TOP QUARK PAIR IN ASSOCIATION WITH AN EXTRA JET: PHENOMENOLOGICAL ANALYSIS AT THE TEVATRON

By

Moh'd Ahmad Hussein

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Physics and Astronomy

2011

ABSTRACT

TOP QUARK PAIR IN ASSOCIATION WITH AN EXTRA JET: PHENOMENOLOGICAL ANALYSIS AT THE TEVATRON

By

Moh'd Ahmad Hussein

The first measurement of the cross section of the top quark pair in association with an extra hard jet ($t\bar{t}$+jet) has been performed with 4.1 fb^{-1} of data collected at CDF. The measurement is an important test of perturbative QCD, as NLO effects play an important role in the calculation of the theoretical cross section. In addition, it is also important as a preview of the LHC, for which almost half of the top quark events will be produced with extra jets. Therefore, this process will be a substantial background for many new physics signals. The measurement is performed using `SecVtx` tagged events in the lepton plus jet channel. A data-driven approach is used to predict the background content, and a 2D likelihood is formed to simultaneously measure the $t\bar{t}$+jet and $t\bar{t}$ without extra jet cross sections. The measured result is $\sigma_{t\bar{t}+\text{jet}} = 1.6 \pm 0.2_{\text{stat}} \pm 0.5_{\text{syst}}$ pb which is in agreement with the recent NLO SM prediction $\sigma_{t\bar{t}+\text{jet}} = 1.791(1)^{+0.16}_{-0.31}$pb .

In order to elucidate the kinematic profile of the extra jet, an isolation algorithm has been developed. The algorithm has extracted correctly the extra jet out from the final state jets more than 60% of the time. This allowed for correcting the measured distributions of the extra jet for purity/efficiency in order to compare them with the MC distributions. The differences in the kinematic of the extra jet using different `SecVtx` requirements and different MC models (`PYTHIA` & `MCFM`) have been studied. The agreement between data and the simulations is reasonable. The fifth and the fourth highest E_T jet in the final state of $t\bar{t}$+jet sample are found to be equally likely the extra jet.

© Copyright by
MOH'D AHMAD HUSSEIN
2011

DEDICATION

To Ahmad Abdullah Hussein; first as a father, then as a mentor, and now as a pillar of support ...

It is he; the first, who recited for me Luqmân's advice to his son as it was revealed:

> O my dear son! Almighty Allah will bring all things to light, be they as small as a grain of a mustard seed, be they hidden inside a rock, or high above in the heavens, or deep down in the earth. Almighty Allah knows the finest mysteries and is well aware of all things [verse 16].
>
> O my dear son! Establish prayers, enjoin good and forbid evil. Endure with patience whatever befalls you, surely, all that is of the matters requiring determination [verse 17].
>
> Do not speak to the people with your face turned away with pride, nor walk proudly on earth; for Almighty Allah does not love any self-conceited boaster [verse 18].
>
> Be moderate in your pace and lower your voice; for the harshest of voices is the braying of a donkey [verse 19].
>
> The Holy Qura'n: Chapter 31

I know that you weren't able to pursue your own PhD for a very noble cause.

I hope that you consider this dissertation yours as well.

ACKNOWLEDGMENTS

This acknowledgment is intended for those who have had a direct involvement with my graduate career.

Aous Abdo: I have to say right off that I can't pay back your unquestionable help. Awsy (as his Mom calls him) the then-research associate was not only among the few nearest and dearest to me but was also my best grad fellow. He was there on August 22, 2004 to pick me up from Metro airport and he was there last Fall to give me his dissertation LaTeX template. In between, Awsy was always there.

Joey Huston: I'll be unfair if I offer my praise to Joey for being a mere adviser or for only admitting me into MSU. With his forbearance, Joey went above and beyond the duties of an adviser. I have learned from him how to multitask and prioritize. I will never forget his peculiar way of examining my histograms (out of his glasses frame) and then pointing to my mistakes! I have benefited from his excellent communication skills with a large collaboration like the one CDF is. The only two things I may have failed to mimic were his astonishing ability of throwing jokes while we were in formal meetings that discuss a very complicated QCD issues, and walking while barefoot in the hallway!

Wayne Repko: Over the past six years, Repko was the ray of advice I always followed. This goes side by side with the conclusive fact that he is among the very few great teachers to learn from and to emulate. He inspired me so much when I took his rigorous QED class; when he rarely looked up at his notes while he was still recovering from a wicked stroke that hit him. The Sunday afternoon coffees I used to share with him once a while are among the delightful memories I will bear forever.

