THE UNIVERSITY OF CHICAGO

MEASUREMENTS OF THE TOP-QUARK DECAY WIDTH AND MASS AT CDF
USING THE TEMPLATE METHOD

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
JIAN TANG

CHICAGO, ILLINOIS
MAY 10, 2012

Copyright © 2012 by Jian Tang

All Rights Reserved

TABLE OF CONTENTS

1	LIST OF FIGURES	vi
2	LIST OF TABLES	viii
3	ACKNOWLEDGMENTS	ix
4	ABSTRACT	xi
5	1 INTRODUCTION	1
6	2 THE THEORY	4
7	2.1 A Brief Description of the Standard Model	4
8	2.2 The Top Quark	6
9	3 THE EXPERIMENT	13
10	3.1 The Accelerator Chain	13
11	3.1.1 The Proton Source	13
12	3.1.2 The Main Injector	15
13	3.1.3 The Antiproton Source	15
14	3.1.4 The Recycler	16
15	3.1.5 The Tevatron	16
16	3.2 The CDF II Detector	17
17	3.2.1 Tracking Systems	18
18	3.2.2 Calorimeters	19
19	3.2.3 Muon Detectors	21
20	3.2.4 Trigger System	21
21	3.2.5 Monte Carlo Generation	22
22	4 THE EVENT SELECTION	27
23	4.1 Objects in an Event	27
24	4.1.1 Leptons	27
25	4.1.2 Jets	29
26	4.1.3 B-Tagging	38
27	4.1.4 Neutrino and Missing Transverse Energy	42
28	4.2 Event Selection	44
29	4.2.1 Trigger	44
30	4.2.2 Offline Event Selection for Top Width Measurement	45
31	4.2.3 Offline Event Selection for Top Mass Measurement	47
32	5 BACKGROUNDS	52
33	5.1 Method II for Lepton + Jets Channel	52
34	5.2 Method I for MET + Jets Final State Events	54
35	5.2.1 Background Parameterization	54
36		

1	5.2.2	Background Estimation	57
2	5.2.3	Candidate Events	58
3	5.2.4	Validation of Background Modeling	59
4	6	EVENT RECONSTRUCTION AND TEMPLATES	65
5	6.1	Top Quark Specific Corrections	66
6	6.2	The Lepton+Jets Channel	67
7	6.2.1	Reconstructed Top Mass m_t^{reco}	67
8	6.2.2	Dijet Mass w_{jj}	70
9	6.3	The MET+Jets Final State Events	73
10	6.3.1	Reconstructed Top Masses m_t^{reco} and $m_t^{\text{reco}(2)}$	73
11	6.3.2	Dijet Mass w_{jj}	74
12	7	THE LIKELIHOOD FIT	78
13	7.1	The Likelihood Fit Function	78
14	7.2	Kernel Density Estimation	83
15	7.2.1	An Introduction	83
16	7.2.2	Multi-Dimentional KDE	88
17	7.3	Local Polynomial Smoothing	91
18	8	STATISTICAL RESULTS	98
19	8.1	Extracting the Limits on Top Decay Width - A Feldman-Cousins Method	98
20	8.2	Top Mass Measurement	101
21	8.2.1	Bias Check	101
22	9	THE SYSTEMATIC EFFECTS	109
23	9.1	Sources of Systematic Effects	109
24	9.1.1	Residual Jet Energy Scale	109
25	9.1.2	b -Jet Energy Scale	111
26	9.1.3	Initial and Final State Radiation	111
27	9.1.4	MC Generators	112
28	9.1.5	Parton Distribution Functions	113
29	9.1.6	Gluon-Gluon Fusion Fraction	114
30	9.1.7	Background	115
31	9.1.8	Multiple Hadron Interaction	117
32	9.1.9	Color Reconnection	117
33	9.1.10	Other Systemtaic Effects	118
34	9.2	Summary of Systematic Effects of Top Mass Measurement	119
35	9.3	Summary of Systematic effects of Top Decay width measurement	119
36	9.3.1	Convolve Systematic Effects into Top Width Measurement	122
37	10	RESULTS	125
38	10.1	Results of Top Quark Width Measurement	125
39	10.2	Results of Top Quark Mass Measurement	128
40	11	CONCLUSIONS	132

1	REFERENCES	133
---	----------------------	-----

LIST OF FIGURES

2	2.1	SM Elementary Particles	5
3	2.2	Top quark decay	6
4	2.3	Tevatron Top Mass Measurement Combination	8
5	2.4	Feynman diagram of one-loop correction to W mass due to the top	9
6	2.5	Feynman diagram of one-loop correction to W mass due to the Higgs	10
7	2.6	M_W and M_{top} constraint contour curves of 68% confidence level	10
8	2.7	$t\bar{t}$ pair production through quark-quark interaction	11
9	2.8	$t\bar{t}$ pair production through gluon-gluon fusion	11
10	2.9	$t\bar{t}$ pair production branching fractions	12
11	3.1	The Fermilab accelerator chain system	14
12	3.2	Tevatron Run II Peak Luminosity as a function of time	17
13	3.3	Tevatron Run II integrated luminosity as a function of time	24
14	3.4	Elevation view of CDF II Detector	25
15	3.5	Schematic view of the particles passing through the tracking and calorimeter system	25
16	3.6	Functional block diagram of the CDF II data flow.	26
17	4.1	Di-jet balance, β as a function of jet η	32
18	4.2	Fractional systematic uncertainty on jet p_T due to the η -dependent corrections	32
19	4.3	Jet E_T versus the number of z -vertices	33
20	4.4	Fractional systematic jet energy uncertainties due to multiple $p\bar{p}$ interactions	34
21	4.5	Fractional absolute corrections on jets with cone $\Delta R = 0.4$	35
22	4.6	Fractional systematic uncertainty on the absolute energy correction	36
23	4.7	Corrections due to OOC effect and its systematic uncertainties	37
24	4.8	Total fractional systematic uncertainty on jet energy scale calibration	39
25	4.9	$t\bar{t}$	41
26	4.10	b -tagging efficiency for top events	43
27	4.11	Feynman diagram for $t\bar{t}$ production in the Lepton + Jets channel	46
28	4.12	Distribution of the neural network input variables	50
29	4.13	Distribution of the neural network input variables (continued)	51
30	5.1	b -tagging rate matrix	56
31	5.2	Distribution of kinematic variables for 1btag events	61
32	5.3	Distribution of kinematic variables for 1btag events(cont.)	62
33	5.4	Distribution of kinematic variables for 2btag events	63
34	5.5	Distribution of kinematic variables for 2btag events(cont.)	64
35	6.1	Distributions of reconstructed m_t^{reco}	71
36	6.2	Distributions of reconstructed w_{jj}	72
37	6.3	Distribution of reconstructed m_t^{reco} with different input top quark masses	75
38	6.4	Distribution of reconstructed $m_t^{\text{reco}(2)}$ with different input top quark masses	76
39	6.5	Distribution of reconstructed w_{jj} with different input JES's	77

1	7.1	One dimensional signal p.d.f.s estimated by KDE for the three variables used in the analysis, m_t^{reco} , w_{jj} , and $m_t^{\text{reco}(2)}$	89
2			
3	7.2	One dimensional background p.d.f.s estimated by KDE for the three variables used in the analysis, m_t^{reco} , w_{jj} , and $m_t^{\text{reco}(2)}$	90
4			
5	7.3	2D probability density function of signals from KDE for top width measurement - 1tag	92
6			
7	7.4	2D probability density function of background from KDE for top width measurement - 2tag	93
8			
9	7.5	2D signal p.d.f.s estimated by KDE for each pair of the variables for M_{top} measurement for $t\bar{t}$ signal events	94
10			
11	7.6	2D background p.d.f.s estimated by KDE for each pair of the variables, events with 4 jets	95
12			
13	8.1	$\Delta\chi^2$ distribution of sample	99
14	8.2	$\Delta\chi_c^2$ vs input top width	100
15	8.3	Data Fit of top width	100
16	8.4	Overlap of Fig. 8.2 and Fig. 8.3	103
17	8.5	Biaschecks of the measured top quark mass before any correction	104
18	8.6	Biaschecks of the measured top quark mass after calibrations	105
19	8.7	Checks of the measured top mass dependencies on the Δ_{JES}	106
20	8.8	Biaschecks of the measured Δ_{JES} vs the input Δ_{JES}	107
21	8.9	Expected statistical uncertainties of measured top quark mass	108
22	9.1	MHI systematics for top width	117
23	9.2	MHI systematics for top mass	120
24	9.3	Color Reconnection	121
25	9.4	Original likelihood function and the likelihood function after incorporating systematic effect	123
26			
27	9.5	Overlap of data fit and plot of $\Delta\chi_c^2$ vs input top width(95% CL) with systematic effect	123
28			
29	9.6	Overlap of Fig. 8.4 and Fig. 9.5(95% CL)	124
30	10.1	2D Log-likelihood fit of data	125
31	10.2	Reconstructed top mass distributions of dataset, overlaid with probability density functions	126
32			
33	10.3	Dijet mass of W boson distributions of dataset, overlaid with probability density functions	127
34			
35	10.4	Top quark width measurement at 68% CL	129
36	10.5	2D likelihood fit contours for the data fitting of top mass measurement	129
37	10.6	Distribution of kinematic variables m_t^{reco} , W_{jj} , and $m_t^{\text{reco}(2)}$ from data, overlaid with their corresponding 1-d p.d.f.s.	130
38			

LIST OF TABLES

1

2	4.1	Offline event selection in the $t\bar{t}$ lepton + jets channel.	46
3	4.2	L+J candidate events in data	47
4	5.1	Expected number of L+J background events	53
5	5.2	Event composition of a MC sample after all cuts are applied.	58
6	5.3	Expected and observed number of events after all cuts for 1tag events.	59
7	5.4	Expected and observed number of events after all cuts for 2tag events.	59
8	5.5	Expected and Observed number of events after all cuts within $4 \leq n_{\text{jets}} \leq 6$. . .	59
9	9.1	JES systematic effects for top quark Mass measurement.	110
10	9.2	JES systematic effects for top quark decay width measurement.	110
11	9.3	b -JES systematic effects for top quark mass measurement.	111
12	9.4	b JES systematic effects for top quark decay width measurement.	112
13	9.5	IFSR systematic effects for top width measurement.	112
14	9.6	IFSR systematic effects for top mass measurement.	112
15	9.7	Generator systematic effects for top width measurement.	113
16	9.8	Generator systematic effects for top mass measurement.	113
17	9.9	PDF systematic effects of top quark decay width measurement	114
18	9.10	PDF systematic effects of top mass measurement	114
19	9.11	Gluon-Gluon Fusion systematic effects for top width measurement.	114
20	9.12	Gluon-Gluon Fusion systematic effects for top mass measurement.	115
21	9.13	Background systematic effects for top width measurement.	116
22	9.14	Background systematic effects for top mass measurement.	116
23	9.15	Trigger Simulation systematic effects.	118
24	9.16	Summary of top mass systematics	119
25	9.17	Summary of top width systematics	120
26	10.1	Cross-check of top mass measurement	131

ACKNOWLEDGMENTS

The acknowledgments, seemingly easy to write since one can say whatever is in mind, is actually very hard for me since thousands of words just are not enough to express how grateful I am to people around me.

First and foremost I am grateful to my Ph.D. supervisor, Young-Kee Kim. Young-Kee is a very compassionate person. Every time I take the shuttle to commute to Fermilab I see her working with her laptop while the rest of the crowd sleeping over in the shuttle. Her passion affects everyone in the group, in a good way, and she is a real mentor because you can learn so much from her, both academically and socially. She has a very tight schedule, but she is able to sit down with every student at Bonjour Cafe every Sunday. I am also very thankful for Young-Kee's support and encouragement of my career decision.

Hyun Su Lee is both a great teacher and friend to me. I started to work with Hyun Su the moment I joined Young-Kee's group four years ago. I am surprised by how many ideas Hyun Su has and how fast he applies the ideas to research. He truly devoted himself to the high energy physics research. I feel very lucky to work with him because I can always get a lot of information every time I got stuck and discuss with him.

It is also very nice to have two of my classmates in the same group. Wesley Ketchum can always take care of things for the group, no matter if it is running for the group meetings or writing comments for other institutes. Satomi Shiraishi shared the office with me for a while, and I already saw her passion in collider physics.

I also would like to thank Ho lin Li, who shared her opinion in divination, and Yangyang Chen, who talked with me a lot on high-ended fashion and politics.

Anton Kapliy was very helpful on my FTK hardware project. I am impressed by how talented he is at hardware coding. Without his help it would be hard to be to finish the hardware project so quickly.

Tingjun Yang gave me a lot of advice on the top quark width measurement. I was able to solve some problem by consulting his experience in neutrino research.

1 I also want to thank all the CDF colleagues who I worked with, including all the conveners
2 and reviewers of my work.

3 Last but not least, I feel very grateful to my family, especially my parents and my wife,
4 who always support me no matter what happens.

ABSTRACT

Measurements of the top quark decay width and mass are presented using the $t\bar{t}$ events produced in $p\bar{p}$ collisions at Fermilab's Tevatron collider and collected by the CDF II detector.

A data sample corresponding to 4.3 fb^{-1} of integrated luminosity is used for the top quark width measurement. Two estimators, the reconstructed top quark mass and the mass of hadronically decaying W boson that comes from the top-quark decay are reconstructed for each event and compared with templates of different input top quark widths (Γ_{top}) and deviations from nominal CDF jet energy scale (Δ_{JES}) to perform a simultaneous fit for both parameters. Δ_{JES} is used for the *in situ* calibration of the jet energy scale at CDF. By applying a Feldman-Cousins limit-setting approach, we establish an upper limit at 95% confidence level (CL) of $\Gamma_{\text{top}} < 7.6 \text{ GeV}$ and a two-sided 68% CL interval of $0.3 \text{ GeV} < \Gamma_{\text{top}} < 4.4 \text{ GeV}$ assuming a top quark mass of $172.5 \text{ GeV}/c^2$, which are consistent with the standard model prediction.

The measurement of the top quark mass uses a data sample of $t\bar{t}$ events in 5.7 fb^{-1} of integrated luminosity collected by the same detector. Candidate events in the top quark mass measurement are required to have large missing transverse energy, no identified charged leptons, and four, five, or six jets with at least one jet tagged as coming from a b quark. This analysis considers events from the semileptonic $t\bar{t}$ decay channel, including events that contain tau leptons. The measurement is based on a multidimensional template method, in a similar way to the top quark width measurement, and the top quark mass is measured to be $M_{\text{top}} = 172.32 \pm 2.37 \text{ (stat)} \pm 0.98 \text{ (syst)} \text{ GeV}/c^2$.

CHAPTER 1

INTRODUCTION

The top quark (t) is the heaviest known elementary particle, with a mass approximately 40 times heavier than the mass of the bottom quark (b), its partner in the weak isospin doublet. It is important for physicists to know the properties of the top quark such as its decay width and mass. A precise measurement of the top-quark width Γ_{top} is a good test of the standard model, whose prediction at the Born level [1] is affected by quantum chromodynamics (QCD) radiative corrections of order 10% [2], as well as by electroweak corrections [3, 4], which are of order 1.5%. On the other hand, The top-quark mass can be used as a consistency check for standard model (SM) parameters. For example, the top quark contributes significantly to electroweak radiative corrections relating its mass (M_{top}) and that of the W boson to the mass of the predicted Higgs boson within the SM and in new physics models [5, 6]. Therefore, combined precision measurements of the W mass and M_{top} provide an important constraint on the Higgs boson mass.

The dominant decay mode of the top quark in the SM produces a W boson and a bottom quark (b). At the Tevatron, the top quark is produced predominantly in $t\bar{t}$ pairs, and the expected signature of a $t\bar{t}$ production event is $t\bar{t} \rightarrow W^+bW^-\bar{b}$, assuming unitarity of the three-generation Cabibbo-Kobayashi-Maskawa (CKM) matrix [7]. Since the W boson subsequently decays either to a quark-antiquark pair or to a lepton-neutrino pair, the final state of $t\bar{t}$ production can be classified by the number of charged leptons produced. In $t\bar{t}$ pair production, one W may decay leptonically to a lepton and its corresponding neutrino, and another hadronically to two quarks. This $t\bar{t}$ decay mode is called the lepton+jets channel. If both W decay leptonically, the $t\bar{t}$ decay is said to be from the dilepton channel, and all-hadronic channel if both W decay hadronically. In this thesis, I will present both the measurements of the top quark decay width and mass in the lepton+jets channel; however, the two measurements use different data events even though they come from the same $t\bar{t}$ decay mode, as will be discussed in Section 4.2.

At leading order the total top-quark width is given by

$$\Gamma_{\text{top}}^0 = |V_{tb}|^2 \frac{G_F M_t^3}{8\pi\sqrt{2}}$$

where V_{tb} , G_F and M_t are the CKM matrix element, Fermi coupling constant and the top-quark mass, respectively. If we take $|V_{tb}|$ to be unity, given a top-quark mass of 172.5 GeV/ c^2 , the equation above gives a Γ_{top} of 1.7 GeV, which would be lowered by around 22% due to corrections from mass effects and strong interactions. A deviation from the SM could indicate a significant contribution of non-SM particles. Novel top-quark decay modes motivated by the large top-quark mass include decay to a charged Higgs $t \rightarrow b + H^+$ [8, 9], decay to its supersymmetric scalar partner stop plus neutralinos [10, 11], and flavor changing neutral current (FCNC) top-quark decays [12]. Therefore, the direct measurement of Γ_{top} is a general way to constrain such processes. The first direct measurement of Γ_{top} was carried out with an integrated luminosity of 1 fb $^{-1}$ of CDF data in the lepton+jets channel [13] and set an upper limit on $\Gamma_{\text{top}} < 13.1$ GeV at 95% CL. Here we increase the data set to 4.3 fb $^{-1}$ using the same type of data events, apply a kernel density estimation (KDE, to be described in Section 7.2) technique to estimate density functions, and determine the jet energy scale (JES) calibration *in situ*. In addition, the methods for setting and incorporating systematic effects are different from the previous direct measurement. We are able to set a lower bound on the top-quark width at 68% CL for the first time.

For the top quark mass measurement, we focus on events with large missing transverse energy ($\cancel{E}_T$) [14] as expected for undetected energetic neutrinos, accompanied by jets. We explicitly veto events with identified high p_T electrons or muons that come from lepton+jets channel, as well as events from all-hadronic channel. This ensures that our result is statistically independent from other top-quark mass measurements [32, 33, 34, 35] and allows for a future combination with them. A previous measurement of M_{top} in this final state used an integrated luminosity of 311 pb $^{-1}$ [15] and yielded $M_{\text{top}} = 172.3 \pm 15.3$ (stat) $\pm$

1 14.4 (syst) GeV/c^2 . Although no identified leptons are explicitly required, our measurement
2 is sensitive to all W leptonic decays. This includes decays to τ leptons, which constitute
3 approximately 40% of the signal sample in our final selection. The frequency of observed
4 τ lepton final states is predicted to be enhanced by new physics models such as a charged
5 Higgs decay [16], therefore, a significantly different measurement of M_{top} in this decay chan-
6 nel could indicate contributions from non-SM [17] physics processes.

CHAPTER 2

THE THEORY

2.1 A Brief Description of the Standard Model

Great progress has been made over the past few decades in studying the elementary particles and their interactions, which resulted in the now widely accepted standard model (SM). Using the language of quantum field theory (QFT), the SM has been extremely successful in explaining and predicting phenomena in particle physics. The SM describes three out of the four fundamental forces in our universe—the strong, weak, and electromagnetic forces (the fourth one being the gravity). The strong force holds quarks together to form protons, neutrons and other hadrons, and it also binds the nucleons to form the nucleus of an atom. The strong interaction is mediated by gluons, which carry color charges that come in three types. The theory that describes the strong interaction is called quantum chromodynamics (QCD). The weak force is responsible for radioactive decay, and is mediated by the W and Z bosons. The typical strength of weak interaction is several orders of magnitude less than that of both the strong and electromagnetic forces. The electromagnetic force governs the interactions between electrically charged particles, and is mediated by photons. In 1968, the electromagnetic force and the weak interaction were unified into the electroweak force. The SM fails to include gravitation in the theory.

There are three generations of quarks in the SM. The first generation contains the up (u) and down (d) quarks, the second the strange (s) and charm (c) quarks, and the third the top (t) and bottom (b) quarks. Leptons also have three generations, the first being the electron (e^-) and electron neutrino (ν_e), the second the muon (μ^-) and muon neutrino (ν_μ), and the third the tau (τ^-) and tau neutrino (ν_τ). Particles in higher generations generally have greater mass and are not stable, and usually decay to lower generation particles. Figure 2.1 shows a summary of the elementary particles in the SM.

