Accurately modelling the $\not{E}_T$ distribution is very difficult as it relies on correct handling of jet simulation as well as accurate detector model. We examine the effect of inaccurate modelling of $\not{E}_T$ on the dilepton Drell-Yan shape by reweighting $\not\!\!E_T$ distribution in the template using data-MC comparison. We select events in W+jets data satisfying the same dilepton cuts as regular dilepton selection and we require that:

- leptons form electron-positron pair or muon-anti-muon pair.
- dilepton mass falls within 15 GeV/c^2 from the Z boson mass
- two or more 15 GeV jets are present

Thus we select a Drell-Yan + 2 jets sample. We compare $\not\!\!E_T$ distribution in data to the one obtained from full Drell-Yan model. The comparison is shown in Figure 8.2. We obtain $\not\!\!E_T$ dependent weights from this comparison (table 8.4). The reweighted template is used in pseudoexperiments yielding mean fitted $\mathrm{M_{top}}$ of 175.74 GeV/c^2 for the dilepton channel and 175.21 GeV/c^2 for the combined fit. This results in 0.23 GeV/c^2 systematic effect on the dilepton fit and 0.03 GeV systematic on the combined fit.

Table 8.4: $\not\!\!E_T$ dependent weights for Drell-Yan shape shift.

$\not\!\!E_T$ bin	scale factor
25-30	1.11
30-35	0.78
35-40	0.81
40-45	1.10
45-50	0.96
$>$50	2.09

The evaluation of systematic uncertainty associated with the lepton + jets backgrounds is described in [18]. We study effects of uncertain background composition by running pseudoexperiments where background is drawn solely from one source: $Wb\bar{b}$, $Wc\bar{c}$, Wc, and W+light jets, single top, QCD or the diboson samples. We take

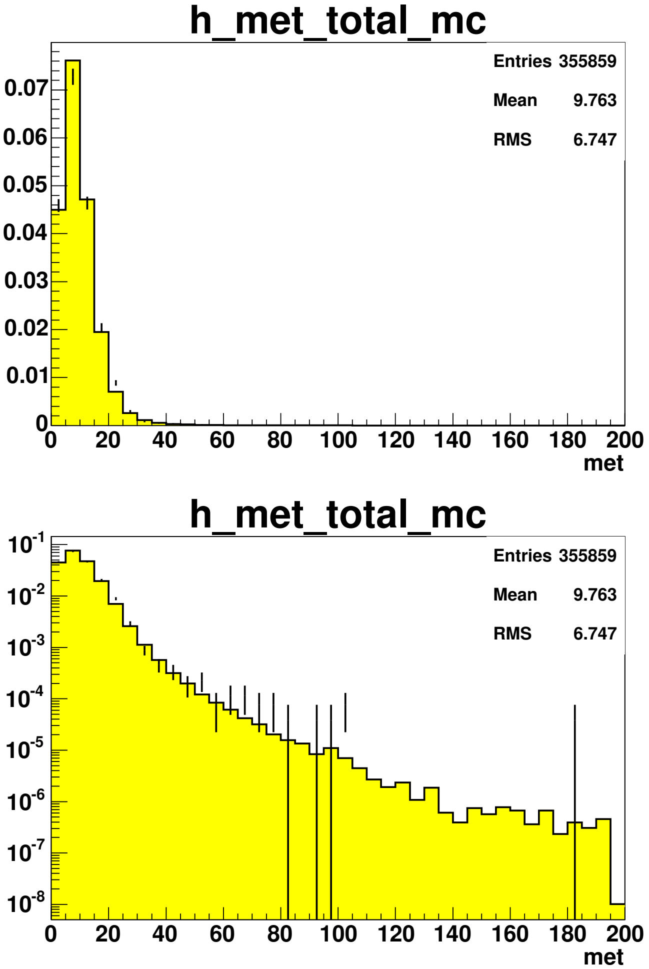


Figure 8.2: Data - Monte Carlo comparison of the $\not{E}_T$ distribution. Linear and log scale. Points are data. Events close to the Z peak are selected.

the largest shift from the nominal result as a systematic. We obtain 0.17 GeV/c^2 systematic for the lepton + jets fit and 0.11 GeV/c^2 for the combined fit. A variation of the momentum transfer scale in the $Wb\bar{b}$ simulation by a factor of two up and down results in a systematic effect of 0.09 GeV/c^2 on the lepton + jets measurement and 0.10 GeV/c^2 on the combined fit. Alternative model for QCD background is used to estimate the associated systematic. In this model a data sample where two of the electron quality cuts are inverted is used. Those 'anti-electrons' are meant to be objects where a quark or gluon jet is closely mimicking an electron. Use of this alternative sample induces very small differences in lepton + jets and combined fit results. The total uncertainty due to lepton + jets background shape is 0.19 GeV/c^2 in the lepton + jets fit and 0.17 GeV/c^2 in the combined fit.

8.5 Lepton Energy Scale

The uncertainty on muon and electron energy reconstruction is estimated to be 1% at the CDF detector. Since measurement of muon energy is based on the track curvature and measurement of electron energy is calorimeter-based we vary the electron and muon response independently. The scaling is performed prior to selection and $\not{E}_T$ calculation. In lepton + jets fit and combined fit up and down electron energy scale adjustment both result in lower than the nominal pseudoexperiment average. We take half difference to the nominal result obtaining 0.03 GeV/c^2 and 0.04 GeV/c^2 systematics for the lepton + jets and combined fits. The electron energy scale systematic in the dilepton fit is 0.23 GeV/c^2. The muon energy scale systematic is 0.09 GeV/c^2, 0.18 GeV/c^2 and 0.06 GeV/c^2 in the lepton + jets, dilepton and combined fits respectively.

8.6 Monte Carlo Event Generator

To evaluate a possible bias resulting from a choice of particular Monte Carlo event generator, we compare results of pseudoexperiments that use pseudodata from a large (4 million events) HERWIG generated sample with input top quark mass of 175 GeV/c^2 to the results of pseudoexperiments from run using the nominal PYTHIA sample and a with input values of $M_{\text{top}} = 175.0$ GeV/c^2. We find large differences 0.75 GeV/c^2 for the lepton + jets fit, 1.33 GeV/c^2 for the dilepton fit, and 0.67 GeV/c^2 for the combined fit. No single cause has been found for such large differences between the two generators in terms of the basic distributions of jet, lepton and neutrino momenta.

8.7 Initial and Final State Radiation

The amount of initial and final state radiation has been studied in the Drell-Yan system [24]. Those studies are used to constrain parameters in the PYTHIA Monte Carlo generator that control amount of gluon radiation. We run pseudoexperiments where the pseudodata has been drawn from samples in which the amount of gluon radiation from the initial and final state have been coherently varied within these constraints . Resulting half difference gives 0.15 GeV/c^2 uncertainty in the lepton + jets channel. Both up and down shifts result in a change in the same direction in the dilepton fit therefore half difference (0.20 GeV/c^2) to the nominal result is taken as a systematic. In the combined fit the statistical uncertainty on the half difference (0.13 GeV/c^2) is greater than the half difference itself (0.10 GeV/c^2) and is taken as a systematic.

8.8 Gluon-Gluon Fusion Fraction

The next to leading order prediction for the fraction of $t\bar{t}$ pair production coming from gluon fusion rather than quark annihilation is 15±5% [15]. However the leading order signal Monte Carlo samples used predict this fraction at 5.6%. We reweight the events in Monte Carlo sample generated at 175 GeV/c^2 so that events coming from the gluon fusion constitute 20% of pseudodata and take the difference in the mean fitted mass from pseudoexperiments as a systematic uncertainty. We obtain 0.16 GeV/c^2 for the combined fit, 0.19 GeV/c^2 for the lepton + jets fit and 0.17 GeV/c^2 for the dilepton fit.