Thomas Rockwell: I came in to HEP having no idea how to copy files in a Unix directory! I looked around in the grad office and was told: It is Tom ... and indeed it was him! When C++ was keeping me down it was he who I would rush into his office. When my makefiles were (making) me crazy it was Tom who (made) them

a real makefiles. I am indebted to Tom not only for his fruitful help with the bash scripts, ROOT, debugging etc, but for the neat smile he used always to welcome me with.

Gene Flanagan: Gene was my second adviser who gave me my first piece of code to play with. Since the first moment I arrived at Fermilab in May 2006 where he was waiting for me near the Highrise, he since then has worked hard trying to transfer his expertise to me. I'd bet that he answered more than a thousand emails from a novice new grad student in the first two years!

Thomas Schwarz: After all, it was TomTrees that I inherited and ran my code over! Tom is the one who did the majority of the work of the background estimation, which many people, including me, have used. I've benefited from his crystal clear reasoning and pinpointed critiques. An Arabic proverb says (more or less) that a pot needs three stones to be stable; Tom was the third of the three advisers that I have; each was an expert in his field.

Brian Martin: Brian is an irreplaceable colleague and friend. Thank you for your help in SpartyJet, for the good coffee you used to prepare, for the joy and the humor you were able to create, and for keep greeting me with my traditional greeting; Salam! I am pretty sure you will be a pioneer scholar and a great father as well (It's soon I guess!). I wish you and Kate the best of luck wherever you end up after Argonne lab.

Brenda Wenzlick: A big thumbs-up to our HEP secretary, for her ability to make things possible. Thank you Brenda for preparing my trips to Fermilab even though I used to give you short notice. I still remember when you came down all of the stairs to charge me up with a fresh cup of coffee while I was setting in the early morning for the qualifier; and how the faculty-member proctor got jealous and insisted on knowing who you were! A special acknowledgment goes also to Debbie Barratt, our very professional graduate program secretary.

Dean Halderson: A truly special praise is owed to Dean, from WMICH, for be-

lieving in me and supporting me throughout my first year in the US. Your kindness kept me rational.

I want to thank JJ Schmidt, Stephen Hahn and William Badgett for their gentleness and true help while I was on shifts at CDF.

Lastly, but certainly not least, in physics we study a lot of paradoxes but I believe that the Joey-Repko-Rockwell-Dean paradox is not addressed in any literature yet! The huge contradiction between the quality of these people, their tranquility, their dignity, compared with the values and creeds of those who are located at 1600 Pennsylvania Avenue is bewildering ...