Three Generations of Matter (Fermions)				
	I	II	III	
mass →	2.4 MeV/c ²	1.27 GeV/c ²	171.2 GeV/c ²	0
charge →	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0
spin →	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
name →	u up	c charm	t top	γ photon
Quarks	4.8 MeV/c ²	104 MeV/c ²	4.2 GeV/c ²	0
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	d down	s strange	b bottom	g gluon
Leptons	<2.2 eV/c ²	<0.17 MeV/c ²	<15.5 MeV/c ²	91.2 GeV/c ²
	0	0	0	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	Z⁰ Z boson
	0.511 MeV/c ²	105.7 MeV/c ²	1.777 GeV/c ²	80.4 GeV/c ²
	-1	-1	-1	±1
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	e electron	μ muon	τ tau	W[±] W boson
				Gauge Bosons

Figure 2.1: Elementary particles in the standard model, including the three generations of quarks and leptons, as well as the force carriers.

2.2 The Top Quark

The existence of the top quark was postulated in 1973 by M. Kobayashi and T. Maskawa, and was discovered at the Fermilab Tevatron by the CDF and DØ detectors in 1995. With an electric charge of $+2/3$ and spin $1/2$, the top quark is the most massive of all observed elementary particles. It has a mass around $172.5 \text{ GeV}/c^2$, about the same mass as an atom of tungsten. Because it is so massive, the top quark has a very short life time, roughly $5 \times 10^{-25} \text{ s}$, so short that it cannot hadronize before it decays. The top quark almost exclusively decays to a W boson and a b quark, as shown in Fig. 2.2. The Ws and Wd final states of the top quark decay are suppressed by the CKM matrix elements V_{tb} and V_{td} , which, assuming unitarity of the three-generation CKM matrix, are less than 0.043 and 0.014, respectively, implying a value of $V_{tb} > 0.999$ [18]. In the SM, the theoretical top decay width at next-to-leading order is shown in Eqn. 2.1 [19]:

Figure 2.2: Feynman diagram of the top quark decay.

$$\Gamma_{\text{top}} = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi} \left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \quad (2.1)$$

where G_F is the Fermi coupling constant, M_W and m_t are the W boson and top quark masses respectively, and α_s is the strong coupling constant.

The top quark mass is a free parameter in the SM, and since its discovery there has

1 been much effort devoted to the precision measurement of the top quark mass. Figure 2.3
 2 shows the combined top quark mass measurement from the CDF and DØ experiment with a
 3 variety of data samples as of July 2011, yielding a combined measurement of $M_{\text{top}} = 173.2$
 4 $\pm 0.9 \text{ GeV}/c^2$.

The precision measurement of the top quark mass has an impact on other high energy physics research. In the search for the Higgs boson, a precise measurement of both the W boson and top quark mass can cast light on the mass range of the Higgs particle. At the tree level, the W boson mass is calculated as

$$M_{W0} = \frac{p_T^\alpha}{2G_F \sin^2 \theta_W}$$

5 where θ_W is the Weinberg angle. At the one-loop correction, the virtual top quark contributes
 6 to the W mass according to the Feynman diagram of Fig. 2.4. The W mass can be written
 7 in the form of Eqn. 2.2.

$$M_W = M_{W0} \frac{1}{\sqrt{1 - \Delta r}} \quad (2.2)$$

For the virtual top quark contribution,

$$(\Delta r)_{\text{top}} \approx -\frac{2G_F m_t^2}{8\sqrt{2}\pi^2 \tan^2 \theta_W} \quad (2.3)$$

8 Thus the one-loop correction to the W mass depends quadratically on the top quark
 9 mass.

The Higgs boson also contributes to the W mass correction Δr at one-loop level, as shown on Fig. 2.5, and the correction can be approximated by Eqn. 2.4, so the correction depends logarithmically on the Higgs mass.

$$(\Delta r)_{\text{Higgs}} \approx \frac{11G_F M_Z^2 \cos^2 \theta_W}{24\sqrt{2}\pi^2} \ln \frac{m_h^2}{M_Z^2} \quad (2.4)$$

Mass of the Top Quark

Figure 2.3: The combined top quark mass measurement at the Fermilab Tevatron as of July 2011. The combined top mass gives $M_{\text{top}} = 173.2 \pm 0.9 \text{ GeV}/c^2$, achieving a precision of less than 1%.

Using the corrections of Eqn. 2.3 and Eqn. 2.4, the mass of the Higgs boson can be expressed in terms of the W mass and the top quark mass, therefore very precise measurements of both M_W and M_{top} can narrow down the search range of the Higgs. Figure 2.6 shows the M_W and M_{top} constraint contour curves of 68% CL. From the direct measurement of the M_W and M_{top} , as shown in blue ellipse on the figure, it tells us that the SM prefers a light Higgs boson.

At the Fermilab Tevatron, the top quarks are produced in $p\bar{p}$ collisions at the energy of $\sqrt{s} = 1.96$ TeV, and the top quarks are produced predominantly in pairs through the QCD processes of $q\bar{q} \rightarrow t\bar{t}$ and $gg \rightarrow t\bar{t}$, whose Feynman diagrams are shown in Fig. 2.7 and Fig. 2.8, respectively. At $\sqrt{s} = 1.96$ TeV, the $t\bar{t}$ production cross section for a top mass of $M_{\text{top}} = 171$ GeV/ c^2 is about 7.61 pb, and approximately 85% of the $t\bar{t}$ events at the Tevatron come from $q\bar{q}$ annihilation, with the rest from gluon-gluon fusion [20].

Since the top quark decays to a b -quark and a W boson, which further decays to either two quarks or a lepton and its corresponding neutrino, $t\bar{t}$ pair production has three desired types of signatures, depending on how the W bosons of the top and anti-top quarks decays. As described in Chapter 1, the three $t\bar{t}$ decay channels are the lepton+jets, dilepton, and all-hadronic channels, and their corresponding branching fractions are shown on Fig. 2.9.

Figure 2.4: Feynman diagram of one-loop correction to W mass due to the top.

Figure 2.5: Feynman diagram of one-loop correction to W mass due to the Higgs.

Figure 2.6: Contour curves of 68% confidence level in the $M_{\text{top}}-M_W$ plane. The green band shows the SM constraint of the M_{top} and M_W for various values of the Higgs mass. The blue ellipse represents the 68% confidence level of direct measurements of the top and W masses, while the red ellipse comes from indirect measurements.

Figure 2.7: Feynman diagram of $t\bar{t}$ pair production through quark-quark interaction.

Figure 2.8: Feynman diagram of $t\bar{t}$ pair production through gluon-gluon fusion.

Top Pair Branching Fractions

Figure 2.9: $t\bar{t}$ pair production branching fractions.

CHAPTER 3

THE EXPERIMENT

The Fermilab Tevatron, completed in 1983 and decommissioned in 2011, was the world's highest energy particle collider before the running of the Large Hadron Collider (LHC) in Europe and the second after LHC. For decades, the Tevatron and two of its associated detectors CDF and DØ have made a lot of contributions to the particle physics world, including the prominent discovery of the top quark in 1995. This Chapter describes the Tevatron accelerator chain and the CDF detector complex.

3.1 The Accelerator Chain

The Fermilab accelerator collides the protons and antiprotons at a center-of-mass energy of 1.96 TeV. A chain of accelerator systems is needed in order to obtain collisions at such a high energy, and the primary parts of the chain includes the proton source, the Main Injector, and antiproton source, the Recycler and the Trvatron ring, as shown in Fig. 3.1.

3.1.1 *The Proton Source*

The collision of protons and antiprotons in the accelerator chain starts with the creation of the proton source. The first accelerator called "Preacc" provides the first stage of acceleration. In this device, hydrogen gas is ionized to create negative ions, each with two electrons and one proton, which are accelerated through a column from an electrically charged dome to the grounded wall to an energy of 750 KeV. The Preacc accelerates beam every 66 milliseconds (15 Hz). After exiting the accelerating column the beam travels through a transfer line and enters the Linear Accelerator (Linac), which is the second stage for accelerating the charged hydrogen ions. The oscillating electric fields in the Linac accelerate the negatively charged hydrogen ions to 400 MeV. The ions then pass through a carbon foil, where the

Figure 3.1: The Fermilab accelerator chain system.

1 electrons of the ions are removed there, leaving only the positively charged protons before
2 entering the next acceleration stage.

3 At the third stage, the Booster, which is a circular accelerator that uses magnets to bend
4 the beam of protons in a circular path and located about 20 feet below the ground, takes
5 the 400 MeV protons from the Linac and accelerates them to 8 GeV. The Booster is the
6 first circular accelerator in the chain of accelerators, and can accelerate beam once every 66
7 milliseconds. The accelerated proton beam in Booster is directed to the Main Injector.

8 *3.1.2 The Main Injector*

9 The Main Injector (MI) is a synchrotron that can accelerate the 8 GeV protons from the
10 Booster to either 120 GeV, which are used for antiproton production, or 150 GeV (for
11 antiprotons as well, the MI receives antiprotons from the Antiproton Source and accelerates
12 them to 150 GeV) and injects them to the Tevatron . MI has a size about seven times the
13 circumference of the Booster, and can accelerate beam every 2.2 seconds.

14 *3.1.3 The Antiproton Source*

15 The antiprotons are not as easily produced as the protons. To produce antiprotons, part of
16 the 120 GeV protons coming from the MI strikes a nickel alloy target and produces many
17 secondary particles, among which the 8 GeV antiprotons are collected and transtered into
18 the Debuncher. The Debuncher is a rounded trangular-shaped synchrotron whose primary
19 purpose is to efficiently capture the high momentum spread antiprotons coming from the
20 target. The Debuncher does not try to accelerate the beam, but maintains the beam at the
21 8 GeV energy. These 8 GeV antiprotons are then transferred to the Accumulator, where the
22 8 GeV antiprotons are stored and cooled until transferred to the Recycler. The Accumulator
23 is also a triangular-shape synchrotron and located is the same tunnel as the Debuncher.

3.1.4 *The Recycler*

The Recycler is a storage ring built along the ceiling of the Main Injector tunnel, and also provides as a cooling device for the antiprotons. The cooling system in the Recycler absorbs momenta from the antiprotons, allowing for more compact and brighter bunches to be sent to the Tevatron. The antiprotons inside of the Recycler are not accelerated and their energy is kept at 8 GeV.

3.1.5 *The Tevatron*

With a circumference of about four miles, the Tevatron is a circular synchrotron with eight accelerating cavities, and is the largest accelerator in the Fermilab accelerator chain system. Both the protons and antiprotons coming from the MI are accelerated in the Tevatron ring from 150 GeV to 980 GeV. The protons and antiprotons circle the Tevatron in opposite directions, and the beams cross each other at the CDF and DØ detectors located along the Tevatron ring. The magnets used in the Tevatron are made from superconducting niobium/titanium alloy that kept at extremely low temperature (4 K).

The Tevatron acts as a storage ring where the protons and antiprotons can collide with each other and produce secondary particles. Once the 980 GeV energy of protons and antiprotons is reached, the two counter-rotating beams pass through each other for hours until the beams fail. Each beam consists of 36 bunches and the bunches collide at a rate of 2.7 MHz. Typically 10^{13} protons and 0.3×10^{13} antiprotons are initially sent to the Tevatron ring, and the typical initial luminosity is approximately $3 \times 10^{32} \text{ cm}^2 \text{ s}^{-1}$. Each stable situation of 980 GeV proton and antiproton collision is called a store. As the particles are lost and the beam emittance increase the instantaneous luminosity decreases. Figure 3.2 shows the Tevatron Run II peak luminosity as a function of time, and Fig. 3.3 shows the total integrated luminosity delivered by the Tevatron as a function of the weeks that the Tevatron has been running. By the time the Tevatron was decommissioned in 2011, a total integrated luminosity of about 12 fb^{-1} was delivered, amounting for a large collection of

1 data for the physics research at the Fermilab Tevatron.

Figure 3.2: Tevatron Run II Peak Luminosity as a function of time.

3.2 The CDF II Detector

2

3 CDF II is a general-purpose detector located at one of the two collision points along the

4 ring of the Tevatron accelerator, and it combines precision tracking for charged particles

5 with fast projective calorimetry and fine grained muon detection. An elevation view of

6 the detector is shown on Fig. 3.4. CDF II detector has a silicon microstrip tracker and

7 a cylindrical drift chamber immersed in a 1.4 T magnetic field, which serve as a charged

8 particle tracking system. Electromagnetic and hadronic calorimeters are used to measure

9 the energies of electrons, photons and jets. Outside the calorimeters lie drift chambers which

10 can detect muons. A cylindrical coordinate system is employed for the detector, where θ

11 and ϕ are the polar and azimuthal angles, respectively, with respect to the proton beam,

12 and pseudorapidity $\eta \equiv -\ln \tan(\theta/2)$. Transverse energy and momentum are defined as

1 $E_T = E \sin \theta$ and $p_T = p \sin \theta$, respectively, where E and p are energy and momentum.
2 Each component of the detector will be described in the following.

3 *3.2.1 Tracking Systems*

4 The particle tracking systems provide precision measurement for tracks of charged particle at
5 the CDF II detector. The secondary particles from the $p\bar{p}$ collisions enter the silicon detector,
6 which consists of seven concentric cylinders of silicon that are arranged like barrels within
7 barrels and fit snugly around the beam pipe. The innermost silicon detector, called layer 00,
8 lies 1.4 to 1.6 cm from the beamline and provides axial tracking information. The silicon
9 vertex detector (SVX II) is a five-layer detector at small radii (extends from 2.5 to 10.6 cm
10 from the beamline) for micro-vertex detection. The SVX II has three cylindrical barrels with
11 a total length of 96 cm. All the five layers are double-sided silicon strip detectors, and three
12 of them combine an $r - \phi$ measurement on one side with 90^0 stereo measurement on the
13 other, and the rest two layers combine $r - \phi$ with small angle stereo at 1.2^0 . The outermost
14 silicon layer is called intermediate silicon layers (ISL). A single ISL layer is placed at a radius
15 of 22 cm in the central region of $|\eta_{det}| < 1.0$, and two layers of ISL are placed in the plug
16 region of $1.0 < |\eta_{det}| < 2.0$ at radii of 20 cm and 28 cm. Particles can pass through as many
17 as all seven layers of the silicon detector. From the collection of positions of the hits left by a
18 particle in the silicon detector, scientist can determine the paths of the particle by connecting
19 those hit information. Since the silicon detector is immersed in a magnetic field, particles
20 that carry electric charges will bend themselves when passing the detector. The slower the
21 particles are and the stronger the magnetic field is, the more they curve. By calculating the
22 amount the particles' curvature, scientists can determine the masses and momenta of them
23 and tell what kinds of particles were produced during the $p\bar{p}$ collision. The best tracking
24 position resolution achieved by the silicon detector system is $9 \mu\text{m}$. One important function
25 of the silicon detector system is to determine the position of the primary interaction of a
26 collision, which provides as basis for particle path determination. Also the high resolution

1 of the silicon detector system can help to tell whether a particle has come from the primary
2 collision, or from the decay of another particle that was produced in the collision.

3 The silicon tracking system has very good resolution in particle tracking, however they
4 are very expensive too. To keep the budget and reasonable tracking resolution, a central
5 outer tracker (COT) is put around the silicon detector system, covering radii between 44
6 and 132 cm and pseudorapidity of $|\eta_{det}| < 1.0$. The COT is an open-cell drift chamber filled
7 with a 1:1 mixture of argon and ethane gas, with thousands of wires strung across it to
8 catch the information left by charged particles that pass it. Inside of the chamber, there
9 are two kinds of wires, the sense wire, which picks up electrons released from the atoms
10 of the gas ionized by passing charged particles, and the field wires, which attract positive
11 ions that are left behind when the electrons are ripped off. Field wires are applied more
12 negative voltage than the sense wires so that electrons move in one direction and positive
13 ions move in the other. Released electrons travel toward the nearest sense wire and transport
14 the tracking information to computers for processing. It is desired that the packed wires are
15 dense enough for the electrons to take less time to reach the nearest wire, so that there is a
16 better chance to distinguish the particles produced in one collision from those of next. The
17 wires in the COT of CDF are 3.56 millimeters apart. The COT has resolution of the hit
18 position of $140 \mu\text{m}$, and momentum of $0.0015 \text{ GeV}/c$.

19 3.2.2 Calorimeters

20 While charged particles can have their momenta measured in the inner tracking system,
21 neutral particles can not since they can not ionize the materials in the silicon detector and
22 COT chamber and can not even bend in the magnetic field. Instead, calorimeters are used
23 to measure the energies of neutral particles, as well as charged particles. There are two types
24 of calorimeters, the electromagnetic calorimeter that measures the energies of electrons and
25 photons, and the hadronic calorimeter that measures the energy deposited by the jets from
26 the quark hadronization. In the central region of $|\eta_{det}| < 1.0$, the calorimeter system contains

the central electromagnetic calorimeter (CEM), central (CHA) and wall (WHA) hadronic calorimeters. In the plug region of $1.1 < |\eta_{det}| < 3.6$, there are plug electromagnetic (PEM) and plug hadronic (PHA) calorimeters.

The calorimeters are a series of wedges arranged into circles along both ends of the detector, and fit snugly around the COT and magnet. Each wedge contains the electromagnetic and hadronic calorimeters, from inside to outside. The electromagnetic calorimeter is a lead/scintillator sampling device with a unit layer composed of 4.5 millimeter lead and 4 millimeter scintillator. The detecting elements are arranged in a tower geometry that points towards the interaction region of the CDF II detector. Energy resolution in the CEM is $13.5\%\sqrt{E_T}$ plus a 2% constant term, and in the plug region the resolution is approximately $16\%\sqrt{E_T}$. The hadronic calorimeter uses steel instead of lead in its scintillator sandwich, and consists of 23 layers of 2 inch iron and 6 millimeter scintillator. The energy resolution for CHA is about $75\%\sqrt{E_T}$, $80\%\sqrt{E_T}$ for WHA and $80\%\sqrt{E_T}$ plus a 5% constant term for the PHA. The calorimeters measure the energy of particles through a mechanism called "showers", which happens when high energy particles collide with a layer of lead or steel and create a cascade of lower-energy and charged particles. These secondary particles then hit other particles that make up the lead or steel and create their own showers. This process continues until almost all the energies of the original particles are deposited in the scintillators of the calorimeter, which convert energies into light that carried away by the electronics. The hadronic calorimeter is much larger than the electromagnetic calorimeter since hadrons usually have to wait until they have a head-on collision with nucleus of an atom in the lead or steel before they start to shower, while the much lighter electrons and photons can shower merely due to coulomb interactions inside the metal, and thus the hadrons need much more space to deposit their energies in the calorimeter.

3.2.3 Muon Detectors

The muon detector are composed of two parts, the muon drift chamber, which works in much a similar way to the COT and accurately measures the position of muons, and the scintillators, which measures the energies of muons. Muons pass through the calorimeters as minimum ionizing particle and deposit very little energy there. Thick walls of steel are put between the muon chambers and calorimeter to stop other particles coming out from the calorimeter so that they are not mixed with muons. The scintillators are put behind the muon chambers. Figure 3.5 shows the relative position of different tracking and calorimeter systems in the CDF II detector and where each type of particles leave their information. In the region of $|\eta_{det}| < 0.6$ lies the central muon detector (CMU), behind which lies the central muon upgrade (CMP). In the region of $0.6 < |\eta_{det}| < 1.0$ there is the central muon extension (CMX).

3.2.4 Trigger System

At the luminosity of $3 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$, three million events are produced every second at the CDF II detector. Most of these events are of no interest to modern high energy physics research since they do not provide new information to what has already been known. A three-level trigger system was designed at CDF to quickly identify potential interesting events and discard the uninteresting ones. Figure 3.6 shows the data flow at CDF II detector and its trigger system, which reduces the event rate from three million per second down to 150 per second.

Level-1 (L1) trigger uses custom designed hardware to find physics objects based on a subset of the detector information and make simple decisions for these objects. L1 trigger performs track finding, which allows a track to be matched to an electromagnetic calorimeter energy cluster for improved electron identification, or to a stub in the muon system for better muon identification and momentum resolution. L1 trigger also takes uses of calorimeter energies, missing energies, etc. to make decisions. The latency, which is defined as the time

require for a trigger to make decisions, for L1 trigger is 5.5 microseconds, and after L1 trigger requirement about 10,000 to 25,000 events per second are output to the Level-2 trigger.

Level-2 (L2) trigger also uses custom hardware, and it makes a limited event reconstruction which can be processed in programmable processors. Two main pieces of hardware are used at L2, the secondary vertex trigger (SVT), which uses the silicon hit information to trigger on the presence of tracks with displaced vertices, and the cluster finder hardware that forms cluster of energies to reconstruct more sophisticated electron and jet candidates. L2 system has improved momentum resolution for tracks, finer angular matching between muons stubs and central tracks, and better identification of electrons and photons. The latency for L2 is about 30 microseconds, and an average of only 900 events pass L2 trigger requirements per second.