8.9 Parton Distribution Functions

In hadron collisions only a fraction of momentum carried by a hadron will be available in the hard scattering process. This is because the hadron momentum is shared among partons within a hadron. The probability of finding a gluon or quark with fraction x of the hadron momentum is described by a probability density function (PDF). Gluons and quarks of different flavors have different PDFs. Since PDFs essentially control the amount of energy available in the collision inaccuracy in the PDFs will have a direct impact on the top quark measurement. The PYTHIA generator uses the PDF set published by CTEQ collaboration (the CTEQ5L version) [56]. To evaluate effect of another PDF set we reweight the nominal PYTHIA sample according this set. We find which partons produced the hard scattering process and what was the momentum transfer in the interaction. We then evaluate the probability of finding those partons given by the PDF from the set under the consideration and the probability given by the CTEQ5L. The ratio of those probabilities is the weight of this event when used in pseudoexperiments.

In this manner we compare the MRST72 [57] set to the default CTEQ5L set giving uncertainty of 0.10 GeV/c^2 for the lepton + jets fit and the dilepton fit and 0.08 GeV/c^2 for the combined fit. We also compare MRST72 and MRST75 PDF sets generated using different values of α_s. This yields 0.09 GeV/c^2 uncertainty for the lepton + jets fit, 0.39 GeV/c^2 for the dilepton fit and 0.11 GeV/c^2 for the combined fit. The CTEQ collaboration incorporated a treatment of correlated experimental uncertainties entering the PDF calculation and produced a set of 40 PDF sets where the 20 dominant components of uncertainties have been varied by $\pm 1\sigma$ [58]. We run pseudoexperiments using each of the 40 PDF sets and add in quadrature the resulting half shifts, resulting in 0.22 GeV/c^2 uncertainty for the lepton + jets fit, 0.36 GeV/c^2 for the dilepton fit and 0.22 GeV/c^2 for the combined fit. The total PDF uncertainty is the sum in quadrature of the results described above: 0.25 GeV/c^2 for the combined and lepton + jets fits and 0.54 GeV/c^2 for the dilepton fit

8.10 Summary of the Systematic Uncertainties

Results of pseudoexperiments where different signal pseudodata was used to model the systematic effects described above are summarized in Table 8.5. We add the systematic effects in quadrature and obtain 1.1 GeV/c^2 total systematic uncertainty for the combined fit and lepton + jets fit and 3.8 GeV/c^2 systematic uncertainty for the dilepton fit. The summary of systematic effects is shown in Table 8.6

Table 8.5: Summary of pseudoexperiment results used to evaluate systematic uncertainties. Mean $M_{\rm top}$ results have units of GeV/c^2 and mean $\Delta_{\rm JES}$ results have units of σ_c. Comb refers to the combined fit, LJ refers to the lepton + jets fit and DIL refers to the dilepton fit.

Sample:	LJ		DIL	Comb	
Mean fitted:	$M_{\rm top}$	$\Delta_{\rm JES}$	$M_{\rm top}$	$M_{\rm top}$	$\Delta_{\rm JES}$
Pythia $M_{\rm top} = 175.0$	175.18	-0.04	175.51	175.24	-0.04
Herwig $M_{\rm top} = 175.0$	174.43	0.32	176.84	174.57	0.33
gg fraction = 20%	174.99	0.02	175.68	175.08	0.03
ADO parameters	175.05	-0.01	175.44	175.10	0.00
SLD parameters	175.08	-0.01	175.55	175.13	0.00
semilep. fractions -1σ	175.21	-0.02	175.72	175.29	-0.02
semilep. fractions +1σ	175.10	-0.05	175.37	175.16	-0.04
b jet $E_T \times 0.99$	174.63	-0.04	174.69	174.65	-0.04
b jet $E_T \times 1.01$	175.72	-0.02	176.14	175.77	-0.01
e $E_T \times 0.99$	175.14	-0.02	175.22	175.15	-0.02
e $E_T \times 1.01$	175.12	-0.03	175.68	175.20	-0.02
μ $p_T \times 0.99$	175.28	-0.05	175.34	175.30	-0.04
μ $p_T \times 1.01$	175.10	-0.02	175.70	175.18	-0.01
more radiation	175.31	0.12	175.91	175.32	0.13
less radiation	175.00	0.02	175.90	175.13	0.03

Table 8.6: Summary of systematics uncertainties. All numbers have units of GeV/c^2. Comb refers to the combined fit, LJ refers to the lepton + jets fit and DIL refers to the dilepton fit

Systematic	LJ	DIL	Comb
Residual JES	0.7	3.5	0.7
Generator	0.8	1.3	0.7
PDFs	0.3	0.5	0.3
b jet energy	0.2	0.2	0.2
Background shape	0.2	0.3	0.2
gg-fusion fraction	0.2	0.2	0.2
initial and final state radiation	0.2	0.2	0.1
MC statistics	0.1	0.5	0.1
lepton energy scale	0.1	0.3	0.1
pileup	0.1	0.1	0.1
Total systematic uncertainty	1.1	3.8	1.1

CHAPTER 9

RESULTS

9.1 Candidate Event Counts

In the data we observe 144 candidate dilepton events and 332 candidate lepton + jets events passing all cuts. All dilepton events are reconstructed by Neutrino Weighting Algorithm and pass the boundary cut. In the lepton + jets channel 104 fail the χ^2 cut or the boundary cut. Tables 9.1 and 9.2 summarize the event counts.

Table 9.1: Observed number of dilepton candidate events in data before and after the boundary cuts.

Sample	before boundary cut	after boundary cut
non-tagged	83	83
tagged	61	61

Table 9.2: Observed number of lepton + jets candidate events candidate events in data before and after the χ^2 and boundary cuts.

Sample	pre-χ^2 cut	pre boundary cut	after all cuts
1-tag	284	237	233
2-tag	152	101	99

9.2 Fit Results

We apply the likelihood fit to the data and measure:

$$\begin{aligned} \mathrm{M_{top}} &= 171.9 \pm 1.7 \text{ (stat.} + \Delta_{\mathrm{JES}} \text{ only)} \ \ \mathrm{GeV}/c^2 \text{ (combined)} \\ \Delta_{\mathrm{JES}} &= -0.12 \pm 0.34\sigma_c \end{aligned}$$

$$\mathrm{M_{top}} = 171.6^{+3.5}_{-3.3} \text{ (stat. only)} \ \ \mathrm{GeV}/c^2 \text{ (dilepton)}$$

$$\begin{aligned} \mathrm{M_{top}} &= 171.8 \pm 1.8 \text{ (stat.} + \Delta_{\mathrm{JES}} \text{ only)} \ \ \mathrm{GeV}/c^2 \text{ (lepton + jets)} \\ \Delta_{\mathrm{JES}} &= -0.09 \pm 0.36\sigma_c \end{aligned}$$

The results above are not corrected for pull width and luminosity effects. The results are surprisingly consistent between the two channels. Note also that the combined fit prefers top quark mass value higher then the values preferred by both the dilepton and lepton + jets fits. This is the effect of *in situ* JES calibration being applied from the lepton + jets channel onto the dilepton channel data. Shift of $-0.1\ \sigma_{\mathrm{c}}$ in Δ_{JES} corresponds to approximately $+0.3$ GeV/c^2 shift in the top mass, therefore after this calibration the dilepton-only fit actually prefers top quark mass marginally higher then the lepton+jets fit.