Contents

List of Tables

List of Figures

Chapter 1

Relevant Theory

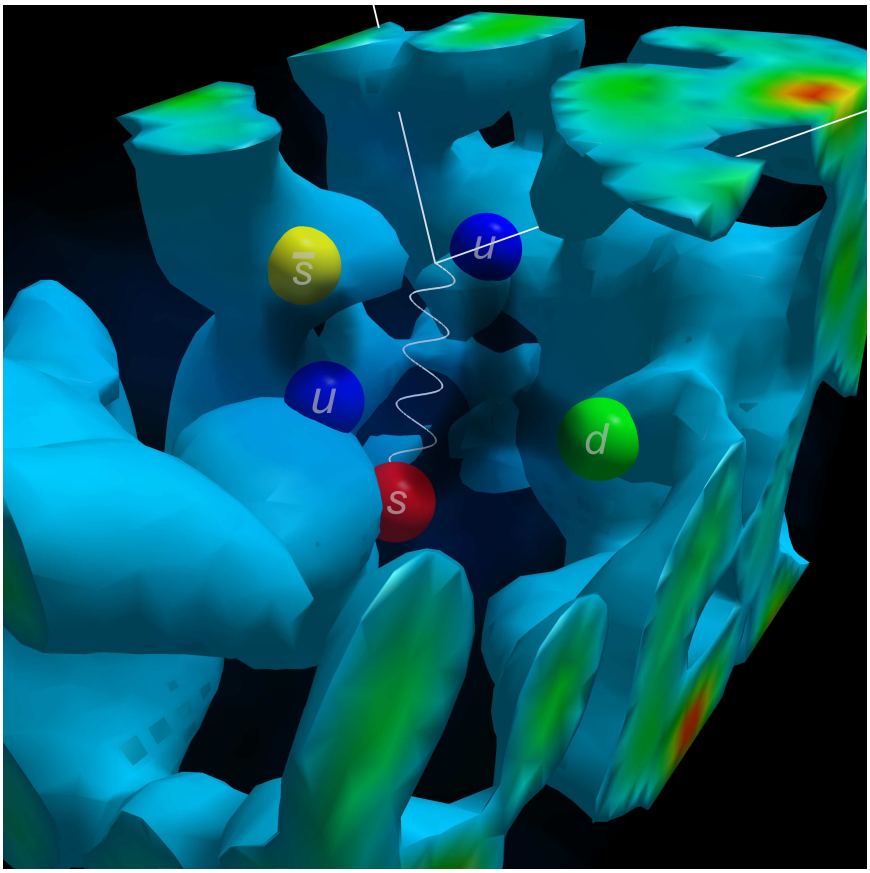


Figure 1.1: Artist's depiction of a proton (uud) fluctuating into a $\Lambda(uds)$ baryon and a $K^+(\bar{s}u)$ meson, in the background of vacuum fluctuations of gluons. The false color scale denotes the intensity of the gluon field at a given moment, proceeding through the rainbow from blue to red (most intense), from [1].

1.1 Introduction

High Energy Physics (HEP) endeavors to enumerate the elementary particles which constitute known matter and to characterize the interactions which govern their behavior. The best theory to date is what has come to be called the Standard Model (SM) of particle physics, which has proved to be a very successful theory at the energy scales attainable by accelerators all the way up to the present-day Tevatron at the Fermi National Accelerator Laboratory (Fermilab) and the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN). Hundreds of measurements varying in analysis technique and detection apparatus have all confirmed the predictions of the SM at a remarkable level of precision.

With the observation of the top quark in 1995, nearly the full catalog of SM particles has been directly observed, with only the Higgs boson evading detection. With the approximate mass of a gold nucleus, the top quark is the heaviest known elementary particle. Due to this feature and others, there is a very active program at the Tevatron to determine the properties of the top quark.

This thesis concerns a first measurement of the top quark pair production in association with an extra hard jet and the kinematics of the extra jet, utilizing 4.1 fb^{-1} of data collected at Fermilab.

I'll begin the thesis by covering the relevant physics of interest underlying the top quark within the SM framework. I'll proceed with a description of the experimental apparatus, both the Collider Detector at Fermilab (CDF) and the series of Fermilab accelerators which culminate in the Tevatron to accelerate protons and anti-protons to a center-of-mass energy of 1.96 TeV. A discussion of how the various decay objects from the top are reconstructed in the detector will follow, along with a discussion of the selection cuts used to obtain the data sample for the measurement, and the Monte Carlo events used in the construction and evaluation of the analysis. Chapter 4 describes both the counting analysis and the kinematic analysis at length, followed

by the results and a brief conclusion.

1.2 The Standard Model of Particle Physics

The SM of particle physics is a quantum field gauge theory describing the dynamics of particles that interact via the strong, electromagnetic and weak interactions based on the SU(3) gauge symmetry of Quantum Chromodynamics (QCD) and the SU(2)×U(1) gauge symmetry of Quantum Electrodynamics (QED). Gravity is not included in the SM, but it's extremely weak compared to the other forces, and not important at the relevant energy scale probed by particle physics.

These SM particles are half-integer spin fermions, which interact via the three forces mediated by integer spin bosons. Each force has at least one force carrier boson. The massless, neutral photon (γ) mediates the electromagnetic force, which is a long-range force and the one most familiar from every day experience. The strong force is mediated by a set of eight bosons collectively referred to as gluons. The gluon (g) is also massless and neutral but itself carries the corresponding charge (color). Strong interactions bind quarks together into nucleons, and the nucleons into atomic nuclei. The weak interaction is mediated by three massive bosons, the W^+, W^- ($\mathrm{M_W}$ = 80.4 GeV/c^2) and neutral Z° ($\mathrm{M_Z}$ = 91.19 GeV/c^2) [1] [2]. Outside of HEP, the weak force shows up primarily through the decay of radioactive nuclei. However, the electromagnetic and weak interactions are unified at high energies (see section 1.3.4).

The fundamental fermions are further categorized into quarks and leptons. Quarks carry color as well as fractional electric charge and participate in all interactions. All of the fermions participate in the weak interactions, and all of the charged fermions interact in the electromagnetic interaction. It is only the weak interactions that can change the "flavor" or type of the quarks, but other interactions leave the flavor intact.