Level-3 (L3) accepts events passing L2 trigger, and uses the full detector resolutions to fully reconstruct events for analysis. The events are then passed to a trigger algorithm that categorizes the events and determines whether the events should be written to mass storage. L3 makes 3-dimensional track reconstruction and tight matching of tracks to calorimeter and muon-system information. The L3 trigger has latency of 1 second, and only about 150 events make it through the trigger requirement per second.

3.2.5 Monte Carlo Generation

The analysis takes uses of a lot of Monte Carlo (MC)-simulated samples, which serve as the basis for the methodology used for both the top quark decay width and mass measurement. The $t\bar{t}$ signal samples are generated by the PYTHIA 6.216 [21], which is a leading-order MC generator. We also use signal MC samples generated by HERWIG 6.510 [22] to evaluate some systematic effects of the analysis. For the top quark decay width measurement, we use PYTHIA to generate samples of background processes such as diboson production (WW, WZ, ZZ), and use MADEVENT [23] to simulate the single-top background sample events in the s -channel and t -channel. W + jets background process is simulated

₁ with ALPGEN 2.10 [24].

Figure 3.3: Tevatron Run II integrated luminosity as a function of time.

Figure 3.4: Elevation view of CDF II Detector.

Figure 3.5: Schematic view of the particles passing through the tracking and calorimeter system.

Figure 3.6: The CDF II trigger system and data flow.

CHAPTER 4

THE EVENT SELECTION

Top-quark properties related analysis are conducted using $t\bar{t}$ pair events produced at the Tevatron. This thesis focuses on the top-quark decay width measurement with events from the "lepton+jets" $t\bar{t}$ pair decay channel, and the top-quark mass measurement with events having the "MET+jets" final state signature. The later essentially also comes from the "lepton+jets" $t\bar{t}$ pair decay channel, and they are triggered by multiple high energy jets. However, it is customary to label events to be "lepton+jets" events if they are triggered by a high momentum (or energy) lepton. So from here, without specification we will just simply call the candidate $t\bar{t}$ events that are triggered by a high momentum electron or muon "lepton+jets" events, and call the events triggered by multiple jets and have missing transverse energy in the final state as "MET+jets" events.

In the following we will examine the event selection details of both analysis. Some of the event selection are common to both analysis, such as the reconstruction and correction of jets, while others are not.

4.1 Objects in an Event

Any physics processes present themselves in the detector as collections of tracks with specific momenta, or hits that deposite energy in the calorimeter, or large missing transverse energy due to energy imbalance in the transverse plane, etc. In the following, we will talk about the objects that are associated with events: leptons, neutrinos, and jets, at the trigger level.

4.1.1 *Leptons*

Central Electrons

Central electrons are defined as follows.

At Level 1, energies in physical calorimeter towers of 0.1×15^0 in the $\eta_{det} - \phi$ space are first summed into 0.2×15^0 trigger towers. The central electron trigger requires at least one trigger tower to have $E_T > 8$ GeV and the ratio of the hadronic to electromagnetic energies in that tower, E_{had}/E_{em} , to be less than $1/8$. Besides, a track with $p_T > 8$ GeV/ c pointing in the direction of the tower must be found in the COT.

At Level 2, a clustering algorithm is run to combine associated energy deposits in the neighboring calorimeter towers. The total E_T of the cluster is required to be less than 16 GeV, its E_{had}/E_{em} is required to be less than $1/8$, and there must be a matching XFT track with $p_T > 8$ GeV/ c .

Finally, at Level 3 an electromagnetic cluster with $E_T > 18$ GeV and $E_{had}/E_{em} < 1/8$ must be found, and a track pointing at the cluster with $p_T > 9$ GeV/ c must be found by the COT track reconstruction algorithm in the Level 3 processors.

Central Muons

Central Muons are divided into two categories, the most central region of the detector ($|\eta_{det}| < 0.6$) where high- p_T muons deposit hits in both the CMU and CMP wire chambers, and the extended central region ($0.6 < |\eta_{det}| < 1.0$) when muons have hits in the CMX wire chambers.

At Level 1, the trigger requires matched hits in one or more CMU projective wire pairs with arrival times within 124 ns of each other, and a pattern of CMP hits on three or four layers consistent in ϕ with the observed CMU hits, and a matching COT track found by the XFT with $p_T > 4$ GeV/ c . The p_T requirement of muons is increased to 15 GeV/ c at Level 2, and 18 GeV/ c at Level 3. These events are referred to as CMUP events. Trigger requirements in the extended central region are similar, and these events are called CMX events.

A complete description of electron and muon identification can be found in [25].

4.1.2 Jets

At the CDF detector, quarks or gluons from a scattering process in $p\bar{p}$ collisions present themselves as clustering of hits in the calorimeter, called jets. These jets carry the energy of the parent partons. Algorithms are applied to calculate the energy of these jets, and these jets are then corrected to correspond to the energy of the parent parton. The precision to which this can be achieved is very important because it determines how precise we can measure the top-quark mass and decay width. In the following, we will talk about the algorithm used to cluster the jets, and the corrections applied to the jet energy.

Jet Clustering

At CDF an algorithm called JETCLU [26] which uses a cone of a fixed radius is used to reconstruct jets. The algorithm starts by selecting 'seed' towers in the calorimeter that have $E_T > 1.0$ GeV. Preclusters are then formed from a chain of contiguous seed towers with decreasing tower E_T . If a tower is outside the 7×7 tower window surrounding the seed, it will be used to form a new precluster.

Preclusters are then used to create clusters. First the E_T weighted centroid of a precluster is calculated and a cone with radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ of 0.4 is created around this centroid. Towers whose E_T is greater than 0.1 GeV and centroid is within the cone are then included in the cluster. A new E_T -weighted centroid of the cluster is then calculated using all the towers within the cone, and a new cone is formed around this new center. This process continues iteratively by adding new or removing old towers until the collection of towers in the cone do not change.

Clusters built this way may overlap with each other. If a cluster is completely contained in another cluster, the smaller one is discarded. If two clusters have some finite overlap, then an overlap fraction is calculated as the ratio of the sum of the E_T of the common towers to the sum of the E_T of the smaller cluster. If the ratio is above a threshold (0.75) then the two clusters are merged into one. Otherwise, the two clusters are kept distinctive, and each

tower in the overlap region is assigned to the cluster closest in the $\eta - \phi$ space, and further iterative computing processes are applied to each cluster until the tower collections in the clusters remain unchanged.

Jet Energy Corrections

The energies and momenta calculated above depend only on the energy deposited in the calorimeters of the detector, and they differ from the values at the parton level. In the following, we describe a series of jet energy correction procedures that are applied to approximate the corresponding quark energies, and a thorough description of the procedure can be found in [27].

η -Dependent Corrections:

Due to the separation of calorimeter components at $|\eta| = 0$ where two halves of the central calorimeter join and at $|\eta| \approx 1.1$ where the plug and central calorimeters join, the CDF calorimeter response is not uniform across η . The η -dependent corrections are introduced so that jets with different η_{det} have the same E_T response as those jets in the well-understood region of the detector, $0.2 < |\eta_{det}| < 0.6$.

The correction procedure uses a "dijet-balancing method" which takes uses of dijet events. First, a jet with $0.2 < |\eta_{det}| < 0.6$ is defined as a "trigger jets" and the other jets is called a "probe jet". When both jets are in the region of $0.2 < |\eta_{det}| < 0.6$, the trigger and probe jets are assigned randomly. The dijet balance is defined as

$$\beta = \frac{p_T^{probe}}{p_T^{trigger}} \quad (4.1)$$

The η -dependent corrections are defined as $1/\beta$ and are determined separately for data and MC and for different jet p_T bins. Figure 4.1 shows the dijet balance of β as a function of jet η for data and PYTHIA MC samples.

Ideally the jet E_T response after the η -dependent corrections should be uniform. Discrepancies from a uniform distribution are due to the limitations of the η - and p_T - dependence of the correction and are taken as part of the systematic uncertainty of the corrections. Further systematic effects take into account the event selection requirements and the fitting procedure. The overall systematic uncertainty from these effects are shown in Fig. 4.2, from which we can see that the uncertainty can be as low as 0.5% for high- p_T jets in the central regions of the detector, and as high as 7.5% for low- p_T jets in the plug and extended region.

Multiple $p\bar{p}$ Interactions:

As the instantaneous luminosity gets higher, more than one $p\bar{p}$ interactions occur in the same bunch crossing, and they increase the energy of the jets from the hard scattering process if their final state hadrons overlap with the jets, thus energies from the multiple interactions must be subtracted from the jets.

The number of reconstructed z -vertices, N_{vtx} , is used to estimate the number of interactions per bunch crossing. To obtain the jet energy correction due to multiple interactions, the transverse energy E_T within a fixed cone of $\Delta R = 0.4$ is measured as a function of the average number of N_{vtx} per bunch crossing, using minimum bias data samples. The cone is defined by randomly selecting a seed tower in the central calorimeter region $0.2 < |\eta| < 0.6$. Figure 4.3 shows the E_T in the cone as a function of N_{vtx} , which shows an increase of 356 MeV in jet E_T for an increase of 1 interaction per bunch crossing. The non-zero interception in the y -axis in Fig. 4.3 is due to the imperfect vertex finding algorithm.

Systematic uncertainties are assigned to the jet energies due to the vertex reconstruction efficiency, and the vertex fake rate due to the possibility that the vertex finding algorithm reconstruct vertices even though no interaction took place. These effects are estimated by repeating the multiple interaction measurement using samples of $W \rightarrow e\nu_e$, minimum bias and another jet sample, and the fractional systematic uncertainties for jets with different cone sizes are plotted in Fig. 4.4. The figure shows that for a jet of cone size 0.4, the

Figure 4.1: Di-jet balance, β as a function of jet η for data and PYTHIA MC samples. Jets correspond to cone size of $\Delta R = 0.4$ and $55 \text{ GeV}/c < p_T < 75 \text{ GeV}/c$.

Figure 4.2: Fractional systematic uncertainty on jet p_T due to the η -dependent corrections as a function of jet η_{det} . Different lines in the figures correspond to different jet p_T bins.

Figure 4.3: Jet E_T versus the number of z -vertices for cone size of $\Delta R = 0.4$.

1 uncertainty for low- p_T jet is only around 1%, however, this correction is not enough for $t\bar{t}$
2 MC samples since the instantaneous luminosity of MC samples does not match data, and
3 this effect will be taken into consideration when we evaluate the systematic effects on the
4 top-quark mass measurement in Section 9.1.

Figure 4.4: Fractional systematic jet energy uncertainties due to multiple $p\bar{p}$ interactions, for jets with cone size ΔR of 0.4, 0.7, and 1.0.

5 Absolute Energy Correction:

6 The absolute energy correction is applied to jet energies after the multiple interaction cor-
7 rection to account for any non-linearity and energy loss in the un-instrumented regions of
8 the calorimeters, so that the energy scale of a jet is independent of the CDF detector and
9 the jet energies should correspond to the total energy of all the hadrons inside the jet cone.

MC simulated dijet samples are used to estimate the correction, and the accuracy of the correction depends on how well jets are modeled by the simulation, that is, the multiplicity and p_T spectrum of the particles inside a jet cone, and the response of the calorimeter to an individual particle. Ideally this correction should be different for jets originated from gluons and quarks since the former on average has a larger multiplicity than the later. However, the corrections are not derived separately since the parton type of each jet in data is not known priorily.

The absolute jet energy is defined as the most probable value for a jet p_T^{jet} , given a particle jet with a fixed $p_T^{particle}$. The probability density function is parameterized as a function of $\Delta p_T = p_T^{particle} - p_T^{jet}$ [27], and an unbinned likelihood fit is used to extract the most probable jet p_T . Figure 4.5 shows the absolute correction for jets with cone radius $\Delta R = 0.4$, from which we see the correction can be as high as 1.4 for low- p_T jets and for high- p_T jets there is an asymptotic correction around 1.1.

Figure 4.5: Fractional absolute corrections on jets with cone $\Delta R = 0.4$.

The uncertainties from the absolute correction depends on how the MC simulates the particle response in data, the modeling of hadronization effects, and the stability of the calorimeter energy scale, etc. The total systematic uncertainty on jet energy scale due to

absolute correction is shown on Fig. 4.6, which shows a rise of uncertainty from 2% at low-
 p_T^{jet} to 3% at high- p_T^{jet} , and the dominant effect comes from the simulation of the calorimeter
response to hardons.

Figure 4.6: Fractional systematic uncertainty on the absolute energy correction. The solid line shows the total uncertainty and others show individual contributions.

4 Out-of-Cone Energy and Underlying Events:

After the absolute corrections, jet energies correspond to the most probable value of energy
within the jet cone, but it is often desirable to reconstruct the energy of the original parton
rather than the jet, e.g. in the measurement of the top quark mass, parton energies are
needed to compute the invariant mass of the decay products. However, the reconstruction of
parton jet energy from the particle jet energy is subject to some difficulties. First, a fraction
of the parton energy maybe lost from the jet cone due to final state gluon radiation (FSR)
at large angles with with respect to the parent parton or due to low- p_T particles that are

bent away from the cone by the strong magnetic field of the detector. These energy loss are called out-of-cone (OOC) effects. Besides, a particle jet can have contributions not related to the actual parent parton, such as particles from the initial state gluon radiation (ISR) or from spectator partons with color connections to other partons of the proton, and these two effects are named Underlying Energy (UE).

Corrections for OOC and UE are determined simultaneously using PYTHIA MC dijet samples, and a parameterization procedure of the difference of the energy between the particle jet and parton is performed in the same manner as in the absolute correction. The study shows that the energy loss due to OOC dominates over the energy increase due to the UE. For jet cone size of $\Delta R = 0.4$, the UE energy is about 0.4 GeV with a systematic uncertainty around 30%. The OOC corrections and associated systematic uncertainties are shown in Fig. 4.7. The systematic uncertainties for the OOC and UE corrections are derived from the comparison of the energy measured in calorimeter towers in certain annuli around the jet cone with samples from data and MC.

Figure 4.7: Corrections due to OOC effect and its systematic uncertainties.

1 Total Jet Energy Scale Uncertainty and Δ_{JES} :

2 Figure 4.8 shows the individual and overall fractional systematic uncertainties on the jet
3 energy scale calibration. The systematic effects quantifies the extend to which we model
4 the jet energy calibrations, and the overall jet energy calibration is referred to as the "jet
5 energy scale" (JES) at CDF. The unit of JES at CDF is defined as σ_c , corresponding to
6 the combined fractional uncertainty on jet energy calibration, and shifts in the JES from
7 the expected value are called Δ_{JES} . The *a priori* Δ_{JES} estimate at CDF is thus $0 \pm 1\sigma_c$
8 by definition. If every correction above is shifted up by one σ uncertainty, then we say
9 that the jet energies are obtained with $\Delta_{\text{JES}} = +1\sigma_c$. Accordingly, if all the calibration
10 procedures are shifted down by one unit of uncertainty, then the jet energies are obtained
11 with $\Delta_{\text{JES}} = -1\sigma_c$.

12 *4.1.3 B-Tagging*

13 The $t\bar{t}$ event signature has two b quarks in the final state. A requirement of the presence of
14 b quarks in the offline selection can enhance the probability of selecting the $t\bar{t}$ signal events.
15 Becasue the b quarks have relatively long lifetime compared with other quarks, and also
16 due to their semi-leptonic decay feature, we can take use of these to identify whether a jet
17 originated from a b quark. In this section, we will talk about a b -tagging algorithm called
18 SECVTX [28].

19 SECVTX is based on the reconstruction of the primary event interaction vertex and a
20 secondary vertex due to displaced tracks of jets. The distribution of primary $p\bar{p}$ interaction
21 vertices along the beam line direction follow a Gaussian with a width of approximately 28
22 cm. In the plane transverse to the beam axis, these vertices are distributed according to
23 a Gaussian with a width of 25 μm in both the x and y axes. The reconstruction of the
24 primary vertices starts by projecting the tracks back to the beam axis, which determines
25 the logitudinal position with a precision of around 0.2 cm. The transverse position of the
26 primary vertices is determined by applying a iterative procedure to the tracks in the events,

Figure 4.8: Total fractional systematic uncertainty on jet energy scale calibration as a function of correction jet p_T .

and depending on the number of tracks and the event topology, has uncertainties in the range of 10-25 μm .

To search for a secondary vertex in a jet, at least three candidate tracks are required to be in the jet cone, and tracks are discarded if they are consistent with coming from the decay $K_s \rightarrow \pi^+\pi^-$ or $\Lambda \rightarrow \pi^-p$. A seed vertex is formed by constraining two tracks to pass through the same space point, and this seed vertex is used as the secondary vertex if at least an additional track is consistent with intersecting the seed vertex. If these steps are not successful in finding a secondary vertex, then a second pass is attempted with more stringent track requirement. A seed vertex is again formed with these tracks and a χ^2 is calculated for this seed vertex, and an iterative procedure is applied to the seed vertex until a seed vertex has at least three associated tracks and an acceptable χ^2 value.

After a secondary vertex is found, the decay length of the secondary vertex L_{xy} , defined as the projection of the two-dimensional vector pointing from the primary vertex to the secondary vertex on the jet axis, as shown in Fig. 4.9, is calculated. L_{xy} is mostly positive for b and c -hadrons, but for random combination of mismeasured tracks it is expected to have a symmetric distribution around $L_{xy} = 0$ [29]. To reduce the chance of falsely tagging a jet as coming from a b quark, SECVTX requires $L_{xy}/\sigma_{L_{xy}} \geq 7.5$ (positive tag), where $\sigma_{L_{xy}}$ is the calculated uncertainty on L_{xy} . If a jet has $L_{xy}/\sigma_{L_{xy}} < -7.5$, then it is negatively tagged. In this analysis, a jet is called b -tagged only if it is positively tagged.

The tagging efficiency for $t\bar{t}$ events is defined as how often at least one of the jets in a $t\bar{t}$ event is positively tagged by the SECVTX algorithm. A detailed description of the measurement of tagging efficiency at CDF can be found in [30]. It is impossible to directly measure the tagging efficiency in $t\bar{t}$ data events, so an indirect method is applied. A data sample of jets whose heavy flavor (b or c -hadrons) fraction can be measured is used to derive the per-jet tagging efficiency in that sample. Since c -hadrons also have finite lifetime, a double-tag requirement that requires a second tagged jet is applied to the sample to remove the charm components, so that the measured tag efficiency is dominated by the contribution

Figure 4.9: An illustration of how to determine if a jet originated from a b -quark.

from bottom. Because jets in $t\bar{t}$ events in general have different characters from those in the calibration sample, such as their energies, pseudorapidities and track multiplicity, the measured efficiency can not be used directly. A matching sample of MC jets is used and its tagging efficiency is measured. The ratio of the tagging efficiency in the data to that in the matching MC sample (scale factor) is calculated to correct the tagging efficiency in $t\bar{t}$ MC samples. Figure 4.10 shows the b -tagging efficiency for top events as a function of both jet E_T and η . The $t\bar{t}$ b -tagging efficiency can be as high as 50%, but falls as jet energy goes higher since high energy jets have high density tracks, which makes it hard to find good quality tracks. The scale factor for $t\bar{t}$ b -tagging efficiency is 0.96 ± 0.05 .

In both the top quark width and mass measurement, we will divide the candidate event samples into subsamples of events that have exactly one tagged jet (1tag event) and that have two or more b -tagged jets (2tag event), in order to improve the statistical sensitivity of the measurements.

4.1.4 Neutrino and Missing Transverse Energy

Neutrinos from the leptonic W boson decays escape the detector undetected, however, we can still extract information about neutrinos from the imbalance of the observed energy in the transverse plane since the total momentum of the $p\bar{p}$ system transverse to the beam direction is zero. This missing transverse energy ($\cancel{E}_T$) is defined with respect to the event's interaction vertex:

$$\cancel{E}_T = \left| \sum_i -\vec{n}_i E_T^i \right| \quad (4.2)$$

where $\vec{n}_i$ is the unit vector in the transverse plane that points to the i -th calorimeter tower in the central, wall and plug region. The $\cancel{E}_T$ is corrected using the muon track momentum when there is a muon identified in the event since muons deposit very little energy in the calorimeter and the momentum measured in the muon chamber has to be used instead.

However, the $\cancel{E}_T$ is a little different in the top mass measurement than in the top width

Figure 4.10: b -tagging efficiency for top events as a function of both jet E_T and η .

measurement, due to different triggers used for both analysis, as will be described in Section 4.2.1. In the top width measurement, the trigger identifies a lepton, and the majority of the $\cancel{E}_T$ comes from the escaping neutrino, while in the top mass measurement the trigger does not require any lepton and offline the analysis rejects events with leptons, therefore the $\cancel{E}_T$ of the top mass measurement has contributions from both the neutrino and lepton, if any.