The negative log-likelihood contours for the combined fit are shown in Fig. 9.1. In this diagram the values of the top quark mass and Δ_{JES} that minimize the likelihood are indicated by 'x'. Along the oval contours the value of negative log likelihood is constant. All other parameters of the fit (that is signal and background expectations n_s and n_b for all the subsamples) are minimized at a grid of (M_{top}, Δ_{JES}) and contours where the value of negative log likelihood increases by 0.5, 2.0 and 4.5 are found. A 1 σ uncertainty can be obtained by drawing two vertical tangents to the inner oval contour. The distances between those two lines and the (M_{top}, Δ_{JES}) point minimizing the likelihood, are the negative and positive uncertainties due to statistics of the data sample and JES. In the same way we can find the 2 and 3 σ uncertainties. Note that interpreting the rise in log-likelihood by 0.5 as the uncertainty assumes Gaussian likelihood form. Non-Gaussian likelihood causes the pull width in pseudoexperiments to be larger than 1 and has to be corrected for (*cf.* Sec. 7.1).

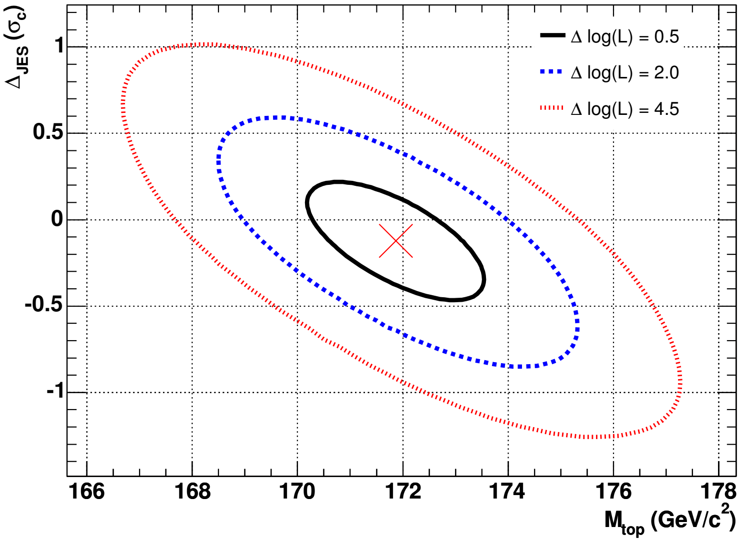


Figure 9.1: Negative log-likelihood contours for the combined fit.

The negative log-likelihood contours for the lepton + jets fit are shown in Fig. 9.2. The interpretation of the contours is the same as for the combined fit.

For the dilepton fit where the *in situ* JES calibration is not applied we show a

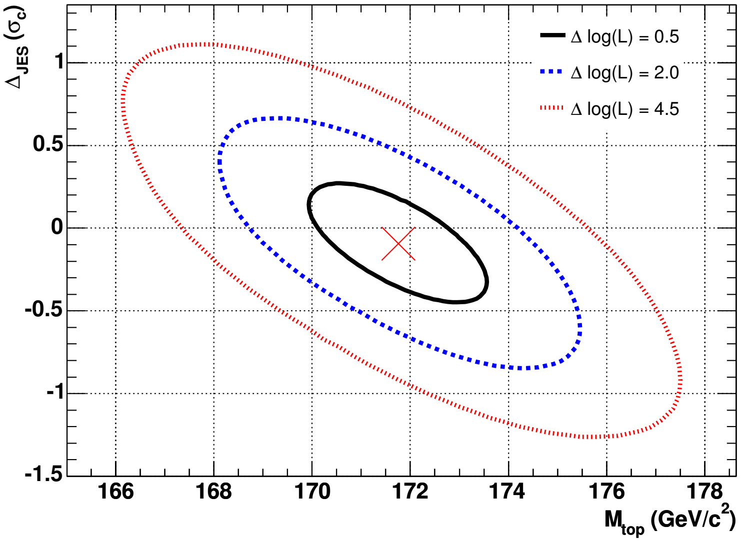


Figure 9.2: Negative log-likelihood contours for the lepton + jets fit

negative log likelihood profile (Fig. 9.3). A scan in the top quark mass is performed. At each point we minimize the log likelihood with respect to the signal and background expectations and plot the resulting offset of the log-likelihood from the minimum. The points where the offset from the minimum reaches 0.5 are interpreted as positive and negative uncertainties.

The distributions of observables in data in the dilepton channel are shown in Fig. 9.4 and from the lepton + jets channel in Fig. 9.5. The normalization of the templates in the figures is set to the fitted signal and background expectation values obtained in the combined fit.

To measure the probability of obtaining statistical uncertainty returned by the fits on the data, we conduct pseudoexperiments where the mean total number of events is set to the number observed in the data. Average number of background events is set to the *a priori* background estimate in those pseudoexperiments as usual. We use the Monte Carlo sample generated at $\mathrm{M_{top}} = 172.0$ GeV/c^2 and $\Delta_{\mathrm{JES}} = 0.0\ \sigma_{\mathrm{c}}$ to draw

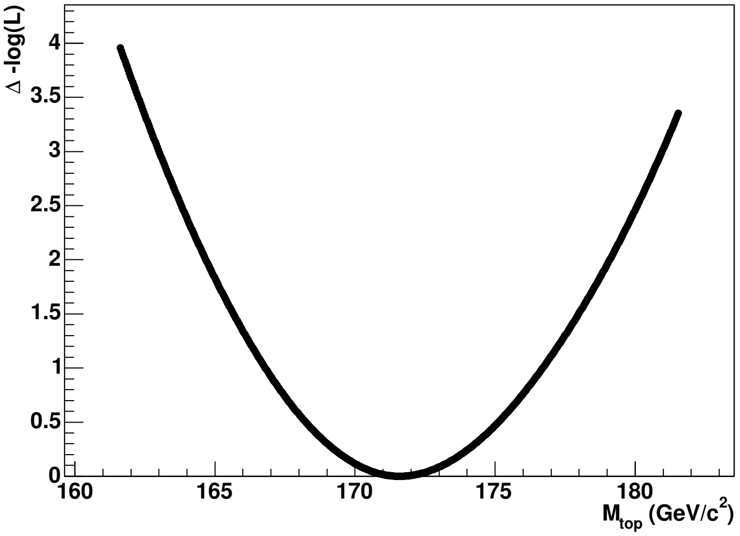
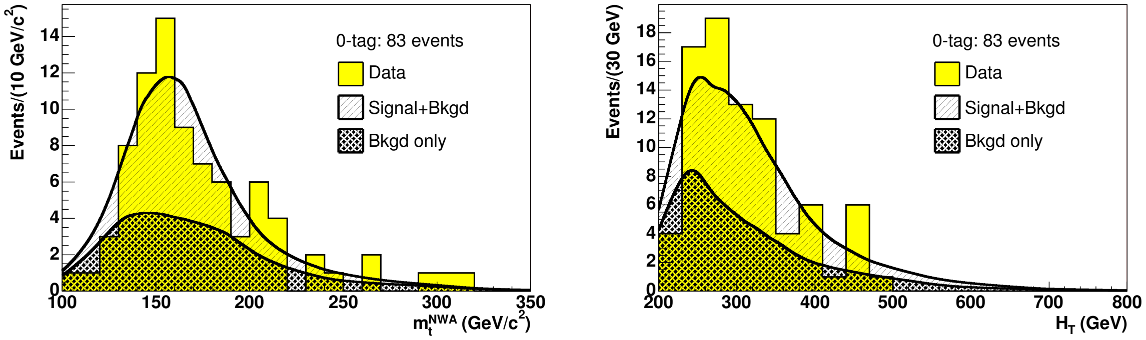