[1] Natural units ($\hbar = c = 1$) will be used from now on through all the thesis. Then energy, momentum and mass will appear as mass, and will normally be stated in GeV (so if translation is required, GeV → GeV/c or GeV/c^2). Examine, for example, Equation (1.12).

Quarks of electric charge $+\frac{2}{3}$ are grouped with quarks of electric charge $-\frac{1}{3}$ into SU(2) doublets. Leptons are colorless, and charged leptons are paired with neutral leptons, called neutrinos, into SU(2) doublets. Each pair is called a "generation" or a "family" of particles. The SM has three of them. The first generation is composed of the up and down quarks, and the electron and electron-type neutrino. These four particles are the constituents of ordinary matter. Their observed stability is a result of their being the lightest particles of their kind and the apparent conservation of lepton and baryon number [2].

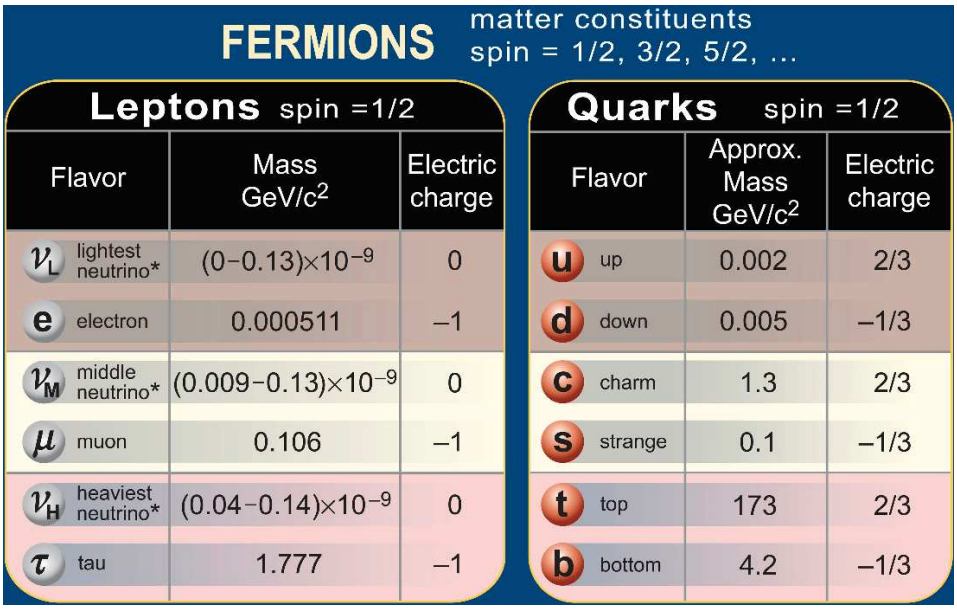
The other two generations are successively more massive and are unstable; they exist only as the product of collisions by cosmic rays and at man-made accelerators. There is a very vast discrepancy between the different masses of the fermions. The hierarchy of the three generations and their basic properties are shown in Figure 1.2. Each of these particles has an antiparticle, identical to the original (mass and spin) but with all the opposite quantum numbers.

It's still an open question why the top quark is so heavy compared to the rest of the particles. Therefore, understanding the properties of the top quark is essential to fundamentally understand the mechanism that gives rise to mass.

1.3 Quantum Gauge Field Theories

The three forces of the SM are described by so called quantum gauge field theories which obey local symmetries and are mediated by field quanta. The quanta of these fields carry spin 1 and are therefore called gauge bosons. I will describe these theories in some detail in the next few pages. There are several texts thoroughly describing the SM. I have relied heavily on [4, 5].

[2] The discovery of the Λ^{++} baryon, which consists of three up quarks (uuu) with parallel spins, seems to violate the Pauli exclusion principle. However, introducing a new quantum number (color) to the quarks will keep the exclusion principle intact. Quarks can then be colored red, green or blue, while anti-quarks are colored anti-red etc.



FERMIONS matter constituents spin = 1/2, 3/2, 5/2, …

Leptons spin =1/2

Flavor	Mass GeV/c^2	Electric charge
ν_L lightest neutrino*	$(0–0.13)\times10^{-9}$	0
e electron	0.000511	–1
ν_M middle neutrino*	$(0.009–0.13)\times10^{-9}$	0
μ muon	0.106	–1
ν_H heaviest neutrino*	$(0.04–0.14)\times10^{-9}$	0
τ tau	1.777	–1

Quarks spin =1/2

Flavor	Approx. Mass GeV/c^2	Electric charge
u up	0.002	2/3
d down	0.005	–1/3
c charm	1.3	2/3
s strange	0.1	–1/3
t top	173	2/3
b bottom	4.2	–1/3

Figure 1.2: The known fermions and their properties in the SM, from [3]. For more precise values, see Reference [2].