4.2 Event Selection

4.2.1 Trigger

The trigger used in an analysis is the first step to select the candidate $t\bar{t}$ events, which happens online. Different triggers aim to select different types of events, although these events may overlap with each other.

Trigger for Top Quark Decay Width Measurement:

The signature of $t\bar{t}$ pair decay in the Lepton + Jets channel (shown in Figure 4.11) has a lepton, 4 quarks, including 2 b quarks, and momentum imbalance in the transverse plane due to the escaping neutrino.

In the detector, the events are selected by a trigger that requires a single high- p_T , isolated electron or muon. The tau lepton is not considered here because it decays to further daughter products, which makes it very difficult to identify. The lepton required in this trigger is the same as described in Section 4.1.1.

Trigger for Top Quark Mass Measurement:

The trigger used for the top quark mass measurement requires multiple jets in the final state. Since CDF Run-II, this multijet trigger has undergone several revisions [31] to cope with the increasing luminosity of Tevatron. The trigger requires the presence of at least

one central tower with $E_T > 10$ GeV at Level 1 trigger, which was strengthened to 20 GeV since June 2006, and at least four calorimetric clusters with $E_T > 15$ GeV at Level 2 and a total transverse energy greater than 125 GeV. The requirement of total transverse energy at Level 2 was increased to 175 GeV after February 2005 to reduce the trigger rate at higher instantaneous luminosity. The jet clustering algorithm was also changed since October 2007. At Level 3, at least four jets with $E_T > 10$ GeV are required. Although this trigger was originally designed to select $t\bar{t}$ candidate events in the all-hadronic channel, where both top and anti-top quarks decay hadronically. However, it is found that this trigger also has large acceptance to $t\bar{t}$ events with a signature of large $\cancel{E}_T$ and jet multiplicity. The motivation of using events from this trigger is to perform top mass measurement with events independent of other CDF top mass analysis.

4.2.2 Offline Event Selection for Top Width Measurement

As mentioned above, the top width measurement uses the events in the lepton + jet channel and triggered by a high- p_T lepton. Offline, further selection requirements are applied to these events to suppress background processes such as the production of a W boson with four or more jets. The identified lepton is required to have $p_T > 20$ GeV/ c for an electron, or $E_T > 20$ GeV for a muon. We require at least four jets in an event, and at least one jet must be b -tagged in an event. b -tags are only allowed for the leading (most energetic) four jets. If an event has a single b -jet (a jet that is b -tagged) it is required to have exactly four tight jets ($E_T > 20$ GeV and $|\eta_{det}| < 2.0$) in total. For events that have two or more b jets, the background contamination is very small, so we loose the requirement on these events and require only three tight jets, while allowing for extra tight or loose jets ($E_T > 12$ GeV and $|\eta_{det}| < 2.4$). The missing transverse energy $\cancel{E}_T$ is at least 20 GeV. A summary of the offline event selection in the lepton + jets channel is in Table 4.1.

Later in Section 6.2 another kinematic requirement will be applied in addition to the above selection, which is $\chi^2 < 9.0$, where the χ^2 is a chi-square value from reconstructing

Figure 4.11: Fynman diagram for $t\bar{t}$ production in the Lepton + Jets channel.

Table 4.1: Offline event selection in the $t\bar{t}$ lepton + jets channel.

	1-tag	2-tag
b -tags (leading 4 jets)	= 1	> 1
Lepton p_T (GeV/c), E_T (GeV)	> 20	> 20
$\cancel{E}_T$ (GeV)	> 20	> 20
4th jet E_T (GeV)	> 20	> 12
Extra jets E_T (GeV)	> 20	Any

Table 4.2: The number of events observed in the Lepton + Jets channel for data of 4.3 fb^{-1} .

	CEM	CMUP	CMX	Total
1-tag == 4 tight jets	294	155	93	542
2-tag	114	62	38	214
2-tag == 3 tight jets (> 0 loose jets)	25	21	8	54
2-tag == 4 tight jets	72	28	21	121
2-tag > 4 tight jets	17	13	9	39
All	408	217	131	756

1 the top quark mass using a chi-square fitter.

2 The top quark decay width measurement uses CDF II data samples that were collected
3 between February 2002 and March 2009, corresponding to a total integrated luminosity of
4 4.3 fb^{-1} . After applying all the event selection requirements to these data samples, 756
5 events are left, with details shown on Table 4.2.2.

6 4.2.3 Offline Event Selection for Top Mass Measurement

7 The offline selection for events of multijet trigger starts with a preliminary event clean up
8 for events passing the trigger requirement, after which a artificial neural network selection
9 is performed.

10 Preliminary Selection:

11 First, events satisfying the trigger requirements are retained only if they were collected with
12 fully-operational tracking detectors, calorimeter and muon systems, and if their primary
13 vertex is located within $\pm 60 \text{ cm}$ of the center of the CDF II detector along the z axis.
14 Jets and $\cancel{E}_T$ are corrected for multiple $p\bar{p}$ interactions, nonuniformities in the calorimeter
15 response along η , and any nonlinearity and energy loss in the uninstrumented regions of the
16 calorimeters, as described in Section 4.1.2. We also require the vertex used for jet clustering
17 to be within $\pm 5 \text{ cm}$ of the primary vertex of the event. We only consider jets with $E_T >$
18 15 GeV and $|\eta| < 2.0$; we define the total transverse energy of the event $\sum E_T$ as the sum

of all jets E_T .

At least three jets are required to be present in the event, any event that has central high- p_T ($p_T > 20$ GeV/ c) electrons or muons are discarded so that this measurement used events totally independent of the CDF top quark mass measurements that uses events with lepton and jets in the final state [32, 33, 34]. Similarly, to avoid overlap with top mass measurement using events coming from the all-hadronic $t\bar{t}$ decay channel at CDF [35], we reject events that have $\cancel{E}_T$ significance $\cancel{E}_T^{sig} < 3$ GeV $^{1/2}$, which is defined as $\cancel{E}_T^{sig} = \cancel{E}_T / \sqrt{\sum E_T}$.

Neural Network Selection:

Due to the trigger type we choose to make this measurement, the dominant background processes come from QCD multijet and W boson + jets events. After the preliminary event selection, we perform a neural network (NN) selection on the events to discriminate the $t\bar{t}$ signal events from background. We use the neural network implementation in ROOT [36] to build the NN. Eight variables are built for each event as inputs to the NN, as listed below:

- E_{T1} : the transverse energy of the leading (most energetic) jet in the event.
- $\Delta\phi Min$: the minimum difference of angle ϕ between the $\cancel{E}_T$ and each jet in the event.
- $\cancel{E}_T^{sig}$: the $\cancel{E}_T$ significance.
- $\sum E_T$: the sum of the transverse energy of all the jets in the event.
- $\sum_3 E_T$: the sum of the transverse energy of all the jets except the leading two jets in the event.
- Sphericity, Aplanarity: energy and topology related variables [37].

We use MC simulated $t\bar{t}$ samples for signal training sample. For background training sample, we use data sample that has at least four jets in each event post the preliminary selection, since this sample has low $t\bar{t}$ signal fraction. We apply an additional kinematic requirement on training samples before the neural network training: events with $\Delta\phi Min <$

1 0.4 are removed. Figure 4.12 and 4.13 show the comparison of the neural network input
2 variables for both the $t\bar{t}$ MC (signal) and multijet data (background) samples.

3 The neural network score output (NNout) for both the $t\bar{t}$ signal and background processes
4 will be shown in Section 5, based on which we will apply a stringent cut on NNout in order to
5 get a significant amount of $t\bar{t}$ events while suppressing a great amount of background rates.

Figure 4.12: Distribution of the neural network input variables for multijet data (Background) and $t\bar{t}$ MC (Signal) samples.

Figure 4.13: (Continued) Distribution of the neural network input variables for multijet data (Background) and $t\bar{t}$ MC (Signal) samples.

CHAPTER 5

BACKGROUNDS

We need to estimate how many background events are present in the candidate data sample we use for the measurement. Background estimation can be done either using MC simulated samples or data, or a combination of both. In the following, we will discuss the background estimation in lepton + jets channel, which uses a hybrid of data- and MC-based method, and in the $\cancel{E}_T$ + jets final state data sample, which takes a data-driven approach.

5.1 Method II for Lepton + Jets Channel

The primary sources of background processes in the lepton + jets channel are W + jets and QCD multijets production. We also consider contributions from Z + jets, diboson and single-top production. A "Method II" approach [38] is used to estimate the rates for each of the background processes. First, for the Z + jets, diboson, single top, and $t\bar{t}$ events we normalize MC simulation events using their respective theoretical cross sections [39, 40, 41]. Then QCD multijet background rate is estimated from data [42]. Finally for W + jets process, its shape is obtained from MC samples while the total number of W + jets events is calculated by subtracting all other background rates and $t\bar{t}$ from the data.

Background processes of WW , WZ , ZZ , single top production and $t\bar{t}$ produce lepton and neutrino, therefore mimicing the $t\bar{t}$ signal process in the lepton + jets channel. We estimate their rates using their theoretical cross section, the luminosity of data sample, trigger efficiency and an overall selection efficiency derived from MC simulation, and the equation used to calculate is:

$$N_{process}^i = \sigma_{theory} \times \epsilon_{evt} \times \epsilon_{tag} \times L, \quad (5.1)$$

where σ_{theory} is the theoretical cross section of the background process, ϵ_{evt} is the event selection efficiency derived from MC samples, ϵ_{tag} is the tagging efficiency, and L is the total

Table 5.1: The sources and expected numbers of background events in the lepton + jets channel, and the number of events observed for 1 b -tag and 2 b -tag samples after all selection requirements.

	1 b -tag	2 b -tag
$W + \text{jets}$	85.6 ± 21.8	9.8 ± 2.9
non- W	24.5 ± 20.6	2.4 ± 1.8
EWK	10.2 ± 0.8	2.4 ± 0.2
Total background	120.2 ± 30.0	14.6 ± 3.4
Observed events	542	214

luminosity used for this analysis. Here the efficiency term ϵ_{evt} includes trigger efficiencies and scale factors for accounting the data - MC lepton identification differences.

The QCD multijet production can also be a source of $t\bar{t}$ background because in QCD events a jet can mimic a lepton or some heavy quarks can decay semileptonically which produces leptons. To determine the fraction of QCD multijet rate, it is found that the $\cancel{E}_T$ spectrum of events coming from QCD production is well differentiable from that of $W + \text{jets}$ events [43]. This is because in QCD events the $\cancel{E}_T$ is due to experimental effects such as jets entering an uninstrumented area of the calorimeter or a muon is faked and its momentum is added during the calculation of $\cancel{E}_T$. We model QCD events by using data that pass all selection criteria where instead of a lepton, an electron-like (called "antielectron") object is required. We then use the shape of $\cancel{E}_T$ from QCD sample and $W + \text{jets}$ MC samples to fit to the $\cancel{E}_T$ shape in the data, thus obtaining the relative fraction of each of the two. The rate of QCD process is then simply calculated as the product of the fitted fraction and the total number of background events in the sample.

Finally, to get the rate for $W + \text{jets}$ production we subtract the contributions from the total number of observed events. The fractions of $W + \text{jets}$ events with heavy flavor quarks (Wc , $Wc\bar{c}$, and $Wb\bar{b}$ events) are determined from MC simulated samples. The rate with which events with a W boson and light flavor quarks contain a misidentified b -jet is determined using data samples triggered by the presence of jets. Table 5.1 summarizes the background compositions.

5.2 Method I for MET + Jets Final State Events

For the analysis of top mass measurement using events of $\cancel{E}_T + \text{jets}$ final state, the dominant background processes come from QCD multijet and $W + \text{jets}$ production. We use a data-driven technique called "Method I" to model the background processes. As discussed in Section 4.2.3, we require at least one b -tagged jet in a candidate event, so we need to estimate how many b -tagged events are in the data. The method used in this analysis is based on the assumption that the probabilities of tagging a b jet in $t\bar{t}$ signal and background processes are different. We first parameterize the b -tagging probabilities, on a *per jet* base, as a function of several jet characteristics using a data sample that is almost pure background processes. Then the probabilities for tagging jets of candidate data events can be calculated based on the parameterization, and the total rate of tagged background can be calculated as a sum of these probabilities, with necessary corrections applied, which will be detailed in the following paragraphs.

5.2.1 Background Parameterization

The parameterization of b -tagging rates as a function of some chosen jet characteristics will predict the number of b -tagged jets from background processes in any given selected sample. The following steps explain the procedure of b -tagging rate matrix determination:

- Select a subsample of data deplete of $t\bar{t}$ contamination.
- Based on the sample, parameterize the b -tagging rate as a function of N variables.
- Build a N -dimensional b -tagging matrix that determines the probability of a jet being b -tagged given the jet's characteristics.
- Predict the amount of expected number of background events by summing over the b -tagging probabilities of all the jets in the selected events.

We parameterize the b -tagging rate as a function of E_T , $\cancel{E}_T^{proj}$, and N_{trk} of a jet, where N_{trk} is the number of good quality SECVTX tracks of a jet, and $\cancel{E}_T^{proj}$ is defined as

$$\cancel{E}_T^{proj} = \cancel{E}_T \cos \Delta\phi(\cancel{E}_T, jet) \quad (5.2)$$

1 The b -tagging probability is defined as the ratio of the number of positive b -tagged jets to the
 2 number of b -taggable jets, where a b -taggable jet is a jet that has at least two good SECVTX
 3 tracks. Figure 5.1 shows the positive and negative b -tagging rate dependences of the three
 4 variables, based on which we will build the three-dimensional b -tagging matrix. The binning
 5 pattern of the b -tagging matrix is also shown in Fig. 5.1.

We then apply this b -tagging rate matrix to higher jet-multiplicities. While the sample we use to derive the b -tagging rate matrix has negligible $t\bar{t}$ contamination, events that have higher jet-multiplicities do not, since we expect about 5% – 10% $t\bar{t}$ contamination in the 4 - 6 jet bin of pretagged data sample. We solve this issue by removing $t\bar{t}$ components using an iterative correction procedure. If we denote N_{exp} as the uncorrected number of expected background events that is evaluated on a sample of N_{evt} events, and N_{obs} as the number of observed background events, then $(N_{obs} - N_{exp})/\epsilon_b$ would be the number of $t\bar{t}$ signal events present in the sample, where ϵ_b is the b -tagging efficiency. Then the following iterative procedure will make a better estimation of the background by multiplying N_{exp} by a *background ratio* at each step of the iteration:

$$N_{exp,i} = N_{exp,i-1} \cdot \frac{N_{evt} - \frac{N_{obs} - N_{exp,i-1}}{\epsilon_b}}{N_{evt}} \quad (5.3)$$

6 The iteration continues until the background estimates changes by less than 0.1 events.

Figure 5.1: Positive and negative b -tagging rates as a function of E_T , $\cancel{E}_T^{proj}$, and N_{trk} respectively for the data sample with exactly three tight jets in each event. The dotted vertical lines represent the binning pattern to be used in the b -tagging matrix.

5.2.2 Background Estimation

As mentioned above, we divide the event sample into two 1tag and 2tag events. An intuitive estimation of background rate can be calculated by tag rate matrix using the following Equation 5.4:

$$n_{exp}^{1tag} \approx \sum_{events} \left\{ \sum_i [p_i \cdot \prod_{j \neq i} (1 - p_j)] \right\} \quad (5.4)$$

$$n_{exp}^{2tag} \approx \sum_{events} \left\{ \sum_{ij} [p_i \cdot p_j \cdot \prod_{j \neq i, j} (1 - p_k)] \right\} \quad (5.5)$$

where the indices i , j , and k run over taggable jets in each event. However, this calculation will give inaccurate background rate because the b tag rate matrix is an inclusive tagging probability and does not take into account the fact that in QCD background real heavy flavor quarks come in pairs and therefore should enhance the two-tagging probability. The solution is to introduce correction factors to make a better estimate of 1-tag and 2-tag background events. We use the events in the background dominant region ($NN_{out} < 0.4$) and parameterize the correction factors as a function of the tagged jet E_T , which is shown in Equation 5.7:

$$C(E_T^1) \equiv n_{obs}^{1tag}(E_T^1) / n_{exp}^{1tag}(E_T^1) \quad (5.6)$$

$$C'(E_T^1, E_T^2) \equiv n_{obs}^{2tag}(E_T^1, E_T^2) / n_{exp}^{2tag}(E_T^1, E_T^2) \quad (5.7)$$

where E_T^i are the transverse energies of tagged jets, $n_{obs}(E_T^i, \dots)$ is the observed number of events with tagged jets of energy $E_T^i, \dots$ evaluated by Equation 5.4.

With these correction factors the new estimated tagged background rate can be rewritten

Table 5.2: Event composition of a MC sample after all cuts are applied.
CDF II Preliminary 8.7 fb⁻¹

Event Type	e+ jets	μ + jets	Had. τ +jets	Lep. τ +jets	DIL (e/ μ)	DIL (e/ μ + τ)	DIL (τ + τ)	All Had.
Composition(%)	28.43	18.72	29.89	13.55	1.22	3.10	1.26	3.82

as Equation 5.8:

$$n_{exp}^{1tag} \approx \sum_{events} \left\{ \sum_i [C(E_T^i) \cdot p_i \cdot \prod_{j \neq i} (1 - p_j)] \right\} \quad (5.8)$$

$$n_{exp}^{2tag} \approx \sum_{events} \left\{ \sum_{ij} [C'(E_T^i, E_T^j) \cdot p_i \cdot p_j \cdot \prod_{k \neq i, j} (1 - p_k)] \right\} \quad (5.9)$$

5.2.3 Candidate Events

We make a final cut on the candidate events: we require that the number of jets (n_{jets}) present in an event to be between $4 \leq n_{jets} \leq 6$, which rejects events with more than 6 jets. The reason is that the majority of the candidate events come from the $t\bar{t}$ semileptonic decay channel. As shown on Table 5.2, about 47% of the events contain an electron or muon in their final products, and 43% include a τ lepton in their decay which further decays to leptons or two jets through weak interaction. Therefore, ideally these candidate events have at most 6 jets in the final decay products. More jets could have been allowed in the candidate events, however due to the trigger type we use (multijet trigger) the QCD multijet process dominates the background of our data sample, it is therefore our intention to reduce the QCD background process by removing events with very large jet multiplicities.

We can know the event composition after all cuts are applied by using the HEPG (which contains sample information at the generator level) information of a MC sample. Table 5.2 shows the composition for each type of events from sample ttop25 (PYTHIA, $M_{top} = 172.5$ GeV/ c^2).

We calculate the estimated number of background events with the iterative correction

1 and b -tagging correction. Table 5.3 and Table 5.4 show the expected and observed events,
2 after all cuts (including $\text{NNout} > 0.9$ for 1tag events, and $\text{NNout} > 0.8$ for 2tag events) are
3 applied, for 1tag and 2tag events respectively. Table 5.5 shows the total number of both
4 expected and observed events within $4 \leq n_{\text{jets}} \leq 6$, which are the candidate events for the
5 top quark mass measurement.

Table 5.3: Expected and observed number of events after all cuts for 1tag events.

Jet Multiplicity	4 jets	5 jets	6 jets	≥ 7 jets
Signal	427.11 ± 50.31	548.44 ± 64.31	252.56 ± 29.59	89.71 ± 10.50
Background	262.90 ± 22.98	292.09 ± 25.51	158.30 ± 14.21	83.32 ± 3.53
Tot Expected	690.00 ± 55.31	840.53 ± 69.19	410.87 ± 32.83	173.03 ± 11.08
Observed	761	880	461	197

Table 5.4: Expected and observed number of events after all cuts for 2tag events.

Jet Multiplicity	4 jets	5 jets	6 jets	≥ 7 jets
Signal	179.00 ± 23.75	255.86 ± 33.94	117.66 ± 15.60	41.55 ± 5.50
Background	43.51 ± 11.36	73.32 ± 19.13	51.76 ± 13.53	36.27 ± 1.38
Tot Expected	222.51 ± 26.33	329.19 ± 38.96	169.42 ± 20.65	77.82 ± 5.67
Observed	225.00	359.00	191.00	98.00

5.2.4 Validation of Background Modeling

6 We also validate the background estimation method by comparing the distributions of kine-
7 matic variables of data with signal plus estimated background. Figure 5.2, 5.3, and Fig. 5.4,
8 5.5 show the comparison of some kinematic variables within the signal region, that is, with
9 all cuts applied (including NNout cuts and $4 \leq n_{\text{jets}} \leq 6$). for both 1btag and 2btag events
10

Table 5.5: Expected and Observed number of events after all cuts within $4 \leq n_{\text{jets}} \leq 6$.
CDF II Preliminary 8.7 fb^{-1}

b-tagging	jet-multiplicity	Signal	Background	Total Expected	Observed
1tag	4 jets	427.12 ± 50.31	262.90 ± 22.98	690.00 ± 55.31	761
	5 or 6 jets	801.01 ± 70.79	450.39 ± 29.20	1251.40 ± 76.58	1341
2tag	4 jets	179.00 ± 23.75	43.51 ± 11.36	222.51 ± 26.33	225
	5 or 6 jets	373.52 ± 37.35	125.09 ± 23.43	498.61 ± 44.09	550

₁ respectively.