Figure 9.3: Negative log-likelihood profile for the dilepton fit

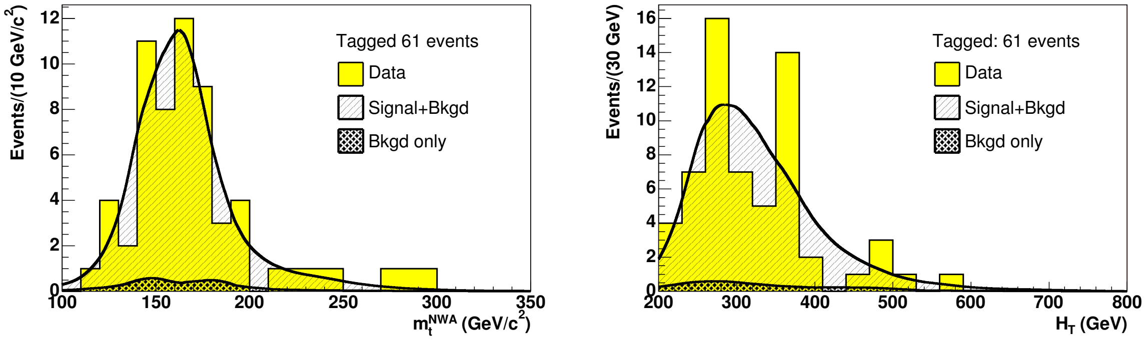
the signal pseudodata. Symmetrized error distributions for all three fits are shown in Fig. 9.6. Both the lepton + jets fit and dilepton fit turn out to be "lucky" with chance of resulting in smaller error of 21% and 14% respectively. Central values returned by separate channel fits are very close causing the uncertainty to be significantly lower than expected with 10% pseudoexperiments obtaining a smaller uncertainty.

9.3 Results After Corrections

As indicated in Section 7.2, the fitted $\mathrm{M_{top}}$ in the dilepton channel and fitted Δ_{JES} values from the lepton + jets and combined fits need to be corrected for instantaneous luminosity bias. The quoted statistical uncertainties also need to be increased.

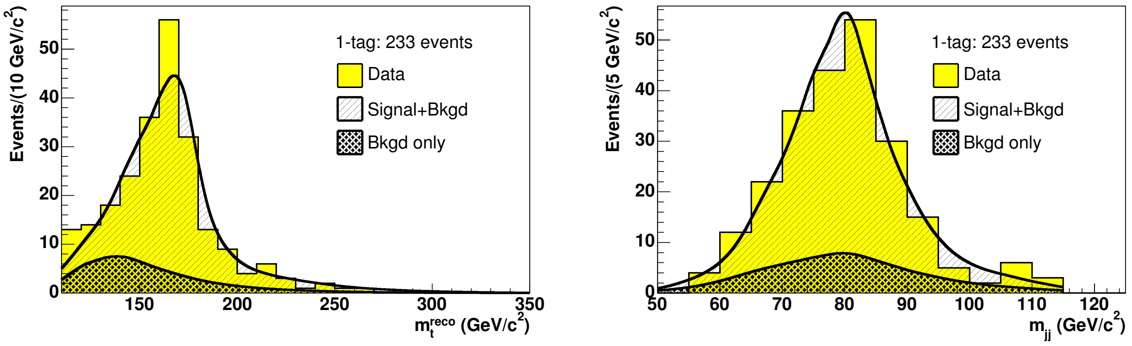


(a) 0-tag m_t^{NWA} (b) 0-tag H_T



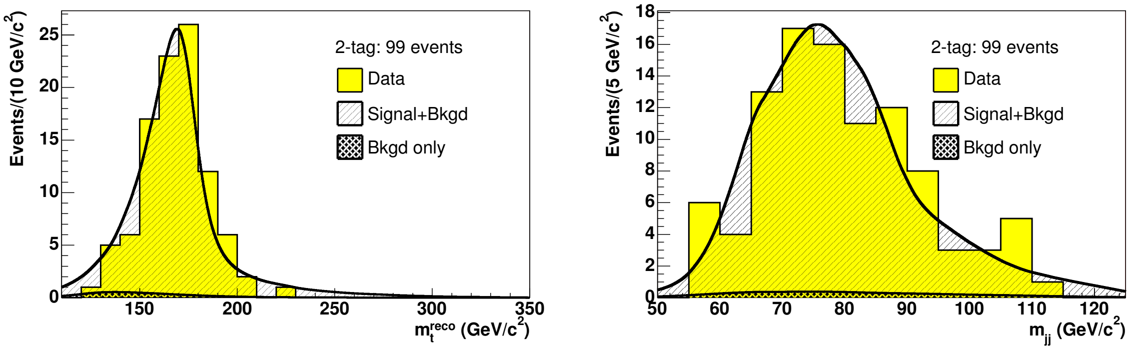
(c) Tagged m_t^{NWA} (b) Tagged H_T

Figure 9.4: One-dimensional dilepton data with density estimates overlaid using $\mathrm{M_{top}} = 172.0$ GeV/c^2, $\Delta_{\mathrm{JES}} = 0.0$, and full background model. The expected numbers of events are set to the values from the constrained fit. Shown are the 0-tag m_t^{NWA} (a) and H_T (b) distributions, and the tagged m_t^{NWA} (c) and H_T (d) distributions.



(a) 1-tag m_t^{reco}

(b) 1-tag m_{jj}



(c) 2-tag m_t^{reco}

(d) 2-tag m_{jj}

Figure 9.5: One-dimensional lepton + jets data with density estimates overlaid using $\mathrm{M_{top}} = 172.0$ GeV/c^2, $\Delta_{\text{JES}} = 0.0$, and a full background model. The expected numbers of events are set to the values from the constrained fit. Shown are the 1-tag m_t^{reco} (a) and m_{jj} (b) distributions, and the 2-tag m_t^{reco} (c) and m_{jj} (d) distributions.