1.3.1 Quantum Electrodynamics (QED)

QED is the gauge theory that describes the electromagnetic force. It's the simplest and the most accurate theory in the SM. The Lagrangian $\mathcal{L}$ for the massless electromagnetic field A_μ interacting with a spin-$\frac{1}{2}$ field ψ of bare mass m is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{1.1}$$

Here, $F_{\mu\nu}$ is the electromagnetic field tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.2}$$

and D_μ is the covariant derivative

$$D_\mu = \partial_\mu + ieA_\mu Q \tag{1.3}$$

where e is the unit of electric charge and Q is the charge operator (i.e. $Q\psi = e\psi$) and A is the gauge field, the quantum of which is the photon. This Lagrangian is invariant under local $U(1)$ gauge transformations:[3]

$$\psi \longrightarrow e^{i\alpha(x)}\psi \tag{1.4}$$

and

$$A_\mu \longrightarrow A_\mu - \frac{1}{g_\mathrm{e}}\partial_\mu\alpha(x) \tag{1.5}$$

The coupling constant g_e is the elementary unit of the charge $e = \sqrt{4\pi\alpha}$, where the current measured vlaue for α is $\frac{1}{137.035999679(94)}$ [2]. Since it is small, we compute all matrix elements as a perturbation series, using α as the expansion parameter (see section 1.5). The $U(1)$ symmetry of QED leads to conservation of electric charge.

1.3.2 Quantum Chromodynamics (QCD)

QCD is the gauge theory that describes the strong force. Unlike QED, the phenomenology of this theory is considerably richer and more complicated; it is different in many respects. The Lagrangian $\mathcal{L}$ for three colored quarks of each quark flavor and eight gluons is

$$\mathcal{L} = -\frac{1}{4}F_a^{\mu\nu}F_{a\mu\nu} + \bar{\psi_j}(i\gamma_\mu D^\mu_{jk} - M_j\delta_{jk})\psi_k \tag{1.6}$$

where the indicies a, j and k refer to color and assume the value $a = 1, \ldots, 8$ and j,k = 1,2,3. The covariant derivative D acting on a quark field is

$$D^\mu_{jk} = \delta jk\,\partial^\mu + ig_\mathrm{s}(T_a)_{jk}\,G^\mu_a \tag{1.7}$$

[3] The invariance of $\mathcal{L}$ is actually under a local phase transformations (i.e. transform the field at every point in space and time) and it is referred, to as gauge transformations for only historical reasons.

where G_a^μ are the gluon fields, T_a are the SU(3) generators and g_s is the strong coupling constant $\sqrt{4\pi\alpha_s}$ analogous to that of QED;

$$T_a = \frac{1}{2}\lambda_a \tag{1.8}$$

The λ matrices are Gell-Mann matrices, which are to SU(3) what the Pauli spin matrices are to SU(2). The field tensor is more complicated than the corresponding QED case:

$$F_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f_{abc}, G_b^\mu G_c^\nu \tag{1.9}$$

The commutators of the λ matrices define the structure constants f_{abc} [4] :

$$[\lambda_a, \lambda_b] = 2if_{abc}\lambda_c \tag{1.10}$$

As a result, QCD predicts that gluons can interact with one another and hence themselves carry color charge. The consequence of this is that the interaction gets stronger at long distances, and weaker at short distances, a trend opposite to QED. This is referred to as asymptotic freedom. Therefore, whereas the QED coupling variation is extremely small over the accessible energy range, the variation in the QCD coupling is substantial and runs as :

$$\alpha_s(q^2) = \frac{12\pi}{(11n - 2f)\ln\left(\frac{q^2}{\Lambda^2_{\rm QCD}}\right)} \tag{1.11}$$

where $\Lambda_{\rm QCD}$ is the QCD energy scale, n and f are the number of the colors and the flavors (3 and 6 respectively). Here it can be seen that for high energy scales [5] ($q^2 \gg \Lambda^2_{\rm QCD}$), α_s is small enough to allow perturbative calculations that are similar to QED [6], but as the interaction energy decreases, the coupling becomes strong and

[4] It is the non-vanishing of the f_{abc} that make QCD a non-Abelian gauge theory.
[5] As it is for this analysis at the Tevatron.
[6] In this case, QCD is said to be renormalizable.