Figure 5.2: Distribution of kinimatic variables for 1btag events.

Figure 5.3: (Continued) Distribution of kinimatic variables for 1btag events.

Figure 5.4: Distribution of kinimatic variables for 2btag events.

Figure 5.5: (Continued) Distribution of kinematic variables for 2btag events.

CHAPTER 6

EVENT RECONSTRUCTION AND TEMPLATES

After the event selection we obtain the candidate events to be used for measurements. The successing step is called event reconstruction, which uses kinematic information from an event to form estimators for either the top quark mass or decay width. We will reconstruct a top quark mass for each candidate event. The reconstructed top quark mass m_t^{reco} is obtained by a χ^2 function fitter for top decay width measurement, and is calculated as the invariant mass of a tri-jet system for the top mass measurement, which will be described later.

As discussed in Section 4.1.2, a series of complicated procedures are involed in correcting the jet energies to the parton level, and any systematic effects from the JES correction procedure will propagate to the systematic uncertainty on the top quark mass or decay width measurement. One effective way of reducing this effect is to introduce a dijet mass m_{jj} , representing the invariant mass of two jets from the hadronic decay of the W boson. The m_{jj} is very sensitive to the Δ_{JES} and is thus used to constrain the calibration of JES during the measurement.

In the top mass measurement, in order to extract more statistical information from an event we reconstruct a third kinematic variable, $m_t^{\text{reco}(2)}$, built in a similar way to m_t^{reco} . We only use two kinematic variables in the top decay width measurement.

For each event sample, m_t^{reco} and m_{jj} (and also $m_t^{\text{reco}(2)}$ in top mass measurement) are built to form a template for that sample. We will compare the data template with MC signal and background templates, and calculate a likelihood function that is used to extract the most probable top quark mass or decay width.

6.1 Top Quark Specific Corrections

Jet energy corrections in Section 4.1.2 are correction to general jets. However, differences exist between the spectra of jets from W decay and the b jets in the $t\bar{t}$ events, as well as between jets coming from b quarks and light-flavor quarks. To account for these effects, we derive corrections specifically to jets in the $t\bar{t}$ semileptonic decay events using MC simulations. Jets with different flavors are separated (whether they came from a b quark or a light quark), and after the relative, multiple interaction and absolute corrections are applied to jet energies they are grouped into different η bins. Jets in each η bin then have their own corrections depending on their p_T values. To obtain the correction, a response factor $F_{resp.}$ is defined for each jet in a certain η bin, as shown in Equation 6.1:

$$F_{resp.} = \frac{p_T^{parton}}{p_T^{jet}} - 1. \quad (6.1)$$

where p_T^{jet} is the p_T of a jet and p_T^{parton} is the p_T of the parton that produced that jet. The correction procedure then tries to parameterize the response factor $F_{resp.}$ as a function of jet p_T for each η bin, for heavy- and light-flavor jets respectively.

Once the response factor $F_{resp.}$ is found, the correction of the jet energy is then given by rearranging Equation 6.1:

$$p_T^{parton} = (F_{resp.} + 1) \times p_T^{jet}. \quad (6.2)$$

The correction can be as high as 50% for low- p_T jets. To calculate the energy of jets, we fix the mass of jets to be 5.0 GeV/ c^2 for those assumed to come from a b quark, and 0.5 GeV/ c^2 for those assumed to come from light-flavor quarks. This top-specific correction apply to both the top mass and decay width measurements.

6.2 The Lepton+Jets Channel

6.2.1 Reconstructed Top Mass m_t^{reco}

Inputs of the χ^2

The high- p_T lepton triggered events in the lepton + jets channel provide a overconstrained kinematics for the $t\bar{t}$ system. There are 6 decay products of the $t\bar{t}$ pair decay, giving 24 quantities that form the 4-vectors of the 6 products. We already know the 4-vectors of the four jets and the charged lepton, which give $4 \cdot 5 = 20$ constraints. We also know the $\cancel{E}_T$ for the events, which gives extra 2 constraints, $\cancel{E}_x$ and $\cancel{E}_y$. Plus, we assume that the neutrino has a mass of 0, bringing the number of constraints to 23. Besides, we can constrain the invariant mass of the two jets from the hadronic W decay to be the W boson mass, within the resolution and W width, and we can make the same constraint on the invariant mass of the lepton and neutrino. Finally, we require that the top quark mass be equal to that of the anti-top quark. With these constraints we create a χ^2 -like function fitter in Eqn. 6.3:

$$\begin{aligned} \chi^2 = & \sum_{i=l,4jets} \frac{(p_T^{i,fit} - p_T^{i,meas})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(p_j^{UE,fit} - p_T^{UE,meas})^2}{\sigma_j^2} \\ & + \frac{(M_{l\nu} - M_W)^2}{\Gamma_W^2} + \frac{(M_{jj} - M_W)^2}{\Gamma_W^2} \\ & + \frac{(M_{bl\nu} - m_t^{reco})^2}{\Gamma_t^2} + \frac{(M_{bjj} - m_t^{reco})^2}{\Gamma_t^2} \end{aligned} \quad (6.3)$$

where the first term constrains the p_t of the lepton and four leading jets in the event to their measured values within their assigned uncertainties σ_i ; the second term does the same for transverse components of unclustered energy (U), which is defined as the energy in the calorimeter not associated with the primary lepton or any of the leading four jets; the third and fourth terms constrain the invariant mass of products from the hadronically and

1 leptonically decaying W boson to be the well measured mass of W 80.42 GeV/ c^2 ; the last two
2 terms constrain the invariant masses of the three-body top decay products to be consistant
3 with the top quark mass, m_t^{reco} , within the top quark decay width which is assumed to be 1.5
4 GeV, where m_t^{reco} is the free parameter to be fitted for the reconstructed top quark mass.
5 The followings describe some inputs of Eqn. 6.3.

6 Jets and Leptons:

7

Here the top-specific corrections discussed in Section 6.1 are applied to jets, and the uncertainties on the jet energies also come from the correction. For a lepton, if it is an electron its energy is determined by the electromagnetic calorimeter and the angle of its 4-vector is determined by the associated track. Electron energies are also corrected for differences in the calorimeter response, and their masses are set to be zero. The resolution of the transverse momentum of an electron is calculated as:

$$\sigma_{p_T^e} = p_T^e \times \sqrt{\left(\frac{0.135}{p_T^e}\right)^2 + (0.02)^2} \quad (6.4)$$

If the lepton is a muon, then the 4-vector uses the momentum of the associated track, and its mass is also set to be zero. Track curvature corrections due to chamber misalignment are applied. The resolution of the transverse momentum of a muon is given as:

$$\sigma_{p_T^\mu} = 0.0011 \times (p_T^\mu)^2 \quad (6.5)$$

Unclustered Energy and Neutrinos:

The unclustered energy contains all of the transverse energy in the calorimeter (towers of $|\eta| < 3.6$) that is not associated with the lepton or any of the leading four jets. If towers are clusterd into a jet, which has $E_T > 8$ GeV and $|\eta| < 2.0$ but is not among the leading four jets, then the tower momenta are replaced with the jet momentum after the jet energy

correction discussed in Section 4.1.2. The momenta for the rest of the towers are multiplied by a factor of 1.4, since this corresponds to the average correction factor for generic jets with E_T of 8 GeV. Finally the energy that contributed to the leading four jets due to underlying events effect is added to the unclustered energy, and energies that were included in the leading four jets due to the OOC correction in Section 4.1.2 are subtracted from the unclustered energy. The last two steps avoid double-counting of energy that was included during the energy corrections of the leading four jets. The uncertainty on the unclustered energy in each event is given by:

$$\sigma^{UE} = 0.4 \times \sqrt{\sum E_T^{uncl}}. \quad (6.6)$$

1 where $\sum E_T^{uncl}$ is the sum of transverse energies in the event, excluding the energies of the
 2 lepton and leading four jets.

The transverse momentum of neutrino is calculated based on the unclustered energy, as well as the energies of the lepton and leading four jets of the event:

$$p_{x,y}^\nu = - \left(\sum p_{x,y}^{jet} + p_{x,y}^{lepton} + p_{x,y}^{UE} \right). \quad (6.7)$$

3 We fix the mass of the neutrino to be zero, and the longitudinal momentum of the
 4 neutrino p_z^ν is a free parameter in the χ^2 function fitter. The initial value of p_z^ν is calculated
 5 by assuming that the neutrino and lepton come from the W decay and thus the invariant
 6 mass of the neutrino and lepton should give the pole mass of W . Note that this generates
 7 a quadratic equation for p_z^ν and produces two solutions for the initial value of p_z^ν . The χ^2
 8 function fitter takes each of the two values and fit the reconstructed top mass m_t^{reco} .

9 Find m_t^{reco} from χ^2 fitting

10 For any given event, we have to decide which jets of the event should correspond to the jets
 11 in the χ^2 fitter, and this brings many possible combinatorics of jet-to-parton assignments.

For a 1tag event, its b -tagged jet should correspond to one of the two b -jets in the χ^2 fitter; two of its non- b -tagged jets should correspond to the W decay jets in χ^2 fitter, which gives 3 possible combinations; besides, there are two solutions for the longitudinal momentum of neutrino, therefore the total number of minimizations to do for a 1tag event is 12. For a 2tag event, this number is 4. Each of the minimizations is done for an event, and we choose m_t^{reco} of an event as the reconstructed mass that has the lowest minimized χ^2 value of all combinations, which we denote as χ_{min}^2 . A requirement of $\chi_{\text{min}}^2 < 9.0$ is imposed to remove those poorly reconstructed events, and the expected statistical uncertainty does not change much over a wide range of the choices of cut.

As described at the beginning of this chapter 6, the template method compares the shapes of observables of data with MC $t\bar{t}$ signal and background samples. Therefore it is essential to this method that different samples produce different shapes of m_t^{reco} . In another word, the distribution of m_t^{reco} must be sensitive to the input parameters of samples. To check if our reconstructed m_t^{reco} has this merit, we plot the distributions of m_t^{reco} from $t\bar{t}$ MC samples of different top quark decay widths on Figure 6.1. The figure does show that, as the input top quark decay width changes the shapes of m_t^{reco} also changes.

6.2.2 Dijet Mass w_{jj}

The dijet mass w_{jj} is reconstructed from the invariant mass of two jets that come from the hadronic W decay, and is very sensitive to Δ_{JES} . In Chapter. 7 we will describe the likelihood fit function that tries to measure the most probable top quark mass M_{top} , and the likelihood function fits both M_{top} and Δ_{JES} simultaneously. Fitting the extra parameter Δ_{JES} in the measurement significantly reduces the systematic effect brought about by JES, which is made possible by the sensitivity of w_{jj} to Δ_{JES} .

We reconstruct w_{jj} by calculating the invariant mass of two non- b -tagged jets of an event. Since there might be more than two non- b -tagged jets in an event, we select any combination of two non- b -tagged jets and choose the pair that produces an invariant mass that is closest

Figure 6.1: Distributions of m_t^{reco} from $t\bar{t}$ samples of input top widths Γ_{top} : 1.5 GeV, 10.0 GeV, and 20.0 GeV. All the $t\bar{t}$ MC samples have the same input top quark mass $M_{\text{top}} = 172.5 \text{ GeV}/c^2$. The top plot shows m_t^{reco} distribution for 1tag events, and the bottom plot shows it for 2tag events.

1 to the world measurement of W boson mass, $80.42 \text{ GeV}/c^2$ [44]. The calculation of w_{jj} is
 2 completely independent of the χ^2 function fitter of m_t^{reco} . For a 1tag event, there are 3
 3 choices of the two-jet pair since there are 3 jets that are non-btagged; for a 2tag events,
 4 there is only one possible choice to reconstruct w_{jj} . Figure 6.2 shows the w_{jj} distribution
 5 for the top width measurement, where different histograms come from $t\bar{t}$ MC samples with
 6 different input Δ_{JES} values.

Figure 6.2: Distributions of w_{jj} from $t\bar{t}$ samples of different input Δ_{JES} values: $-2.0 \sigma_c$, $0.0 \sigma_c$, and $+2.0 \sigma_c$. All the $t\bar{t}$ MC samples have the same input top quark width $\Gamma_{\text{top}} = 1.5 \text{ GeV}$. The top plot shows w_{jj} distribution for 1tag events, and the bottom plot shows it for 2tag events.

6.3 The MET+Jets Final State Events

The event signature for the top quark mass measurement has a large missing transverse energy and large jet-multiplicities, and based on the studies on $t\bar{t}$ MC samples we found that most of the selected events come from the $t\bar{t}$ semi-leptonic decay channel. The event reconstruction for events in the $\cancel{E}_T + \text{jets}$ final state has uses a similar method to reconstruct w_{jj} as events in the lepton + jets channel, but reconstructs the top quark mass quite differently.

6.3.1 Reconstructed Top Masses m_t^{reco} and $m_t^{\text{reco}(2)}$

Since the event signature of $\cancel{E}_T + \text{jets}$ events has merely multiple jets and a large $\cancel{E}_T$, the event reconstruction here is quite different from that in the top quark decay width measurement where the lepton information is well known. Instead of using a function fitter to reconstruct the top quark mass, we form the reconstructed top-quark mass (m_t^{reco}) from the invariant mass of the three jets whose momentum sum yields the largest p_T among all possible choices out of the leading four or five (if there are more than four jets in the event) jets. It is found that the invariant mass constructed this way has a large correlation with the hadronically decaying top-quark mass. To enhance the probability that the jets used to compute m_t^{reco} come from the hadronically decaying top quark, we add two constraints to the calculation of m_t^{reco} : first, m_t^{reco} must contain the two jets that contribute to w_{jj} , to be described in Section 6.3.2; second, in 2tag events, one jet of m_t^{reco} must be b -tagged. A third variable, $m_t^{\text{reco}(2)}$, is defined as the invariant mass of three jets, two of which are the pair which defines w_{jj} . The third jet of $m_t^{\text{reco}(2)}$ is required to be the most energetic jet of those not forming m_t^{reco} , and is required to be b -tagged in a 2-tag event. The variable $m_t^{\text{reco}(2)}$ provides complementary information on the top quark mass, and is particularly important in events where the three jets used to compute m_t^{reco} were not the actual decay products of the hadronically-decaying top quark.

To give a sense of how the resolution power of these observables behaves, Fig. 6.3 and

1 Fig. 6.4 show the distribution of reconstructed m_t^{reco} and $m_t^{\text{reco}(2)}$, with different input top
2 quark masses, while the input ΔJES is the same.

3 6.3.2 *Dijet Mass w_{jj}*

4 In a similar way to the w_{jj} reconstruction in the top quark decay width measurement, we
5 reconstruct the W boson mass from its hadronically decay jets to constrain *in situ* the JES.
6 For each possible jet-to-quark assignment in an event we calculate the invariant mass of two
7 non-btagged jets, then we choose the one that has the value closest to the world average W
8 mass measurement, $80.42 \text{ GeV}/c^2$, to be the reconstructed w_{jj} . Again, since we divide events
9 into 1tag and 2tag subcategories, the number of jet combinatorics are different depending
10 on if an event is 1tag or 2tag and on how many jets are present in the event. For example,
11 if an event contains 5 jets and has one jet tagged as coming from a b quark, then we can
12 choose any two jets out of the four non-btagged jets to reconstruct w_{jj} , giving a total number
13 of combinatorics of six, and the final reconstructed w_{jj} is the one in the six invariant masses
14 that is closest to the world W boson mass measurement.

Figure 6.3: Distribution of reconstructed m_t^{reco} with different input top quark masses, after all cuts are applied, both 1tag (top) and 2tag (bottom) are shown.

Figure 6.4: Distribution of reconstructed $m_t^{\text{reco}(2)}$ with different input top quark masses, after all cuts are applied, both 1tag (top) and 2tag (bottom) are shown.

Figure 6.5: Distribution of reconstructed m_{jj} with different input JES's, after all cuts are applied, both 1tag (top) and 2tag (bottom) are shown.

CHAPTER 7

THE LIKELIHOOD FIT

With the variables formed in Chapter 6, we can use their distributions from $t\bar{t}$ signal and background samples to build the probability density functions (p.d.f.s) for each sample, and construct the likelihood function based on them to simultaneously determine M_{top} and Δ_{JES} for the top quark mass measurement, or Γ_{top} and Δ_{JES} for the top quark decay width measurement. In the following, we will describe the likelihood function in Section 7.1. The p.d.f.s that the likelihood function is built upon will be discussed in Section 7.2. We will describe the Kernel Density Estimation (KDE) method used to estimate the p.d.f.s of the kinematic variables created in Section 6. KDE estimates the p.d.f.s for variables from samples of discrete values of Δ_{JES} and M_{top} or Δ_{JES} and Γ_{top} , and we will use local polynomial smoothing (LPS) to smooth and interpolate the p.d.f.s for arbitrary values of these parameters, described in Section 7.3.

7.1 The Likelihood Fit Function

The likelihood fit functions for the top quark width and mass measurements are very similar except that the former measures the top decay width and the later top mass. The form of the likelihood function is given by Eqn. 7.1:

$$\mathcal{L} = e^{-\frac{\Delta_{\text{JES}}^2}{2}} \times \prod_{\text{sample}} \mathcal{L}_{\text{sample}} \quad (7.1)$$

where the first term constrains the measurement of the Δ_{JES} to its nominal value of zero within the uncertainty of $1\sigma_c$; the second term $\prod_{\text{sample}} \mathcal{L}_{\text{sample}}$ is a product of $\mathcal{L}_{\text{sample}}$, to be described in Eqn. 7.4, for all subsamples of the events. For both top quark decay width and mass measurements there are two subsamples, 1tag and 2tag events, so the product has the

form of Eqn. 7.2:

$$\prod_{sample} \mathcal{L}_{sample} = \mathcal{L}_{1tag} \times \mathcal{L}_{2tag} \quad (7.2)$$

1 These $\mathcal{L}_{sample}$ are core to the measurement, and each of them has the following format:

$$\begin{aligned} \mathcal{L}_{sample} &= e^{\frac{(n_b - n_{b0})^2}{2\sigma_{n_{b0}}^2}} \\ &\times \frac{\exp(-(n_s + n_b))(n_s + n_b)^N}{N!} \\ &\times \prod_{i=1}^N \frac{n_s P_s(\vec{x}_i; \Theta, \Delta_{JES}) + n_b P_b(\vec{x}_i; \Delta_{JES})}{n_s + n_b} \end{aligned} \quad (7.3)$$

$$(7.4)$$

2 where the product runs over all events in a given subsample with N representing the
3 total number of events in that subsample. The observable $\vec{x}$ represent $(m_t^{\text{reco}}, w_{jj}, m_t^{\text{reco}(2)})$
4 for top mass measurement and $(m_t^{\text{reco}}, w_{jj})$ for top width measurement; Θ represents M_{top}
5 for the top quark mass measurement, and Γ_{top} for the top quark decay width measurement.
6 The quantities P_s and P_b designate the signal and background p.d.f.s determined by KDE
7 and LPS, to be described in Section 7.2 and 7.3 respectively. n_s represent the expected
8 number of signal events, and n_b the expected number of background events. The first term
9 in Eqn. 7.4 is a constraint to the expected number of background events, where n_{b0} is the *a*
10 *priori* estimate for the expected number of background events obtained from the background
11 rate estimation and $\sigma_{n_{b0}}$ is the uncertainty on the estimate, both of which come from the
12 background rate estimation in Section 5. The second term in Eqn. 7.4 is present because
13 the number of signal and background events follow the Poisson statistics. The last term is
14 crucial to the measurement since it contains the information of each event in the subsample
15 by applying the p.d.f.s of signal and background to the calculated observables $\vec{x}$ of each event.
16 The following Section 7.1 demonstrate how Eqn. 7.4 is derived.

1 Derivation of the Likelihood Fit Function

2 To make the derivation of the likelihood fit function easier to understand, we start by some
3 definitions to be used in the derivation:

$D \equiv$ Observed data.

$n_s \equiv$ Expected number of signal events.

$n_b \equiv$ Expected number of background events.

$N_s \equiv$ Actual number of signal events in data D.

$N_b \equiv$ Actual number of background events in data D.

4 Here D can represent the vector observable $\vec{x}$ in Eqn. 7.4. As in Eqn. 7.4 let us use Θ to
5 denote Γ_{top} in the top width measurement and M_{top} in the top mass measurement.