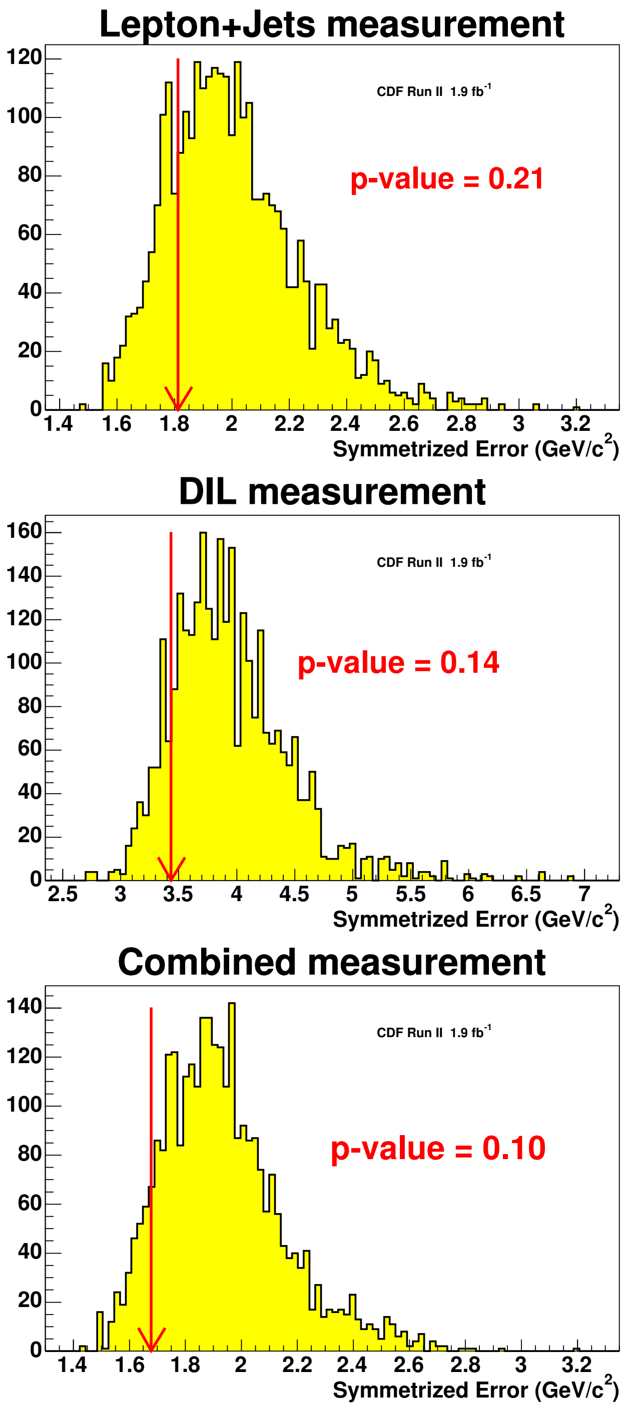


Figure 9.6: Expected errors and probability to get values equal to or smaller than the measured errors for lepton + jets (top), dilepton (middle) and combined (bottom) fits. Uncertainty obtained in data is shown by the red arrow

After all corrections we obtain the following results:

$$\begin{aligned}
\mathrm{M_{top}} &= 171.9 \pm 1.7\ (\text{stat.} + \text{JES}) \pm 1.1\ (\text{other syst.})\ \mathrm{GeV}/c^2 \\
&= 171.9 \pm 2.0\ \mathrm{GeV}/c^2\ (\text{combined}) \\
\mathrm{M_{top}} &= 171.2\ ^{+3.6}_{-3.4}\ (\text{stat.}) \pm 3.8\ (\text{syst.})\ \mathrm{GeV}/c^2 \\
&= 171.2\ ^{+5.3}_{-5.1}\ \mathrm{GeV}/c^2\ (\text{dilepton}) \\
\mathrm{M_{top}} &= 171.8 \pm 1.9\ (\text{stat.} + \text{JES}) \pm 1.1\ (\text{other syst.})\ \mathrm{GeV}/c^2 \\
&= 171.8 \pm 2.2\ \mathrm{GeV}/c^2\ (\text{lepton} + \text{jets})
\end{aligned} \tag{9.3.1}$$

The Jet Energy Scale measurements are:

$$\begin{aligned}
\Delta_{\mathrm{JES}} &= -0.17 \pm 0.35\ (\text{stat.} + \mathrm{M_{top}}\ \text{only})\ \sigma_{\mathrm{c}} \quad (\text{combined}) \\
\Delta_{\mathrm{JES}} &= -0.13 \pm 0.37\ (\text{stat.} + \mathrm{M_{top}}\ \text{only})\ \sigma_{\mathrm{c}} \quad (\text{lepton} + \text{jets})
\end{aligned} \tag{9.3.2}$$

9.4 Cross Checks

We perform fit for $\mathrm{M_{top}}$ while removing the jet energy constraint prior or the background expectation priors. We also split the data into subsamples based on tag count, lepton type and data taking period. The results are summarized in Tables 9.3, 9.4, 9.5 and 9.6. The results quoted are not corrected for pull width or instantaneous luminosity effects.

Both lepton + jets and combined fits are robust against releasing the JES prior. The fitted $\mathrm{M_{top}}$ does not change, however the uncertainty on $\mathrm{M_{top}}$ is increased by about 0.1 GeV/c^2 in the lepton + jets fit. The fitted Δ_{JES} value is as expected slightly lower in the fit with no JES prior but the change is very small.

Removing the background priors from the likelihood has essentially no effect on the fitted top quark mass across the two channels and in the combined fit. There are small variations in the fitted signal and background expectations from the values returned when the priors are applied. Notably the 2-tag lepton + jets sample prefers no background. The 0-tag dilepton sample prefers about 30% less background than given by the *a* priori estimate, but that is within the uncertainty returned by the fit. Sometimes the indicated error on the background expectation is "-". This signifies that n_b is consistent with 0 and the error estimation routine fails.

Fits using the data only from subsamples separated by multiplicity of b tagged jets are consistent among one another and with the combined result. Separation of data by lepton or dilepton type gives results that are consistent among one another. Dividing the data based on data taking period gives a difference of about 9 GeV/c^2 in the fitted $\mathrm{M}_{\mathrm{top}}$ in the dilepton channel. This difference is actually only at the level of 1.3 σ so is not a cause for concern. Similarly the difference seen in the Δ_{JES} fit in the first and second half of the data in the lepton + jets and combined results is not statistically significant.

Table 9.3: Cross-checks on the data. Nominal fit and fits with JES or background constraints removed. Comb refers to the combined fit, LJ refers to the lepton + jets fit and DIL refers to the dilepton fit.

Fit	Comb	M_{top} (GeV/c^2)		Δ_{JES} (σ_c)	
		LJ: 1tag n_s	2tag n_s	1tag n_b	2tag n_b
		DIL: 0tag n_s	tagged n_s	0tag n_b	tagged n_b
	LJ-only	M_{top} (GeV/c^2)		Δ_{JES} (σ_c)	
		1tag n_s	2tag n_s	1tag n_b	2tag n_b
	DIL-only	M_{top} (GeV/c^2)			
		0tag n_s	tagged n_s	0tag n_b	tagged n_b
Nominal	Comb	171.9 ± 1.7		-0.12 ± 0.34	
		$184.1\ ^{+17.7}_{-17.1}$	$96.4\ ^{+10.4}_{-9.7}$	$47.4\ ^{+10.2}_{-10.0}$	3.4 ± 1.9
		$43.8\ ^{+10.4}_{-9.8}$	$56.9\ ^{+8.2}_{-7.5}$	$41.5\ ^{+6.5}_{-6.4}$	3.9 ± 1.0
	LJ	171.8 ± 1.8		-0.09 ± 0.36	
		$184.0\ ^{+17.8}_{-17.1}$	$96.4\ ^{+10.4}_{-9.7}$	$47.5\ ^{+10.2}_{-10.1}$	3.4 ± 1.9
	DIL	$171.6\ ^{+3.5}_{-3.3}$		-	
		$43.7\ ^{+10.4}_{-9.8}$	$56.9\ ^{+8.2}_{-7.5}$	$41.6\ ^{+6.5}_{-6.4}$	3.9 ± 1.0
No JES prior	Comb	171.9 ± 1.7		$-0.14\ ^{+0.36}_{-0.37}$	
		$184.1\ ^{+17.7}_{-17.1}$	$96.4\ ^{+10.4}_{-9.7}$	$47.4\ ^{+10.2}_{-10.0}$	3.4 ± 1.9
		$43.8\ ^{+10.4}_{-9.8}$	$56.9\ ^{+8.2}_{-7.5}$	$41.5\ ^{+6.5}_{-6.4}$	3.9 ± 1.0
	LJ	171.8 ± 1.9		$-0.11\ ^{+0.39}_{-0.38}$	
		$184.1\ ^{+17.8}_{-17.1}$	$96.4\ ^{+10.4}_{-9.7}$	$47.5\ ^{+10.2}_{-10.1}$	3.4 ± 1.9
No background prior	Comb	171.9 ± 1.7		$-0.11\ ^{+0.35}_{-0.34}$	
		$176.1\ ^{+21.5}_{-20.7}$	$99.0\ ^{+10.3}_{-9.6}$	$57.0\ ^{+18.6}_{-17.5}$	$0.0\ ^{+2.2}_{-}$
		$54.8\ ^{+15.8}_{-15.2}$	54.3 ± 12.2	$28.2\ ^{+15.2}_{-14.0}$	$6.7\ ^{+11.1}_{-}$
	LJ	171.8 ± 1.8		-0.06 ± 0.36	
		$175.7\ ^{+21.6}_{-20.7}$	$99.0\ ^{+10.3}_{-9.6}$	$57.3\ ^{+18.6}_{-17.6}$	$0.0\ ^{+2.2}_{-}$
	DIL	171.5 ± 3.4		-	
		$54.6\ ^{+15.8}_{-15.2}$	54.2 ± 12.3	$28.4\ ^{+15.3}_{-14.0}$	$6.8\ ^{+11.2}_{-}$

Table 9.4: Cross-checks on the data. Fits on dilepton and lepton + jets subsamples separated by tagging multiplicity. LJ refers to the lepton + jets fit and DIL refers to the dilepton fit. Background prior is applied

Fit	Comb	M_{top} (GeV/c^2)		Δ_{JES} (σ_c)	
		LJ: 1tag n_s	2tag n_s	1tag n_b	2tag n_b
		DIL: 0tag n_s	tagged n_s	0tag n_b	tagged n_b
	LJ-only	M_{top} (GeV/c^2)		Δ_{JES} (σ_c)	
		1tag n_s	2tag n_s	1tag n_b	2tag n_b
	DIL-only	M_{top} (GeV/c^2)			
		0tag n_s	tagged n_s	0tag n_b	tagged n_b
0-tag	DIL	$170.1^{+6.4}_{-7.6}$		-	
		$43.7^{+10.5}_{-9.8}$	-	$41.6^{+6.5}_{-6.4}$	-
Tagged	DIL	$172.2^{+4.4}_{-4.0}$		-	
		-	$57.0^{+8.2}_{-7.5}$	-	3.9 ± 1.0
1-tag	LJ	$169.1^{+3.1}_{-2.6}$		$-0.17^{+0.48}_{-0.57}$	
		$187.6^{+17.9}_{-17.2}$	-	44.8 ± 10.3	-
2-tag	LJ	$173.6^{+2.6}_{-2.3}$		$0.20^{+0.47}_{-0.50}$	
		-	$96.3^{+10.4}_{-9.7}$	-	3.5 ± 1.9

Table 9.5: Cross-checks on the data. Fits on dilepton and lepton + jets subsamples separated by type of leptons. LJ refers to the lepton + jets fit and DIL refers to the dilepton fit. Background prior is not applied

Fit	Comb	$M_{\rm top}$ (GeV/c^2)		$\Delta_{\rm JES}$ (σ_c)	
		LJ: 1tag n_s	2tag n_s	1tag n_b	2tag n_b
		DIL: 0tag n_s	tagged n_s	0tag n_b	tagged n_b
	LJ-only	$M_{\rm top}$ (GeV/c^2)		$\Delta_{\rm JES}$ (σ_c)	
		1tag n_s	2tag n_s	1tag n_b	2tag n_b
	DIL-only	$M_{\rm top}$ (GeV/c^2)			
		0tag n_s	tagged n_s	0tag n_b	tagged n_b
ee	DIL	169.0 ± 8.0		-	
		$5.6\ ^{+7.6}_{-}$	$9.0\ ^{+3.3}_{-2.7}$	$15.4\ ^{+8.1}_{-7.3}$	$0.0\ ^{+1.3}_{-}$
$e\mu$	DIL	$173.6\ ^{+5.2}_{-4.0}$		-	
		$33.3\ ^{+11.1}_{-10.7}$	$27.3\ ^{+7.3}_{-9.8}$	$9.7\ ^{+10.5}_{-8.9}$	$1.7\ ^{+9.4}_{-}$
$\mu\mu$	DIL	$167.9\ ^{+7.5}_{-6.8}$		-	
		$16.2\ ^{+7.1}_{-7.8}$	$14.9\ ^{+7.2}_{-7.1}$	$2.8\ ^{+7.9}_{-}$	$8.1\ ^{+7.6}_{-5.7}$
e	LJ	172.2 ± 2.7		-0.09 ± 0.51	
		$95.1\ ^{+16.0}_{-15.2}$	$54.0\ ^{+7.7}_{-7.0}$	$35.9\ ^{+14.0}_{-13.0}$	$0.0\ ^{+3.0}_{-}$
μ	LJ	$171.3\ ^{+2.4}_{-2.3}$		$-0.04\ ^{+0.46}_{-0.47}$	
		$81.0\ ^{+14.7}_{-13.9}$	$45.0\ ^{+7.0}_{-6.4}$	$21.0\ ^{+12.5}_{-11.4}$	$0.0\ ^{+1.6}_{-}$

Table 9.6: Cross-checks on the data. Combined, lepton + jets and dilepton fits separated by data taking period. Comb refers to the combined fit, LJ refers to the lepton + jets fit and DIL refers to the dilepton fit. Background prior is not applied

Fit	Comb	M_{top} (GeV/c^2)		Δ_{JES} (σ_c)	
		LJ: 1tag n_s	2tag n_s	1tag n_b	2tag n_b
		DIL: 0tag n_s	tagged n_s	0tag n_b	tagged n_b
	LJ-only	M_{top} (GeV/c^2)		Δ_{JES} (σ_c)	
		1tag n_s	2tag n_s	1tag n_b	2tag n_b
	DIL-only	M_{top} (GeV/c^2)			
		0tag n_s	tagged n_s	0tag n_b	tagged n_b
First 1 fb^{-1}	Comb	$171.7\,^{+2.3}_{-2.4}$		$0.45\,^{+0.55}_{-0.50}$	
		$90.8\,^{+15.6}_{-14.8}$	$48.0\,^{+7.3}_{-6.6}$	$41.2\,^{+13.9}_{-13.0}$	$0.0\,^{+1.7}_{-}$
		$32.8\,^{+11.4}_{-10.9}$	$17.6\,^{+8.1}_{-7.8}$	$9.2\,^{+10.8}_{-}$	$8.4\,^{+8.1}_{-6.4}$
	LJ	$172.2\,^{+2.5}_{-2.4}$		$0.59\,^{+0.52}_{-0.55}$	
		$89.7\,^{+15.5}_{-14.7}$	$48.0\,^{+7.3}_{-6.6}$	$42.3\,^{+13.9}_{-13.0}$	$0.0\,^{+1.7}_{-}$
	DIL	166.1 ± 5.0		-	
		$34.5\,^{+10.6}_{-10.2}$	$16.9\,^{+7.8}_{-7.7}$	$7.5\,^{+9.8}_{-}$	$9.1\,^{+8.1}_{-6.3}$
Last 0.9 fb^{-1}	Comb	171.7 ± 2.7		$-0.70\,^{+0.53}_{-0.59}$	
		$89.1\,^{+14.9}_{-14.2}$	$51.0\,^{+7.5}_{-6.8}$	$11.9\,^{+12.2}_{-11.1}$	$0.0\,^{+2.7}_{-}$
		$20.3\,^{+11.2}_{-10.4}$	$35.0\,^{+6.3}_{-6.9}$	$20.7\,^{+11.4}_{-10.3}$	$0.0\,^{+5.5}_{-}$
	LJ	$170.2\,^{+3.1}_{-3.0}$		$-0.61\,^{+0.54}_{-0.64}$	
		$90.5\,^{+15.1}_{-14.3}$	$51.0\,^{+7.5}_{-6.8}$	$10.5\,^{+12.2}_{-}$	$0.0\,^{+2.6}_{-}$
	DIL	$175.2\,^{+5.3}_{-4.7}$		-	
		$22.1\,^{+11.1}_{-10.5}$	$35.0\,^{+6.3}_{-6.4}$	$18.9\,^{+11.3}_{-10.0}$	$0.0\,^{+4.9}_{-}$