N_s and N_b are unknown and they must satisfy:

$$N_s + N_b = N(\text{the total number of events in the subsample.})$$

6 We define the likelihood $\mathcal{L}_{\text{sample}}$ of a subsample as the probability of seeing the observed
7 data D given the values of n_s , n_b , Θ and Δ_{JES} . Since N_s and N_b are unknown, the likelihood
8 can be written as

$$\begin{aligned} \mathcal{L}_{\text{sample}} &= \sum_{N_s=0}^N p(D, N_s, N_b | n_s, n_b, \Theta, \Delta_{\text{JES}}) \\ &= p(D | N_s, N_b, n_s, n_b, \Theta, \Delta_{\text{JES}}) p(N_s, N_b | n_s, n_b, \Theta, \Delta_{\text{JES}}) \end{aligned} \quad (7.5)$$

9 where we have used the conditional probability equation in the probability theory.

Since the signal and background processes are independent and both should follow the Poisson statistics, the term $p(N_s, N_b | n_s, n_b, \Theta, \Delta_{\text{JES}})$ can be written as the product of two

Poisson forms:

$$p(N_s, N_b | n_s, n_b, \Theta, \Delta_{\text{JES}}) = \text{Pois}(N_s, n_s) \cdot \text{Pois}(N_b, n_b) \quad (7.6)$$

where

$$\text{Pois}(N, n) = \frac{e^{-n} n^N}{N!}$$

- 1 is the probability of observing N events for a Poisson distribution with mean n .
- 2 To find out what $p(D | N_s, N_b, n_s, n_b, \Theta, \Delta_{\text{JES}})$ is, we see that this term means the prob-
- 3 ability of seeing the observed data given that there are N_s signal events and N_b background
- 4 events in the data, not knowing which N_s of them are signal and which N_b are background.
- 5 We can decompose the situation into all possible ways of picking N_s signal events out of the
- 6 N data events in D . Again, let us first define some notations:

$w \equiv$ An element of the set of all 2^N ways to divide the N events
in D into signal and background events.

$w_{N_s} \equiv$ An element of the set of all $\binom{N}{N_s}$ ways of picking
 N_s signal events from the N events in D .

$S(w) \equiv$ the set of signal events in combination w .

$B(w) \equiv$ the set of background events in combination w .

$N_s(w) \equiv$ the number of signal events in combination w .

$B_b(w) \equiv$ the number of background events in combination w .

1 Then

$$\begin{aligned}
p(D|N_s, N_b, n_s, n_b, \Theta, \Delta_{\text{JES}}) &= p(D|N_s, N_b, \Theta, \Delta_{\text{JES}}) \\
&= \sum_{w_{N_s}} p(D|w_{N_s}, N_s, N_b, \Theta, \Delta_{\text{JES}}) p(w_{N_s}|N_s, N_b, \Theta, \Delta_{\text{JES}}) \\
&= \sum_{w_{N_s}} p(D|w_{N_s}, \Theta, \Delta_{\text{JES}}) p(w_{N_s}|N_s, N_b) \tag{7.7}
\end{aligned}$$

The third equal sign holds because w_{N_s} already indicates that there are N_s signal events in data D , and w_{N_s} is independent of the parameters Θ and Δ_{JES} . It is trivial to see that w_{N_s} follows a Binomial distribution and thus

$$p(w_{N_s}|N_s, N_b) = \frac{1}{\binom{N}{N_s}}$$

Combining Equations 7.5, 7.6, and ?? we have

$$\mathcal{L}_{\text{sample}} = \sum_w p(D|w, \Theta, \Delta_{\text{JES}}) \frac{1}{\binom{N}{N_s(w)}} \text{Pois}(N_s(w), n_s) \text{Pois}(N_b(w), n_b) \tag{7.8}$$

Finally, $p(D|w, \Theta, \Delta_{\text{JES}})$ denotes the probability of seeing data D given parameters Θ , Δ_{JES} , and W , which specifies whether an event in D is a signal or background for every event in D . Therefore we can write $p(D|w, \Theta, \Delta_{\text{JES}})$ as

$$p(D|w, \Theta, \Delta_{\text{JES}}) = \left(\prod_{i \in S(w)} f_s(D_i, \Theta, \Delta_{\text{JES}}) \right) \left(\prod_{j \in B(w)} f_b(D_j, \Delta_{\text{JES}}) \right) \tag{7.9}$$

2 where f_s denotes the probability for an event to be a signal and f_b the probability for an
3 event to be a background. The first term on the right of Eqn. 7.9 runs over all events that
4 are signals, and the second term runs over all events that are backgrounds.

The likelihood can now be written as

$$\begin{aligned} \mathcal{L}_{sample} = & \sum_w \left(\prod_{i \in S(w)} f_s(D_i, \Theta, \Delta_{\text{JES}}) \right) \left(\prod_{j \in B(w)} f_b(D_j, \Delta_{\text{JES}}) \right) \\ & \cdot \frac{1}{\binom{N}{N_s(w)}} \text{Pois}(N_s(w), n_s) \text{Pois}(N_b(w), n_b) \end{aligned} \quad (7.10)$$

By expanding the Poisson and Binomial equations and some simple combination of terms in Equation 7.10, the likelihood can be rewritten as

$$\begin{aligned} \mathcal{L}_{sample} &= \frac{\text{Pois}(N, n_s + n_b)}{(n_s + n_b)^N} \sum_w \left(\prod_{i \in S(w)} f_s(D_i, \Theta, \Delta_{\text{JES}}) \right) \left(\prod_{j \in B(w)} f_b(D_j, \Delta_{\text{JES}}) \right) \\ &= \text{Pois}(N, n_s + n_b) \prod_{i=1}^N \frac{n_s f_s(D_i, \Theta, \Delta_{\text{JES}}) + n_b f_b(D_i, \Delta_{\text{JES}})}{n_s + n_b} \end{aligned} \quad (7.11)$$

The fact that the second equal sign in Eqn. 7.11 holds can be seen by expanding a product of N terms of simple sum, and rewriting it as a sum of 2^N terms, each of which is a product of N simple terms (f_s and f_b).

This proves Eqn. 7.4.

7.2 Kernel Density Estimation

7.2.1 An Introduction

To use Eqn. 7.11 we need the expression for the p.d.f.s of f_s and f_b , which evaluate the probability for an event to be a $t\bar{t}$ signal and background respectively. For simplicity we will first discuss how to evaluate f_s and f_b in the one-dimensional case, and then extend to the multi-dimensional cases.

A p.d.f $f(x)$ can be estimated in a parametric way (here $f(x)$ is the p.d.f of some reconstructed observables, in our case m_t^{reco} and w_{jj} etc., for a given sample), by assuming a functional form of the true density $f(x)$. We can fit the function to the distribution of x

1 and find out the parameters of the function. However, the main disadvantage of the para-
2 metric approach is that it lacks flexibility, that is, each parametric family of distributions
3 imposes restrictions on the forms that $f(x)$ can have. For instance, the Gaussian function
4 requires the distribution to be bell-shapes and symmetrical, which is not guaranteed for the
5 variables we use in the analysis. Secondly, parametric method may fail to capture some of
6 the features of the distributions. Thus we can think of non-parametric methods to estimate
7 density functions. One simple but useful way is to use histograms, however histograms are
8 lack of continuity and this disadvantage is severe in the region where there are not enough
9 statistics, making the density estimation not accurate.

A technique called Kernel Density Estimation (KDE) was developed in the 1950s. KDE takes uses of two objects, the kernel function and the smoothing parameter (or called bandwidth). The kernel function is choosen carefully such that it satisfies some basic requirements, which is to be described later. KDE does not try to fit the kernel functions to the data distrubtion. Instead, it assigns a kernel function to each data point, and the estimated density function is the sum of kernel functions of all the data points. The kernel function of each data point has a bandwidth, which determines how wide the kernel function should expand for this particular data point. Therefore data points may or may not have the same bandwidth. A typical one-dimentional kernel density estimator looks like Eqn. 7.12:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) = \frac{1}{n} \sum_{i=1}^n K_h(x, X_i) \quad (7.12)$$

where $K(x)$ and $K_h(x)$ are the kernel functions, n is the number of data events, i runs over all data points, and h is the bandwidth (here we are discussing the simplest case where all data points have the same bandwidth).

$$K_h(x, X_i) = \frac{1}{h} K\left(\frac{x - X_i}{h}\right) \quad (7.13)$$

10 Choices of kernel functions are based on the MISE analysis, which is short for Mean

1 Integrated Squared Error analysis. Much of the follow derivations are based on [45].

We first calculate the bias of the $\hat{f}(x)$:

$$Bias\hat{f}(x) = E[\hat{f}(x)] - f(x)$$

In Eqn. 7.17, X_i is any of the random variables from data points, thus

$$E[\hat{f}(x)] = EK_h(x, X), \text{ and } Var\hat{f}(x) = \frac{1}{n}VarK_h(x, X)$$

2 To calculate the expectation,

$$\begin{aligned} EK_h(x, X) &= \int \frac{1}{h} K\left(\frac{x-t}{h}\right) f(t) dt = \int K(w) f(x-hw) dw \\ &= f(x) \int K(w) dw - hf'(x) \int wK(w) dw + \frac{1}{2}h^2 f''(x) \int w^2 K(w) dw \end{aligned} \quad (7.14)$$

$$(7.15)$$

To calculate the variance of $K_h(x, X)$, we need to calculate

$$E\left[\frac{1}{h}K\left(\frac{x-X}{h}\right)\right]^2 = \int \frac{1}{h^2} K^2\left(\frac{x-t}{h}\right) f(t) dt = \int \frac{1}{h} K^2(w) f(x-hw) dw \approx \frac{f(x)R(K)}{h} \quad (7.16)$$

For Eqn. 7.15, if the kernel K satisfies

$$\int K(w) = 1, \int wK(w)dw = 0, \int w^2 K(w)dw \equiv \sigma_K^2 > 0 \quad (7.17)$$

where σ_K^2 is the variance of K , then the bias of $\hat{f}(x)$ will be on the order of $O(h^2)$:

$$Bias\hat{f}(x) = \frac{1}{2}\sigma_K^2 h^2 f''(x) + O(h^4) \implies ISB = \frac{1}{4}\sigma_K^4 h^4 R(f'') + O(h^6)$$

From Eqn. 7.15 and 7.16, the variance can be calculated as

$$Var \hat{f}(x) = \frac{f(x)R(K)}{nh} - \frac{f(x)^2}{n} + O\left(\frac{h}{n}\right)$$

Thus the integrated variance is

$$IV = \frac{R(K)}{nh} - \frac{R(f)}{n} \quad (7.18)$$

The Asymptotic Mean Integrated Square Error (AMISE) is defined as the sum of ISB and IV at the leading order. We ignore the second term of the right hand side of Eqn. 7.18 since it does not depend on the bandwidth or kernel function, thus the AMISE is

$$AMISE = \frac{R(K)}{nh} + \frac{1}{4}\sigma_K^4 h^4 R(f'') \quad (7.19)$$

Equation 7.19 tells how the AMISE behaves as the bandwidth changes: for very small values of h , the IV term dominates the AMISE, while for large values of h the ISB term dominates. We can find a bandwidth h^* that minimizes Eqn. 7.19:

$$h^* = \left[\frac{R(K)}{n\sigma_K^4 R(f'')} \right]^{1/5} \quad (7.20)$$

With the choice of h^* for the bandwidth, the corresponding AMISE is given by

$$AMISE^* = \frac{5}{4}[\sigma_K^4 R(K)]^{4/5} R(f'')^{1/5} n^{-4/5} \quad (7.21)$$

Different choices of kernels give different values for Eqn. 7.21. In the analysis for both the top quark mass and decay width measurement we use the Emanechnikow kernel, defined in Eqn. 7.22:

$$K(t) = \begin{cases} \frac{3}{4}(1 - t^2), & |t| < 1 \\ 0, & \text{Otherwise} \end{cases} \quad (7.22)$$

1 which gives $R(K) = 3/5$ and $\sigma_K^2 = 1/5$.

Equation 7.20 involves the real density function $f(x)$. In practice, function $f(x)$ is not known, thus making Eqn .7.20 hard to implement in practice. However, a method called *oversmoothing* [45] gives an upper limit for the optimal choice of bandwidth, which is defined as

$$h_{OS} = 3 \left[\frac{R(K)}{35n\sigma_K^4} \right] \sigma_f$$

where σ_f is the standard variance of the distribution $f(x)$. To use oversmoothing in practice, we use the standard variance of the data x to replace σ_f . h_{OS} is used as the first step to approximate the density function. In the following paragraph, we will talk about how to determine the bandwidth for each data point. It is not necessary to use the same bandwidth for each data point in KDE, and thus the density function estimator can be written as

$$\hat{f}(x) = \sum_{i=1}^n \frac{n}{h_i} K \left(\frac{x - X_i}{h_i} \right)$$

where h_i is the bandwidth that is determined for each data point. This way the kernel function is called a adaptive kernel. For the region where the data distribution is very dense, h_i should be small enough to capture the details of the distribution. But for region where the distribution is sparse, h_i should be large enough not to catch the information caused by statistical fluctuation. There has been much effort contributed to search for the optimal choice for h_i . Here we choose the square root law by [46]: $h_i \propto \frac{1}{\sqrt{f(x_i)}}$. Here the value $f(x_i)$ is unknown since the true density function $f(x)$ is usually not a priori, and we estimate $f(x_i)$ by using a pilot density function $\hat{f}_{pilot}(x)$, which is a p.d.f obtained from KDE by using a constant bandwidth h_{OS} for every data point. We follow [47] and use

$$h_i = h_{OS} \sqrt{\frac{G}{\hat{f}_{adapt}(x_i)}} \quad (7.23)$$

2 where G is the geometric mean of all the $\hat{f}(x_i)$ of data points.

When using Eqn. 7.23, h_i could become very large at the region where $\hat{f}_{adapt}(x_i)$ is small, thus points associated with these h_i 's can contribute much density estimation to other points that are very distant, which is not desired and should be avoided. We follow [47] and use

$$\hat{f}_{adapt}(x_i) \rightarrow \max(\hat{f}_{adapt}(x_i), 0.1 \cdot \hat{f}_{adapt}^{max}(x)),$$

1 where $\hat{f}_{adapt}^{max}(x)$ is the maximum density estimate of $\hat{f}_{adapt}(x_i)$. This has the effect that the
2 maximum value of h is $\sqrt{10}$ times the value of the smallest h

3 Now that we have chosen our kernel and the rules to calculate bandwidth for each data
4 point, we are ready to show how KDE is implemented in the analysis. Figure 7.1 shows
5 the one-dimensional histograms and KDE-estimated p.d.f.s for the three kinematic variables
6 used in the top quark mass measurement, m_t^{reco} , w_{jj} , and $m_t^{\text{reco}(2)}$. Figure 7.2 shows the
7 same type of plots for the background sample used in the top quark mass analysis.

8 7.2.2 Multi-Dimentional KDE

The general multivariate kernel estimator is [48]

$$\hat{f}(\mathbf{x}) = \frac{1}{n|\mathbf{H}|^{1/2}} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{X}_i}{\sqrt{\mathbf{H}}}\right)$$

In practice, however, productive kernels are used due to their simplicity. A d -dimentional productive kernel is defined as

$$\hat{f}(\mathbf{x}) = \frac{1}{nh_1 \cdots h_d} \sum_{i=1}^n \left\{ \prod_{j=1}^d \mathbf{K}\left(\frac{\mathbf{x}_{ij} - \mathbf{X}_{ij}}{h_j}\right) \right\}$$

9 where the same (univariate) kernel is used for each dimention but with different smoothing
10 parameters. For the top quark width measurement we use two-dimentional KDE ($d = 2$)
11 since there are two kinematic variables, m_t^{reco} and w_{jj} ; for the top quark mass analysis, we

Figure 7.1: One dimensional signal p.d.f.s estimated by KDE for the three variables used in the analysis, $m_t^{\text{reco}}, w_{\text{jj}}$, and $m_t^{\text{reco}(2)}$, overlaped with their histogram distributions. The events here comes from a sample with $M_{\text{top}} = 172.0 \text{ GeV}/c^2$ and $\Delta_{\text{JES}} = 0.0$. The left column is for 1btag events and right column is for 2btag events.

Figure 7.2: One dimensional background p.d.f.s estimated by KDE for the three variables used in the analysis, m_t^{reco} , w_{jj} , and $m_t^{\text{reco}(2)}$, overlapped with their histogram distributions. The events here comes from the background sample. The left column is for 1btag events and right column is for 2btag events.

1 use three-dimensional ($d = 3$) since we use a third variable, $m_t^{\text{reco}(2)}$.

Reference [48] shows that for d -dimensional KDE, the oversmoothing rule is

$$h_{OS}^d = \left[\frac{d \cdot R(K)^d}{nC_f \sigma_K^4} \right]^{1/(d+4)} \sigma_f$$

where

$$C_f = \frac{16\Gamma(\frac{d+8}{2})d(d+2)}{\pi^{d/2}(d+8)^{(d+6)/2}}$$

2 Two-Dimensional KDE in Top Width Measurement

3 The top quark decay width uses two-dimensional KDE to estimate the p.d.f.s. Figure 7.3
 4 show the 2-dimensional signal probability density functions for lepton + jets channels at
 5 $\Gamma_{\text{top}} = 1.5 \text{ GeV}$ and $\Delta_{\text{JES}} = 0\sigma_c$, and Fig. 7.4 shows background probability density function
 6 at $\Delta_{\text{JES}} = 0\sigma_c$.

7 Three-Dimensional KDE in Top Mass Measurement

8 For the top quark mass measurement, we use three-dimensional KDE to estimate the p.d.f.s.
 9 Since it is hard to the three dimensional plots of p.d.f, we only show the two-dimensional
 10 projections for each pair of the three kinematic variables. Figure 7.5 shows the p.d.f.s for
 11 each pair of the variables of $t\bar{t}$ signal sample with $M_{\text{top}} = 172.0 \text{ GeV}/c^2$ and $\Delta_{\text{JES}} = 0.0$.

12 Figure 7.6 shows the p.d.f.s for background sample.

13 7.3 Local Polynomial Smoothing

14 The families of probability density functions $P_s(\vec{x}; M_{\text{top}}, \Delta_{\text{JES}})$ (or $P_s(\vec{x}; \Gamma_{\text{top}}, \Delta_{\text{JES}})$) and
 15 $P_b(\vec{x}; \Delta_{\text{JES}}$ should be continuous on the parameters to be fitted, M_{top} and Δ_{JES} . However,
 16 we only have limited number of $t\bar{t}$ signal MC samples and background samples. In fact, for
 17 the top quark mass measurement, we have 76 $t\bar{t}$ MC samples generated by PYTHIA, with the

Figure 7.3: 2D probability density function of signals from KDE ($\Gamma_{\text{top}} = 1.5$ GeV and $\Delta_{\text{JES}} = 0$). Top plot, 1-btag; bottom plot, 2-btag.

Figure 7.4: 2D probability density function of background from KDE ($\Gamma_{\text{top}} = 1.5 \text{ GeV}$ and $\Delta_{\text{JES}} = 0$). Top plot, 1-btag; bottom plot, 2-btag.

Figure 7.5: Two dimensional signal p.d.f.s estimated by KDE for each pair of the variables: m_t^{reco} vs W_{jj} (top), m_t^{reco} vs $m_t^{\text{reco}(2)}$ (middle) and W_{jj} vs $m_t^{\text{reco}(2)}$ (bottom), for 1btag (left) and 2btag (right) events respectively. Events here comes from a signal MC sample with $M_{\text{top}} = 172.0 \text{ GeV}/c^2$ and $\Delta_{\text{JES}} = 0.0$.

Figure 7.6: Two dimensional background p.d.f.s estimated by KDE for each pair of the variables: m_t^{reco} vs W_{jj} (top), m_t^{reco} vs $m_t^{reco(2)}$ (middle) and W_{jj} vs $m_t^{reco(2)}$ (bottom), for 1btag (left) and 2btag (right) events respectively. Events here comes from the background sample, which is modeled from data.

¹ M_{top} ranging from 150 to 240 GeV/ c^2 , and each M_{top} is associated 29 correlated MC samples
² with different Δ_{JES} values, ranging from $-3\sigma_c$ to $+3\sigma_c$. As for background, there is just one
³ sample since it is modeled from the data. For the top quark decay width measurement,
⁴ there are 16 MC samples with Γ_{top} ranging from 0.1 GeV to 30 GeV, each associated with
⁵ 29 different Δ_{JES} values, as in the top mass measurement. Therefore we only have limited
⁶ amount of density functions on the parameter grid of $(M_{\text{top}}, \Delta_{\text{JES}})$ or $(\Gamma_{\text{top}}, \Delta_{\text{JES}})$. We
⁷ need to extend the family of density functions so that it is continuous on the parameter grid,
⁸ and we do so by employing local polynomial smoothing (LPS) [50]. We briefly describe this
⁹ technique and the followings are excerpt from Reference [49].