CHAPTER 10

CONCLUSIONS

We present the first top quark mass measurement using data from two decay channels simultaneously. In the 1.9 fb^{-1} of data from $t\bar{t}$ decays into dilepton and lepton + jets channels we measure:

$$\begin{aligned} \mathrm{M_{top}} &= 171.9 \pm 1.7 \text{ (stat. + JES)} \pm 1.1 \text{ (other syst.) GeV}/c^2 \\ &= 171.9 \pm 2.0 \text{ GeV}/c^2 \text{ (combined)} \end{aligned} \tag{10.0.1}$$

As a consistency check we perform measurements using the data from dilepton channel and lepton + jets channel separately finding an agreement between the results:

$$\begin{aligned} \mathrm{M_{top}} &= 171.2\ ^{+3.6}_{-3.4} \text{ (stat.)} \pm 3.8 \text{ (syst.) GeV}/c^2 \\ &= 171.2\ ^{+5.3}_{-5.1} \text{ GeV}/c^2 \text{ (dilepton only)} \\ \mathrm{M_{top}} &= 171.8 \pm 1.9 \text{ (stat. + JES)} \pm 1.1 \text{ (other syst.) GeV}/c^2 \\ &= 171.8 \pm 2.2 \text{ GeV}/c^2 \text{ (lepton + jets only)} \end{aligned} \tag{10.0.2}$$

The analysis includes full treatment of correlations in systematics between the two channels and does not assume Gaussian likelihoods or symmetric errors in the channels being combined. The jet energy scale calibration is extracted form hadronic decay of the W boson in the lepton + jets channel and is applied to data from both channels.

The method presented here is expected to achieve statistical precision of about 1.2 GeV/c^2 with full Tevatron dataset with no method improvements. The analysis can and should be extended by an addition of the all-hadronic decay channel providing a single robust treatment of the systematic effects in all channels. We expect a decrease of approximately 5% in overall uncertainty resulting from addition of the all-hadronic channel, however future improvements are hard to predict as systematic uncertainties may be reduced.

This measurement is an important contribution to the puzzle of creation of mass. Shown in Fig. 10.1 are the same W boson mass and top quark mass measurements as in Fig. 1.2, but now also the top quark mass result of this analysis has been added. As apparent from the figure, this measurement offers an improvement in overall uncertainty on the top quark mass and seems to indicate slightly higher preferred Higgs boson mass than either of the two previously published separate dilepton and lepton + jets measurements. The measurement presented here is an important milestone in achieving more precise top quark mass results. These future precision measurements of the top quark mass as well as the measurements of the W boson mass will allow us to test if the current understanding of the electroweak symmetry breaking mechanism is correct. The Tevatron will continue to run through the Spring of 2009 providing a total data set about three times larger than used for this analysis. With such large data set a measurement of the top quark mass with total uncertainty of about 1 GeV/c^2 (using both CDF and DØ data) and a measurement of the W boson mass with uncertainty lower than 25 MeV/c^2 will be possible. The Higgs boson may also be discovered using the Tevatron data. If the Higgs is not found at the Tevatron, it will be found at the LHC - provided that it does exist. Given those measurements we will then be finally able to answer the question "Why do things have mass?", or we may find out that our current understanding is incorrect.

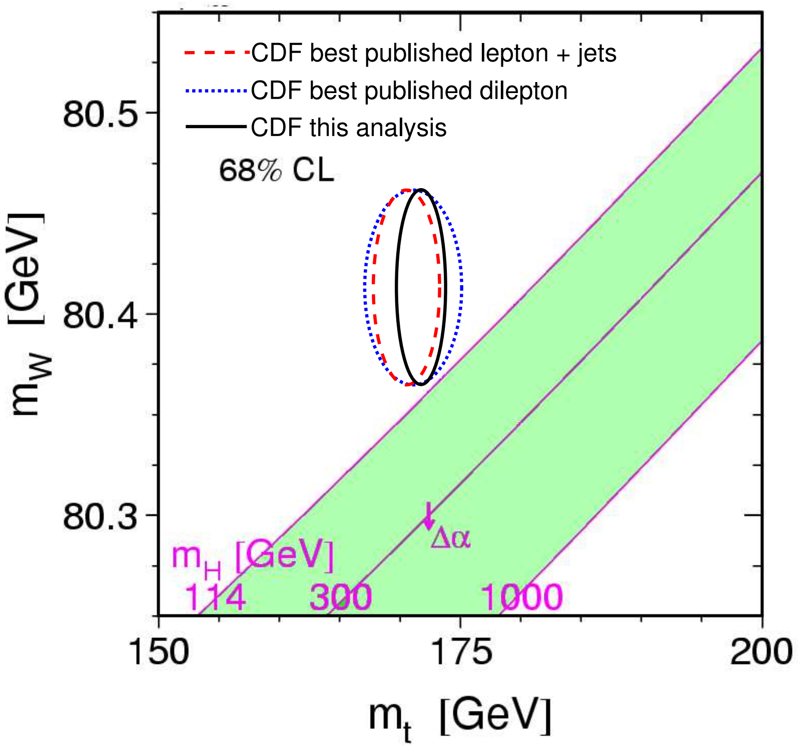


Figure 10.1: Dependence of the Higgs boson mass on the top quark mass and the W boson mass. Adapted from [3]. Ellipses indicate CDF II W boson mass measurement [4] and most precise CDF II (published or submitted) dilepton [5] (dotted) and lepton + jets [6] (dashed) top quark mass results. Solid ellipse indicates the top quark mass result presented in this analysis and the W boson mass result used in the other two ellipses. One dimensional 68% confidence levels are indicated by the ellipses.

REFERENCES

[1] F. Abe *et al.* Evidence for top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. D*, 50(5):2966, Sep 1994.

[2] W.-M. Yao *et al.* Review of particle physics. *J. Phys. G*, 33:1, 2006. and 2007 partial update for the 2008 edition.

[3] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, The LEP Electroweak Working Group. Precision electroweak measurements and constraints on the Standard Model. *hep-ex/arXiv:0712.0929.* Updated July 2008.

[4] T. Aaltonen *et al.* First measurement of the W boson mass in run II of the tevatron. *Phys. Rev. Lett.*, 99(15):151801, 2007.

[5] T. Aaltonen *et al.* Measurement of the top quark mass with dilepton events selected using neuroevolution at CDF. *arXiv:hep-ex/0807.4652*, 2008. Submitted to Phys. Rev. Lett.

[6] T. Aaltonen *et al.* Precise measurement of the top quark mass in the lepton + jets topology at CDF II. *Phys. Rev. Lett.*, 99(18):182002, 2007.

[7] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, The LEP Working Group for Higgs Boson Searches. Search for the Standard Model Higgs boson at LEP. *Phys. Lett. B*, 565:61, 2003.

[8] The Tevatron Electroweak Working Group. A combination of CDF and DØ results on the mass of the top quark. *hep-ex/arXiv:0808.1089*, 2008.

[9] S. Heinemeyer *et al.* Precise prediction for M_W in the MSSM. *hep-ph/arXiv:0604147v2.* Updated March 2008.

[10] S. Heinemeyer *et al.* Electroweak precision observables in the Minimal Supersymmetric Standard Model. *Physics Reports*, 425:265, 2006.

[11] A. Djouadi *et al.* Supersymmetric contributions to electroweak precision observables: QCD corrections. *Phys. Rev. Lett.*, 78(19):3626, May 1997.

[12] A. Djouadi *et al.* Leading QCD corrections to scalar quark contributions to electroweak precision observables. *Phys. Rev. D*, 57(7):4179, Apr 1998.

[13] S. Heinemeyer *et al.* Leading electroweak two-loop corrections to precision observables in the mssm. *J. High Energy Phys.*, 2002(10):072, 2002.

[14] J. Haestier *et al.* Electroweak precision observables: Two-loop yukawa corrections of supersymmetric particles. *arXiv:hep-ph/0508139v1.*

[15] M. Cacciari *et al.* The $t\bar{t}$ cross-section at 1.8 and 1.96 TeV: a study of the systematics due to parton densities and scale dependence. *J. High Energy Phys.*, 04:068, 2004.

[16] B. W. Harris *et al.* The fully differential single top quark cross section in next-to-leading order QCD. *Phys. Rev. D*, 66:054024, 2002.

[17] K. Chetyrkin *et al.* Second order QCD corrections to $\Gamma(t \to Wb)$. *Phys. Rev. D*, 60:114015, 1999.

[18] J. Adelman. *Measurement of the Top Quark Mass at CDF Using the Template Method in the Lepton + Jets Channel.* PhD thesis, The University of Chicago.

[19] L. Lyons *et al.* How to combine correlated estimates of a single physical quantity. *Nucl. Instrum. Meth. A*, 270:110, 1988.

[20] adopted from www.fnal.gov.

[21] P. Derwent *et al.* Debuncher cooling performance. Technical report, FNAL, FERMILAB-CONF-05-458-AD.

[22] Jim Morgan. The Antiproton Source Rookie Book.

[23] D. Acosta *et al.* Measurement of the J/ψ meson and b-hadron production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1960$ GeV. *Phys. Rev. D*, 71:032001, 2005.

[24] A. Abulencia *et al.* Top quark mass measurement using the template method in the lepton + jets channel at CDF II. *Phys. Rev. D*, 73:032003, 2006.

[25] A. Sill *et al.* CDF Run II silicon tracking projects. *Nucl. Instrum. Meth. A*, 447:1, 2000.

[26] The CDF II Collaboration. The CDF II detector, technical design report. *FERMILAB-Pub-96/390-E*, 1996.

[27] T. Affolder *et al.* CDF central outer tracker. *Nucl. Instrum. Meth. A*, 526:249, 2004.

[28] A. Abulencia *et al.* Measurements of inclusive W and Z cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *J. Phys. G*, 34(12):2457.

[29] L. Balka *et al.* The CDF central electromagnetic calorimeter. *Nucl. Instrum. Meth. A*, 267:272, 1988.

[30] S. Bertolucci *et al.* The CDF central and endwall hadron calorimeter. *Nucl. Instrum. Meth. A*, 267:301, 1988.

[31] M. Albrow *et al.* The CDF plug upgrade electromagnetic calorimeter: test beam results. *Nucl. Instrum. Methods A*, 480:524, 2002.

[32] G. Apollinari *et al.* Shower maximum detector for the CDF plug upgrade calorimeter. *Nucl. Instrum. Methods A*, 412:515, 1998.

[33] D. Acosta et al. The performance of the CDF luminosity monitor. *Nucl. Instrum. Methods*, A494:57.

[34] T. Sjostrand *et al.* Pythia 6.4 physics and manual. *J. High Energy Phys.*, 05:026, 2006.

[35] G. Corcella *et al.* Herwig 6.5. *J. High Energy Phys.*, 01:010, 2001.

[36] F. Maltoni *et al.* Madevent: Automatic event generation with MadGraph. *J. High Energy Phys.*, 0302:027, 2003.

[37] M. L. Mangano *et al.* ALPGEN, a generator for hard multiparton processes in hadronic collisions. *J. High Energy Phys.*, 07:001, 2003.

[38] T. Sjostrand *et al.* High-energy-physics event generation with pythia 6.1. *Comput. Phys. Commun.*, 135:238, 2001.

[39] M. L. Mangano *et al.* Matching matrix elements and shower evolution for top-quark production in hadronic collisions. *J. High Energy Phys.*, 01:013, 2007.

[40] Darin E. Acosta *et al.* Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with secondary vertex $b-$tagging. *Phys. Rev. D*, 71:052003, 2005.

[41] F. Abe *et al.* The topology of three jet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev.*, D45:1448, 1992.

[42] A. Bhatti *et al.* Determination of the jet energy scale at the Collider Detector at Fermilab. *Nucl. Instr. Meth. A*, 566:375, 2006.

[43] The CDF Collaboration. Top dilepton cross section in 2.0 fb^{-1} using the DIL selection. *CDF Public Note CDF9291*, 2008.

[44] B. Abbott *et al.* Measurement of the top quark mass in the dilepton channel. *Phys. Rev. D*, 60(5):052001, 1999.

[45] A. Abulencia *et al.* Top quark mass using template methods on dilepton events in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 73:112006, 2006.

[46] S. Sabik. *Measurement of the Top Quark Mass in the Dilepton Channel Using the Neutrino Weighting Algorithm at CDF II.* PhD thesis, University of Toronto, 2006.

[47] F. James *et al.* Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput. Phys. Commun.*, 10:343, 1975.

[48] E. Brubaker. *A Measurement of the Mass of the Top Quark in Lepton + Jets Events at CDF.* PhD thesis, University of California, Berkeley, 2004.

[49] S. Snyder. *Measurement of the Top Quark Mass at DØ.* PhD thesis, State University of New York at Stonybrook, 1995.

[50] Clive Loader. *Local Regression and Likelihood.* Springer, 1999.

[51] David W. Scott. *Multivariate Density Estimation: Theory, Practive and Visualization.* Wiley-Interscience, 1992.

[52] Ian S. Abramson. On bandwidth variation in kernel estimates - a square root law. *The annals of Statistics*, 10(4):1217, 1982.

[53] M.G. Bowler. A Note on the Fragmentation Functions of Discrete Particles on the Mennessier String. *Z. Phys. C*, 22:155, 1984.

[54] DØ Collaboration Y. Peters. Private communication.

[55] Precision electroweak measurements on the Z resonance. *Physics Reports*, 427(5-6):257–454, 5 2006.

[56] H. L. Lai *et al.* Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. *Eur. Phys. J. C*, 12:375, 2000.

[57] A. D. Martin *et al.* Parton distributions and the LHC: W and Z production. *Eur. Phys. J. C*, 14:133, 2000.

[58] J. Pumplin *et al.* New generation of parton distributions with uncertainties from global QCD analysis. *J. High Energy Phys.*, 07:012, 2002.