¹⁰ The idea of LPS is to approximate the value of the p.d.f at a point with a second-order
¹¹ polynomial. Suppose we want to estimate the p.d.f $P(\mathbf{x}; \theta)$ for any parameter θ , where $\mathbf{x}$
¹² is the random variable vector, and could represent $m_t^{\text{reco}, w_{\text{jj}}}$, and $m_t^{\text{reco}(2)}$ in our analysis,
¹³ and θ , in the case of $t\bar{t}$ signal samples, could represent the parameter vector of $(M_{\text{top}}, \Delta_{\text{JES}})$
¹⁴ in the top mass analysis, or $(\Gamma_{\text{top}}, \Delta_{\text{JES}})$ in the top decay width analysis, and in the case
¹⁵ of background sample, just represent Δ_{JES} . For convinence, we omit the notation of $\mathbf{x}$ and
¹⁶ denote $P(\theta)$ as the p.d.f. We are already given the density estimation Y_k from KDE, where
¹⁷ k runs over the discrete parameter grid of θ , and we will use these Y_k 's to expand the density
¹⁸ estimation to cover the continuous space of θ .

For a two-dimentional vector ν , we first denote a functional vector basis $\mathbf{F}(\nu)$, defined as

$$\mathbf{F}(\nu) = \begin{pmatrix} 1 \\ \nu_1 \\ \nu_2 \\ \nu_1\nu_2 \\ \frac{1}{2}\nu_1^2 \\ \frac{1}{2}\nu_2^2 \end{pmatrix}$$

To estimate the density function at any value of θ , we approximate it as a linear combina-

tion of the components of $\mathbf{F}(\theta - \boldsymbol{\Theta}_k)$, where k runs over all the discrete parameter grids that we already have. We determine the coefficients of the linear combinations by minimizing the following equation:

$$\sum_k \omega(\boldsymbol{\Theta}_k) (Y_k - \mathbf{c} \cdot \mathbf{F}(\theta - \boldsymbol{\Theta}_k))^2$$

where $\mathbf{c}$ represent the vector of coefficients to be determined, and $\omega(\theta_k) = W(u_k)$, with

$$u_k = \sqrt{\sum_{d=1}^{N_d} \left(\frac{\boldsymbol{\Theta}_{k,d} - \theta_d}{h_d} \right)^2}$$

and

$$W(u) = \begin{cases} (1 - |u|^3)^3 & \text{for } |u| < 1, \\ 0 & \text{Otherwise.} \end{cases}$$

1 The weight $\omega(\boldsymbol{\Theta}_k)$ is present to ensure that points near the evaluation point θ contribute
2 more to the estimation and points far away from θ contribute little. The constant h_d is
3 a smoothing parameter, and large values of h_d give more weight to points distant from θ .
4 The values of h_d in our analysis are selected such the the measurement if unbiased and the
5 evaluated corresponding statistical uncertainties are correct.

CHAPTER 8

STATISTICAL RESULTS

8.1 Extracting the Limits on Top Decay Width - A Feldman-Cousins Method

We use a Feldman-Cousins approach to set limits on the top quark decay width. The key feature in constructing confidence intervals using Feldman-Cousins scheme is to define the *ordering principle*. In the original paper of Feldman-Cousins [?], an ordering principle is defined per Pseudo-Experiment as

$$R \equiv \Delta\chi^2 = \chi^2(\Gamma_{Input}) - \chi^2(\Gamma_{BestFit}) \quad (8.1)$$

where Γ_{Input} is the input top width of the MC sample, and $\Gamma_{BestFit}$ is the measured top width of a Pseudo Experiment; $\chi^2(\Gamma_{Input})$ is the χ^2 value at input top width, while $\chi^2(\Gamma_{BestFit})$ is the χ^2 value at the best fit top width of this single Pseudo-Experiment. We use the minimization of negative log-likelihood fit to get the best fit value of measured top width, but suppose the likelihood function to be in Gaussian regime, $-2\log(\text{likelihood})$ should follow the χ^2 distribution. Therefore in Equation 8.1 we simply use

$$\chi^2 = -2\log L \quad (8.2)$$

where L is the likelihood function. There is a $\Delta\chi^2$ value for each Pseudo-Experiment. For a MC sample we run thousands of Pseudo-Experiments therefore we have a distribution of $\Delta\chi^2$ for this sample. Figure 8.1 shows the $\Delta\chi^2$ distribution of sample `twid50` ($\Gamma_{input} = 5.0$ GeV, $\Delta JES = 0$). With this distribution we need to find a critical $\Delta\chi^2$ value $\Delta\chi_c^2$ so that the interval $[0, \Delta\chi_c^2]$ covers 95% of the events. If the distribution is really a χ^2 distribution one would naively expect $\Delta\chi_c^2$ to be 3.84, as calculated from statistics. However, in reality $\Delta\chi_c^2$

could deviate from 3.84 for several reasons: a) Physical boundary effects; b) the deviation of likelihood function from Gaussian regime.

Figure 8.1: $\Delta\chi^2$ distribution of sample twid50 ($\Gamma_{input} = 5.0$ GeV).

Each MC sample will have a $\Delta\chi_c^2$ value. Now that we have 21 MC samples, we can draw a plot of $\Delta\chi_c^2$ vs input top width. Figure 8.2 shows this plot, with no systematic effects considered. We will need this plot together with the data fit to get limit(s) for top width.

From data fit, which is done by minimizing $\chi^2 = -2\log(\text{likelihood})$, we get a plot in Fig. 8.3. In order to extract the limit(s) of top width we overlap plots of Fig. 8.2 and Fig. 8.3, and the point(s) at intersection of these two plots will give the limit(s) of top quark width, as shown in Fig. 8.4, which gives an upper limit of top width $\Gamma_{top} < 6.63$ GeV for pure statistics.

Figure 8.2: $\Delta\chi^2_c$ vs input top width.

Figure 8.3: Data Fit of top width.

8.2 Top Mass Measurement

8.2.1 Bias Check

It is possible that our method has biases to the measured top quark mass. To estimate the bias, we run pseudo experiments (PE) on signal samples with different input M_{top} and Δ_{JES} . Since the background sample is estimated from real data events, we use the same background events for different MC signal samples when running PE's. The number of background events used to run PE is the number we obtain from background estimation, while the number of signal events drawn to run PE is calculated from the theoretical calculation with cross section of $\sigma = 7.4$ pb, at $M_{\text{top}} = 172.5$ GeV/ c^2 . Uncertainties on the measured top quark mass in the biascheck come from the bootstrap method [51].

Figure 8.5 shows that our method biases the top mass to a certain degree. The figure tells us that we can use a linear function to fit the measured top mass vs the input top mass, therefore we apply this linear fit as a calibration to our measured top quark mass from now on. The figure also shows the pullwidth plot of measured top mass, in which the pullwidth is defined, for each ensemble of PE's for a MC sample, as the sigma of the Gaussian fit of the mass pull distribution.

We apply the linear calibration to our measured top quark mass. The distributions of the average mass residual (the deviation from the input top-quark mass) and the width of the pull (the ratio of the residual to the uncertainty) for simulated experiments are corrected to be zero and unity, respectively. The correction is $M_t^{\text{corr}} = 1.24 \times M_t^{\text{meas}} - 40.6$ GeV/ c^2 , where M_t^{meas} is the raw value from the likelihood fit and M_t^{corr} is the corrected value of the measurement. The measured uncertainty is correspondingly increased by 20% to correct the width of the pull distribution. Figure 8.6 shows the biascheck and pullwidth after calibration.

We then inspect the measured top mass dependencies on Δ_{JES} . Three top mass MC samples are chosen: $M_{\text{top}} = 167.5$ GeV/ c^2 , 172.0 GeV/ c^2 , and 177.5 GeV/ c^2 , each with five

1 different Δ_{JES} : 0.0, ± 0.4 , ± 1.0 . By applying the same calibration described above, we get
2 the plots shown on Fig. 8.7, which tell us that our measured top mass after calibration has
3 negligible dependency on Δ_{JES} . Figure 8.9 shows the expected statistical uncertainties of
4 the measured top quark mass after the calibration.

5 Figure 8.8, which presents the biascheck of measured Δ_{JES} vs input Δ_{JES} , also validates
6 our measurement of Δ_{JES} . Figure 8.9 shows the expected statistical uncertainties of the
7 measured top quark mass after the calibration.

Figure 8.4: Overlap of Fig. 8.2 and Fig. 8.3, in which the intercept(s) gives the top width limit(s).

Figure 8.5: Biaschecks of the measured top quark mass before any correction/calibration. Residual plot is on the left, and the pullwidth plot on the right. All the samples have the same $\Delta_{JES} = 0.0$.

Figure 8.6: Biaschecks of the measured top quark mass after calibrations. Residual plot is on the left, and the pullwidth plot on the right. All the samples have the same $\Delta_{\text{JES}} = 0.0$.

Figure 8.7: Checks of the measured top mass dependencies on the Δ_{JES} . Plots shown here are after calibration.

Figure 8.8: Biaschecks of the measured Δ_{JES} vs the input Δ_{JES} . Different color represents different input top masses.

Figure 8.9: Expected statistical uncertainties of measured top quark mass. Plots shown here are after calibration.

CHAPTER 9

THE SYSTEMATIC EFFECTS

There are many effects in the analysis that could affect the precision of the measurement, and we quantify these effects as systematic uncertainties on the measurement. The evaluation of systematic effects at CDF is described in the following: we tune a systematic effect on the measurement both up and down by one sigma, and if their effects on the measurement are in the opposite way, meaning that the measurements after shifting the systematic effect in both directions lead to results distant from the nominal result (without shifting the systematic effect) in opposite directions, then we take half of their difference (difference between results after shifting the systematic effect) as the uncertainty caused by that systematic effect. If shifting the systematic effect results in measurement shifting in the the same direction from the nominal, we will take half the largest difference from the nominal.

In evaluating the systematic effects, we perform pseudo experiments in the same way as we did for the bias checks. To be conservative, if the associated statistical uncertainty (due to limited event statistics of samples) is larger than the systematic uncertainty itself, we take the former as the value for that systematic effect.

In the following sections, the common sources of systematic effects for both the top quark decay width and mass measurements will be described first, and some specific effects and treatment of systematic effects for both measurements will be explained later.

9.1 Sources of Systematic Effects

9.1.1 *Residual Jet Energy Scale*

Jet energy scale has a large influence on both the top quark decay width measurement since a change of jet energy scale will change the jet energies and possibly the shapes of the reconstructed variables, the evaluated likelihood, and thus affect the final fitted value of the top quark decay width or mass. The use of W resonance for the *in situ* JES calibration

Table 9.1: JES systematic effects for top quark Mass measurement.

Levels	$\Delta M_{top}(Gev/c^2)$
Level 1	0.05
Level 4	0.03
Level 5	0.45
Level 6	0.04
Level 7	0.20
Level 8	0.04
Total	0.50

Table 9.2: JES systematic effects for top quark decay width measurement.

Levels	$\Delta \Gamma_{top}(GeV)$
Level 1	0.18
Level 4	0.04
Level 5	0.24
Level 6	0.04
Level 7	0.31
Level 8	0.09
Total	0.45

captures the overall JES shift but fails to fully account for the delicacies of JES, since each level of the JES corrections behaves differently as a function of η and p_T . To study any possible residual systematic effect due to the JES, we generate MC samples by shifting the JES up and down by one σ_c at each correction level while keeping other correction levels intact. Event selection and reconstruction are rerun for these generated samples. We then run PEs for them, with the JES prior term removed in the likelihood function, and take half of the difference between results from the $+\sigma$ and $-\sigma$ shifts as the systematic uncertainty for each of the JES levels. Table 9.1.1 shows a summary of the residual JES systematic uncertainties for the top quark mass measurement at each level, and the quadrature of values due to different levels of jet energy correction gives a total of $0.44 \text{ GeV}/c^2$ systematic uncertainty for the top mass measurement.

Table 9.2 summarize the shifts of the measured top quark decay width due to different levels of jet corrections, and the total shift is 0.45 GeV .

Table 9.3: b -JES systematic effects for top quark mass measurement.

Sources	$\Delta M_{\text{top}}(\text{GeV}/c^2)$
Calorimeter response	0.12
Semi-leptonic decay	0.11
b -fragmentation	0.24
Total	0.29

9.1.2 b -Jet Energy Scale

The analysis assumes that the JES is the same for all flavors of jets. However, the fragmentation, the calorimeter response, and the semileptonic decay feature of b -jets make it different from generic jets. The b -hadron resulting from the b -quark fragmentation carries a larger fraction of the momentum from the parent quark than for light quark jets. This is because the b -quark is much heavier and thus is slightly decelerated when forming a b -hadron. We estimate this effect by varying the parameters used in the Bowler fragmentation model [52] in the PYTHIA MC generator using two kinds of parameters from the SLD and LEP results [53]. The calorimeter may also respond to b -jets and light flavor jets differently since the charged particle fraction and momentum spectrum of b -jets is different from that of light jets. The effect from the calorimeter response is evaluated by shifting the E_T of jets identified as from b quark up and down by 1% and multiply the result by 0.2 since the expected uncertainty from this effect is 0.2%. The uncertainty of the b -quark semileptonic decay ratio may also affect the b -jet energy, and we evaluate by reweighting the nominal $t\bar{t}$ signal sample to vary the ratio up and down by one sigma. The summary of systematic effect due to b -jet energy scale is on Table 9.3 for the top quark mass measurement and on Table 9.4 for the top quark decay width measurement.

9.1.3 Initial and Final State Radiation

Effects due to uncertain modeling of initial-state radiation (ISR) and final-state radiation (FSR) are studied. A control sample of Drell-Yan events is used to measure the amount

Table 9.4: b JES systematic effects for top quark decay width measurement.

Sources	$\Delta\Gamma_{\text{top}}(\text{GeV})$
Calorimeter response	0.21
Semi-leptonic decay	0.01
b -fragmentation	0.12
Total	0.24

Table 9.5: IFSR systematic effects for top width measurement.

Type	subcategories	Measured $\Gamma_{\text{top}}(\text{GeV})$
Nominal	ttop25(PYTHIA)	2.13
More IFSR	dtops1	2.90
Less IFSR	dtops2	2.41
Total 0.24		

1 of ISR present in the sample as a function of the dilepton invariant mass. This quantity and
 2 its uncertainty is then extrapolated to the $t\bar{t}$ mass region. We use two samples to estimate
 3 the radiation effect, one with increased ISR and FSR, the other with decreased ISR and
 4 FSR, and the systematic uncertainty is on Table 9.5 and Table 9.6 for the top width and
 5 mass measurement, respectively.

9.1.4 MC Generators

7 Most of the MC simulated sample used in the analysis are generated by the PYTHIA MC
 8 generator, and it is important to check whether the result obtained from other MC gener-
 9 ators is consistent with PYTHIA. We perform pseudo experiments generated with HERWIG,
 10 which includes spin correlations between the top and antitop quarks and has a different

Table 9.6: IFSR systematic effects for top mass measurement.

Type	subcategories	Measured $M_{\text{top}}(\text{GeV}/c^2)$
Nominal	ttop25(PYTHIA)	172.71
More IFSR	dtops1	172.86
Less IFSR	dtops2	172.57
Total 0.21		

Table 9.7: Generator systematic effects for top width measurement.

Type	Measured $\Gamma_{\text{top}}(\text{GeV})$
Nominal ttop25(PYTHIA)	2.53
dtops0(HERWIG)	2.12
Total 0.41	

Table 9.8: Generator systematic effects for top mass measurement.

Type	Measured $M_{\text{top}}(\text{GeV}/c^2)$
Nominal ttop25(PYTHIA)	172.71
dtops0(HERWIG)	172.06
Total 0.65	

fragmentation model. The results of different MC generators are shown on Table 9.7 and Table 9.8, for the top width and mass measurement, respectively.

9.1.5 Parton Distribution Functions

The parton distribution function (PDF) describes the probability of finding a gluon or quark with a certain fraction of the hadron momentum. Gluons and quarks of different flavors have different PDFs and have their own uncertainties, thus affecting the measurement. We perform PEs for a total of 46 different PDF sets. The default version of CTEQ5L is used by the PYTHIA generator. We also use two sets of MRST [54] PDFs, MRST72 which uses the same value of α_s as CTEQ5L, and MRST75 which uses a different value. We also use CTEQ6L and CTEQ6L1 that use different α_s values, and CTEQ6M and variation of 20 different eigenvectors in CTEQ6M up and down. To evaluate the total systematic uncertainty due to PDFs, we take the difference between, CTEQ5L and MRST72 (due to the different PDF set used), MRST72 and MRST75 (due to the different values of α_s used), and each pair of eigenvectors in CTEQ6M added in quadrature (due to the uncertainty in the PDF set itself), and the results are summarized in Table 9.9 and Table 9.10, for the top width and mass measurement, respectively.

Table 9.9: PDF systematic effects of top quark decay width measurement

Type	Measured $\Gamma_{\text{top}}(\text{GeV})$
MRST72 - CTEQ5L (different groups)	0.20
MRST75 - MRST72 (QCD scale)	0.03
20-Eigenvector (CTEQ6M)	0.18
Total	0.27

Table 9.10: PDF systematic effects of top mass measurement

Type	Measured $M_{\text{top}}(\text{GeV}/c^2)$
MRST72 - CTEQ5L (different groups)	0.01
MRST75 - MRST72 (QCD scale)	0.06
20-Eigenvector (CTEQ6M)	0.19
Total	0.20

9.1.6 Gluon-Gluon Fusion Fraction

The fraction of $t\bar{t}$ pair production coming from gluon-gluon fusion at the next-to-leading order calculation is $15\% \pm 5\%$ at the Tevatron. However the PYTHIA and HERWIG MC generators have 5.6% $t\bar{t}$ productions from the gluon-gluon fusion and the rest from the quark-antiquark annihilation. We therefore reweight the gluon-gluon events in the signal sample to 20% and perform PEs. The results are shown in Table 9.11 and Table 9.12, for the top width and mass measurement, respectively.

Table 9.11: Gluon-Gluon Fusion systematic effects for top width measurement.

Type	Measured $\Gamma_{\text{top}}(\text{GeV})$
Nominal ttop25(PYTHIA)	2.13
20% gg-fraction	2.45
Total < 0.32	

Table 9.12: Gluon-Gluon Fusion systematic effects for top mass measurement.

Type	Measured $M_{top}(GeV/c^2)$
Nominal ttop25(PYTHIA)	172.71
20% gg-fraction	172.71
Total < 0.01	

9.1.7 Background

Top width measurement

Modeling of background processes may induce systematic effect to the measurement. We examine several sources of background modeling effect. The first is the uncertainty on the fraction of different background types used in the analysis. We run PEs by drawing background events from only one of the background types, and if a particular background type consists of several n -parton samples, all the samples are used. For single-top PEs, both the s -channel and t -channel samples are used. For diboson PEs, WW , WZ , and ZZ samples are used. We take the largest shift from the nominal PE result as the systematic effect of background fraction.

Possible mismodeling of background can be caused by the choice of Q^2 in the W production process, where Q^2 is the square of the momentum transfer in the hard scatter process. We run PEs by drawing background events from only the $Wb\bar{b}$ samples and change Q^2 by factors of 2^2 and 0.5^2 .

Finally, we consider the modeling of the QCD background. We consider an alternative method for modeling the QCD process that is different from the one described in Section 5.1. We form templates from antielectron (fails two kinematic cuts of quality electrons) samples, which consist of events where the isolated electron is replaced by objects in the detector that look like but are not electrons, and run PEs with these templates.

Table 9.13 shows the width shifts due to background effects.

Table 9.13: Background systematic effects for top width measurement.

Categories	$\Delta\Gamma_{\text{top}}(\text{GeV})$
Background Fraction	
Light Flavor	0.04
Wbb	0.03
Wcc	0.02
Diboson	0.06
Single top	0.11
QCD	0.01
Background Shape	
Q^2	0.01
QCD (antiele)	0.01
Total	0.11

Table 9.14: Background systematic effects for top mass measurement.

Type	subcategories	Measured $M_{top}(GeV/c^2)$	$\Delta M_{top}(GeV/c^2)$
Nominal	ttop25(PYTHIA)	172.71	
Background Uncertainty	more	172.50	0.10
	less	172.53	
Background Shape	up	172.57	0.07
	down	172.59	
Total 0.12			

1 top mass

2 To take into account the possibility that we might have mismodeled the background, we
3 reweight the background shape up and down by one sigma according to the H_T of a back-
4 ground event and take half the difference as the systematic effect due to midmodeling. This
5 approach we take here is different from the top width measurement since the background
6 samples used in the top mass measurement is modeled from the data sample instead of using
7 MC generators. We also vary the expected number of background events within its uncer-
8 tainty, as shown in Table 5.5, and count half the difference as another systematic effect. We
9 find the total systematic effect due to background modeling to be $0.12 \text{ GeV}/c^2$, summarized
10 in Table 9.14.

9.1.8 Multiple Hadron Interaction

We estimate systematic uncertainties due to the difference in the instantaneous luminosity distribution between MC samples and data. The estimation is done by studying the dependence of the extracted top width or top mass on the number of z vertices (n_z) in the MC events. We group ttop25 (PYTHIA, $M_{\text{top}} = 172.5 \text{ GeV}/c^2$, $\Delta_{\text{JES}} = 0.0$) MC events based on n_z and extract the measured top mass as a function of n_z . As shown in Fig. 9.1 and Fig. 9.2, there is a sizable n_z dependence of the both the measured top width and top mass.

We evaluate the systematic uncertainty due to multiple hadron interaction based on the difference in the number of average interaction vertices between data and MC sample, and the fitting of points on Fig. 9.1 and Fig. 9.2, which gives a top width shift of 0.26 GeV, and 0.16 GeV/c^2 for the top mass measurement.

Figure 9.1: Multiple hadron interaction systematic effect for top width measurement.

9.1.9 Color Reconnection

Color reconnection (CR) is the effect that color connection of $t\bar{t}$ decay products to the $p\bar{p}$ beam remnants, which are the quarks in the initial $p\bar{p}$ system but are not involved in the

Table 9.15: Trigger Simulation systematic effects.

Type	Measured $M_{top}(GeV/c^2)$
Nominal ttop25(PYTHIA)	172.71
signal weight up	172.57
signal weight down	172.57
Total 0.14	

primary interaction. The energies of the $t\bar{t}$ system may be affected by the CR effect. We account the CR effect by running templates of two different sets of tune parameters of CR effect, namely Tune Apro and Tune ACRpro, where Tune Apro is a similar tune with nominal sample using PYTHIA 6.4.19 instead of PYTHIA 6.2.16 and Tune ACRpro includes CR effect. The systematic uncertainty for the top mass measurement due to CR effect is $0.20\text{ GeV}/c^2$, and the shift of top width is 0.41 GeV .

9.1.10 Other Systematic Effects

Trigger simulation and Calibration for top mass measurement

We count in the possible effects on top mass measurement due to trigger simulation. We reweight the trigger simulation up and down by one sigma and find the difference in the measured top masses, which gives $0.14\text{ GeV}/c^2$.

Jet Resolution Effect for the top quark decay width measurement

We talked about the jet energy calibration in Section 4.1.2, which is the procedure of correcting the calorimeter jet energy back to the original parton energy, and the associated systematic effect due to the correction procedure was also described and taken into account in the measurements. However, there are also uncertainties for any individual measured jet energy, which we call jet energy resolution. Jet resolution plays an important role in the top decay width measurement since we may fail to quote a two-sided measurement of the top

Table 9.16: Summary of systematic effects. All numbers are after calibration.
CDF II Preliminary 5.7 fb⁻¹

Systematic (GeV/c ²)	ΔM_{top}
Residual JES	0.50
Generator	0.65
PDFs	0.20
b jet energy	0.29
Background	0.12
gg fraction	< 0.01
Radiation	0.21
Trigger simulation	0.14
Multiple Hadron Interaction	0.16
Color Reconnection	0.20
Total Effect	0.98

width if a bad jet energy resolution is present. Jet resolution is studied at CDF using photon + jets and dijet samples. We evaluate the systematic effect to top width measurement by worsening the jet resolution by smearing jet energies with a Gaussian function, which has a σ that is 5% of our default jet resolution. We again use nominal MC sample of input top width 1.5 GeV to calculate the width shift due to jet resolution effect, and take the difference of measured top width from PEs, which is 1.1 GeV.

9.2 Summary of Systematic Effects of Top Mass Measurement

Table 9.16 summarizes all systematic effects on the top mass measurement.

9.3 Summary of Systematic effects of Top Decay width

measurement

Table 9.17 summarizes the shifted measured top width due to all systematic effects. The total shift of top width is calculated by taking the quadrature of width shifts of subcategories of sysematics, which is 1.61 GeV.

Figure 9.2: Multiple hadron interaction systematic effect for top mass measurement.

Table 9.17: Summary of shifted measured top width due to systematic effects. All numbers have units of GeV.

Systematic	$\Delta\Gamma_{\text{top}}$ (GeV)
Residual JES	0.29
Jet Resolution	1.10
Generator:	0.41
PDFs	0.27
b jet energy	0.24
Background shape	0.11
gg fraction	0.32
Radiation	0.24
Lepton energy	0.17
Multiple Hadron Interaction	0.26
Color Reconnection	0.85
Total Effect	1.61

Figure 9.3: Color Reconnection.

9.3.1 Convolve Systematic Effects into Top Width Measurement

To incorporate systematic effects into the limit(s) on Γ_{top} we use a convolution method for folding systematic effects into the likelihood function [55, 56]. First, we convolute the original likelihood function $\mathcal{L}_0$ (pure statistic) with a Gaussian function related to systematic effects to obtain a new likelihood function $\mathcal{L}$:

$$\mathcal{L}(\Gamma_{\text{top}}|x) = \int d\tilde{\Gamma}_{\text{top}} \mathcal{L}_0(\tilde{\Gamma}_{\text{top}}|x) \frac{e^{-\frac{1}{2}(\frac{\tilde{\Gamma}_{\text{top}} - \Gamma_{\text{top}}}{\sigma})^2}}{\sqrt{2\pi}\sigma} \quad (9.1)$$

where x represents data and σ is equal to the total top width shift due to systematic effects. From Section 9.1 we know that $\sigma = 1.61$ GeV. After the convolution of systematics we expect the shape of likelihood function to be broader than original one. Figure 9.4 shows an example from MC sample twid50 ($\Gamma_{\text{input}} = 5.0$ GeV, $\Delta JES = 0$). The red points represent the likelihood function after incorporating systematic effects while blue points the original likelihood function. Note that after convolution the shape of likelihood function should extend to region of negative top width, however we cut it at 0 GeV (physical boundary) and this doesn't affect the minimization of likelihood function.

Second, we shift this new likelihood function horizontally according to a Gaussian distribution with $\sigma = 1.61$ GeV. In fact, with no boundary effect considered, the first step changes the shape of likelihood function while the second step changes the best fit top width of the likelihood function. Then we repeat what was done in Section 8.1: (1) get the $\Delta\chi^2$ distribution for each MC sample and find the critical $\Delta\chi^2_{\text{c}}$ value; (2) plot $\Delta\chi^2_{\text{c}}$ vs input top width; (3) overlap this plot with data fit and find the limit(s) of top quark width. Note that in the data fit the likelihood function is also convoluted with the same Gaussian function. Figure 9.5 shows the overlap of $\Delta\chi^2_{\text{c}}$ vs input top width plot and data fit, which gives an upper limit of top quark width $\Gamma_{\text{top}} < 7.6$ GeV. To help compare the results with and without systematics incorporated we put together Fig. 8.4 and Fig. 9.5 and see how systematic effects affect the result of top width upper limit, as shown in Fig. 9.6.

Figure 9.4: Original likelihood function(blue points) and the likelihood function after incorporating systematic effect(red points). MC sample twid50 ($\Gamma_{input} = 5.0$ GeV)

Figure 9.5: Overlap of data fit and plot of $\Delta\chi_c^2$ vs input top width(95% CL) with systematic effect.

Figure 9.6: Overlap of Fig. 8.4 and Fig. 9.5(95% CL).

CHAPTER 10

RESULTS

10.1 Results of Top Quark Width Measurement

We use dataset collected at CDF II detector, corresponding to a total integrated luminosity of 4.3fb^{-1} , to measure the top quark decay width. After performing the log-likelihood fit of data, the best fit gives $\Gamma_{\text{top}}^{\text{datameas}} = 1.90_{-1.50}^{+1.88}\text{ GeV}$ and $\Delta_{\text{JES}} = 0.07_{-0.21}^{+0.20}$, as shown in Figure 10.1. The data fit in Fig. 9.5 is the actually the 1D projection of the 2D log-likelihood fit for measured top width. As already stated in Section 9.3.1, using a Feldman-Cousins limit setting approach, from the overlap of data fit and $\Delta\chi_c^2$ distribution we obtain an upper limit of top quark width $\Gamma_{\text{top}} < 7.6\text{ GeV}$ at 95% CL, with systematic effects incorporated.

Figure 10.1: 2D Log-likelihood fit of data, giving $\Gamma_{\text{top}}^{\text{datameas}} = 1.90_{-1.50}^{+1.88}\text{ GeV}$ and $\Delta_{\text{JES}} = 0.07_{-0.21}^{+0.20}$.

Figure 10.2 and 10.3 show the distributions of reconstructed top mass m_t^{reco} and dijet mass w_{jj} for the dataset, overlaid with probability density functions from input $\Gamma_{\text{top}} = 1.5$

1 GeV and the full lepton+jets backgrounds.

Figure 10.2: Reconstructed top mass distributions of dataset, overlaid with probability density functions from input $\Gamma_{\text{top}} = 1.5$ GeV and full lepton+jets backgrounds. Top: 1-tag events of data; Bottom: 2-tag events of data.

2 We also measure top width at 68% confidence level, which is shown in Figure 10.4, and
3 we have central limits of top quark width : $0.6 \text{ GeV} < \Gamma_{\text{top}} < 3.9 \text{ GeV}$ of pure statistic
4 result, and $0.3 \text{ GeV} < \Gamma_{\text{top}} < 4.4 \text{ GeV}$ with systematic effects incorporated.

Figure 10.3: Dijet mass of W boson distributions of dataset, overlaid with probability density functions from input $\Gamma_{\text{top}} = 1.5$ GeV and full lepton+jets backgrounds. Top: 1-tag events of data; Bottom: 2-tag events of data.

10.2 Results of Top Quark Mass Measurement

We extract the most probable top quark mass of data from the likelihood function. Figure 10.5 is the two-dimensional likelihood fit for data sample. After the calibration we mention before, we measure

$$\begin{aligned} M_{top} &= 172.32 \pm 2.37 \text{ GeV}/c^2, \text{ calibrated} \\ \Delta_{JES} &= 0.28 \pm 0.42 \text{ calibrated} \end{aligned}$$

Figure 10.6 shows the three kinematic variables from data, overlaid with their corresponding one-dimensional p.d.f.s from sample *ttop25* (with input $M_{top} = 172.5 \text{ GeV}/c^2$ and $\Delta_{JES} = 0.0$) plus the background fully modeled.

We run a variety of cross-checks on our measurement, which corresponds to the top quark mass measurement using data samples that come from different collection period of the CDF II detector, and subsamples of data events such as 1tag and 2tag events, as well as using slightly different likelihood functions, details of which are summarized on Table 10.1.

Figure 10.4: Top quark width measurement at 68% CL.

Figure 10.5: 2D likelihood fit contours for the data fitting for top mass measurement.

Figure 10.6: Distribution of kinematic variables m_t^{reco} , W_{jj} , and $m_t^{\text{reco}(2)}$ from data, overlaid with their corresponding 1-d p.d.f.s from MC sample *ttop25* (with input $M_{\text{top}} = 172.5 \text{ GeV}/c^2$ and $\Delta_{JES} = 0.0$) plus the background fully modeled. Both 1tag (left) and 2tag (right) events are displayed.

Table 10.1: Cross checks

Sub data events	$M_{top}(\text{GeV}/c^2)$		Δ_{JES}	
	nsig(1tag)	nsig(2tag)	nbkg(1tag)	nbkg(2tag)
p0-p28 data(Nominal)	172.3 ± 2.4		0.3 ± 0.4	
	662.4 $^{+42.1}_{-41.5}$	234.5 $^{+19.4}_{-18.8}$	456.0 $^{+28.1}_{-28.0}$	47.2 $^{+10.0}_{-9.9}$
p0-p28 data(without bkgd constraint)	172.0 ± 2.6		0.3 ± 0.5	
	491.4 $^{+68.5}_{-68.6}$	219.4 $^{+28.1}_{-27.8}$	647.6 $^{+70.7}_{-68.7}$	63.5 $^{+25.8}_{-24.1}$
p0-p9 data	163.6 ± 13.4		0.1 ± 1.6	
	- $^{+-}_{--}$	37.9 $^{+10.5}_{-9.9}$	265.5 $^{+15.6}_{-16.1}$	27.8 $^{+9.1}_{-8.9}$
p10-p13 data	178.1 ± 13.9		-0.25 ± 1.2	
	- $^{+-}_{--}$	11.4 $^{+8.7}_{-7.9}$	156.7 $^{+13.5}_{-12.9}$	33.5 $^{+8.4}_{-8.2}$
p14-p17 data	FAILURE		FAILURE	
	-	-	-	-
p18-p28 data	172.6 ± 3.1		0.4 ± 0.5	
	303.7 $^{+34.9}_{-34.3}$	111.8 $^{+14.9}_{-14.3}$	399.0 $^{+27.5}_{-27.3}$	40.3 $^{+9.6}_{-9.5}$
1btag	171.8 ± 2.7		0.6 ± 0.5	
	663.0 $^{+42.1}_{-41.5}$	-	455.6 $^{+28.1}_{-28.0}$	-
2btag	173.2 ± 3.7		-0.2 ± 0.6	
	-	234.4 $^{+19.4}_{-18.8}$	-	47.3 $^{+9.9}_{-9.9}$

CHAPTER 11

CONCLUSIONS

We present measurements of the top quark decay width and top quark mass using a template method. For the top width measurement, we use events in the lepton+jets $t\bar{t}$ pair decay channel that correspond to an integrated luminosity of 4.3 fb^{-1} collected by the CDF II detector at the Fermilab Tevatron. We build the reconstructed top mass, m_t^{reco} , and the dijet mass of W boson, w_{jj} , for MC and background samples, as well as for data set. By taking a Feldman-Cousins limit setting method, we report an upper limit of top quark width $\Gamma_{\text{top}} < 7.6 \text{ GeV}$ at 95% CL, corresponding to a lower limit on the top quark lifetime of $\tau_{\text{top}} > 8.7 \times 10^{-26} \text{ s}$. We also measure the central limit of top quark width at 68% CL: $0.3 \text{ GeV} < \Gamma_{\text{top}} < 4.4 \text{ GeV}$.

For the top quark mass measurement, we use data that correspond to an integrated luminosity of 5.7 fb^{-1} and have a $\cancel{E}_T + \text{jets}$ signature in the final state. With a similar template method used in the top width measurement, we measure the top quark mass to be

$$\begin{aligned}
 M_{\text{top}} &= 172.32 \pm 2.37 \text{ (stat + JES)} \pm 0.98 \text{ (other syst)} \text{ GeV}/c^2 \\
 &= 172.32 \pm 1.80 \text{ (stat.)} \pm 1.54 \text{ (JES)} \pm 0.98 \text{ (syst.)} \text{ GeV}/c^2 \\
 &= 172.32 \pm 2.56 \text{ GeV}/c^2 \\
 \Delta_{JES} &= 0.28 \pm 0.42\sigma_c
 \end{aligned}$$

REFERENCES

- [1] M. Jezabek and J.H. Kuhn, Phys. Rev. **D 48**, 1910 (1993); **49**, 4970 (1994).
- [2] M. Jezabek and J.H. Kuhn, Nucl. Phys. **B 314**, 1 (1989).
- [3] A. Denner and T. Sack, Nucl. Phys. **B 358**, 46 (1991).
- [4] G. Eilam, R. R. Mendel, R. Migneron, and A. Soni, Phys. Rev. Lett. **66** 3105 (1991).
- [5] ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavor Working Groups, arXiv:1012.2367v2.
- [6] H. Flücher *et al.*, Eur. Phys. J. C **60**, 543 (2009).
- [7] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652-657 (1973).
- [8] V. Barger, J. L. Hewett and R. J. N. Phillips, Phys. Rev. **D 41**, 3421 (1990).
- [9] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, 101803 (2009).
- [10] K. I. Hikasa and M. Kobayashi, Phys. Rev. **D 36**, 724 (1987).
- [11] C. S. Li, J. M. Yang, and B. Q. Hu, Phys. Rev. **D 48**, 5425 (1993).
- [12] J. L. Diaz-Cruz, M. A. Perez, G. Tavares-Velasco, and J. J. Toscano, Phys. Rev. **D 60**, 115014 (1999).
- [13] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **102**, 042001 (2009).
- [14] The transverse momentum p_T and transverse energy E_T of a particle are defined as $|\vec{p} \sin \theta|$ and $E \sin \theta$ respectively, where θ is the polar angle of the particle momentum with respect to the proton beam direction. The missing transverse energy, an imbalance of energy in the plane transverse to the beam direction, is defined as $\cancel{E}_T = |\sum_{towers} E_T \hat{n}_T|$

, where $\hat{n}_T$ is the unit vector normal to the beam and pointing to a given calorimeter tower and E_T is the transverse energy measured in that tower.

[15] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **75**, 111103(R) (2007).

[16] J. F. Gunion, H. E. Haber, F. E. Paige, W.-K. Tung, and S. S. D. Willenbrock, Nucl. Phys. B **294**, 621 (1987).

[17] G. L. Kane and S. Mrenna, Phys. Rev. Lett. **77**, 3502 (1996).

[18] Particle Data Group, pdg.lbl.gov, RPP2011-REV-TOP-QUARK.

[19] M. Jezabek and J.H.Kuhn, Nucl.Phys.B 314,1 (1989)

[20] M. Cacciari *et al.*, Sov. Phys. JETP **04**, 068 (2004).

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

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

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

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

[25] arXiv:hep-ex/0508029v36 (2007).

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

[27] A. Bhatti *et al.*, Nuclear Instruments and Methods in Physics Research A 566 (2006) 375-412.

[28] T. Affolder *et al.* (CDF Collaboration), "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV". Phys. Rev. D 64, 032002 (2001).

- 1 [29] F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **57**, 5382 (1998).
- 2 [30] D. Acosta *et al.* (CDF Collaboration), "Measurement of the $t\bar{t}$ production cross section
3 in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with secondary vertex b
4 -tagging". Phys. Rev. D **71**, 052003 (2005).
- 5 [31] T. Aaltonen *et al.* (CDF Collaboration), "Measurement of the $t\bar{t}$ production cross section
6 in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using events with large missing transverse energy and
7 jets". Phys. Rev. D **84**, 032003 (2011).
- 8 [32] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **83**, 111101 (2011).
- 9 [33] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **79**, 092005 (2009).
- 10 [34] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **81**, 031102 (2010).
- 11 [35] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **81**, 052011 (2010).
- 12 [36] R. Brun and F. Rademacher, ROOT version 4.00/08,
13 <http://root.cern.ch/root/html/TMultiLayerPerceptron.html>.
- 14 [37] J. D. Bjorken and S. J. Brodsky, Phys. Rev. D **1** (1970) 1416.
- 15 [38] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **105**, 012001 (2010).
- 16 [39] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).
- 17 [40] B. W. Harris *et al.* Phys. Rev. D **66**, 054024 (2002).
- 18 [41] S. Moch and P. Uwer, Nucl. Phys. B, Proc. Suppl. **183**, 75 (2008).
- 19 [42] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **77**, 011108 (2008).
- 20 [43] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **82**, 112005 (2010).
- 21 [44] S. Eidelman *et al.* (Particle Data Group Collaboration), Phys. Lett. B **592**, 1 (2004).

- 1 [45] David W. Scott, *Multivariate Density Estimation*, 1992
- 2 [46] Ian S. Abramson, The Annals of Statistics, Vol.10, No.4, 1217-1223 (1982).
- 3 [47] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman and
4 Hall, 1986.
- 5 [48] Stephan R. Sain, *Multivariate Locally Adaptive Density Estimation*, Computational
6 Statistics & Data Analysis 39, 165-186 (2002).
- 7 [49] Fedorko, thesis
- 8 [50] C. Loader, *Local Regression and Likelihood* (Springer, New York, 1999).
- 9 [51] B. Efron, Ann. Stat. **7**, 1 (1979).
- 10 [52] M. Bowler, Z. Phys. C **22**, 155 (1984).
- 11 [53] Y. Peters, K. Hamacher and D. Wicke, Fermilab publication FERMILAB-TM-2425-E
12 (2009).
- 13 [54] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C **4**, 462
14 (1998).
- 15 [55] J. Berger, B. Liseo, and R. Wolpert, Statist. Sci. **14**, 1 (1999).
- 16 [56] L. Demortier, *Proceedings of the Conference on Advanced Statistical Techniques in Par-*
17 *ticle Physics, Durham, 2002*, (Institute for Particle Physics Phenomenology, University
18 of Durham, UK, 18-22 March 2002); IPPP/02/39 (2002), p. 145.