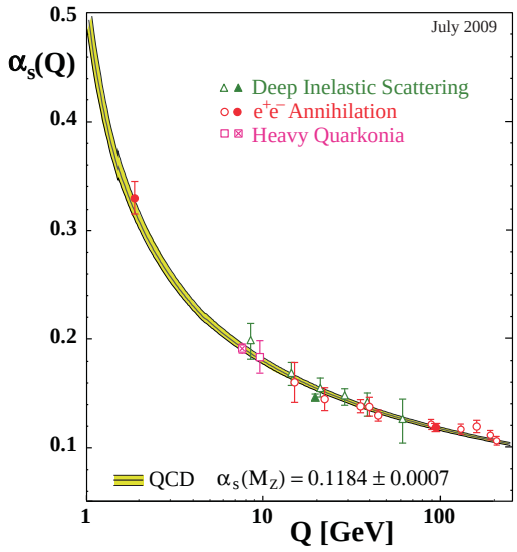


Figure 1.3: Measurement of α_s as a function of energy, from [2].

diverges at $\Lambda_{\rm QCD} \approx 0.2$ GeV making perturbative calculations impossible [7]. The variation of α_s with energy is shown in Figure 1.3 .

The strong coupling at low energy is responsible for the phenomenon of quark confinement; in nature, quarks are not observed in an isolated state. Instead, they are confined in colorless bound states of combinations of two or three quarks and anti-quarks, the hadrons. Hadrons come in one of two types: *baryons* (which are fermionic bound state of three quarks) and *mesons* (which are bosonic bound states of a quark and an anti-qurak). For example, the negative charged pion (π^-) is composed of d and $\bar{u}$ quarks. On the other hand, the proton is comprised of two u and one d quarks as depicted in Figure 1.1.

Hadronization

These (uud) constituent quarks of the proton are the "valence" quarks [8]. Other virtual quark–anti-quark pairs appear inside the proton due to gluon interactions

[7] Lattice QCD techniques are developed to cope with this behavior; the space-time continuum is replaced by a finite lattice of discrete points. The equations of QCD are then solved numerically hoping to achieve realistic results in the limit as the lattice spacing shrinks to zero.

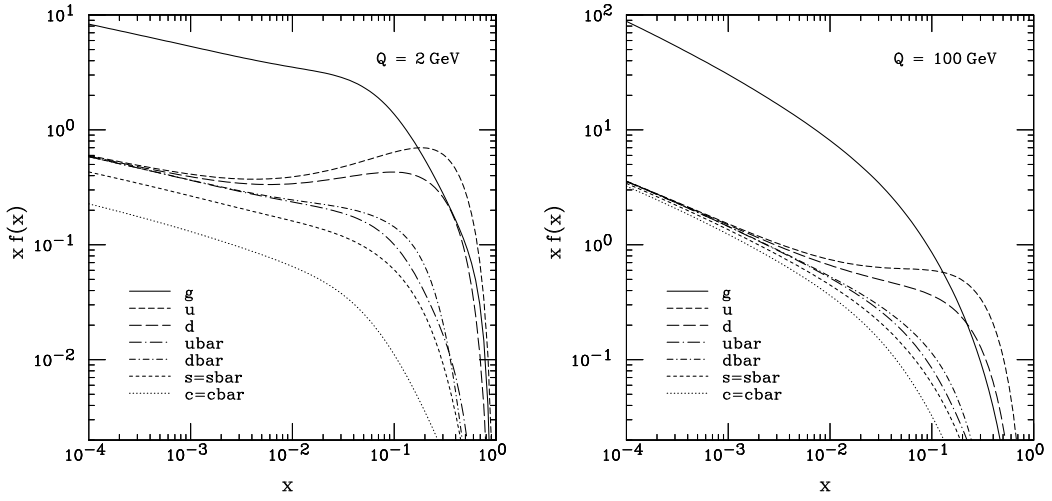
[8] They contribute to the quantum number of the proton.

(they are created and annihilated in the vacuum). These quarks are called the "sea" quarks [9]. Thus, when a parton receives a large amount of energy in a collision, as it moves away from other particles, the binding energy between them increases. When the energy is sufficiently large, a new quark–anti-quark pair is created from vacuum. Experimentally, this process repeats until the parent parton's energy is expended, resulting in not a single parton but collimated sprays of partons moving approximately in the same direction. This process is called parton shower. The interaction of these showers with the detector result in the measured jets. The jets play a very key role in HEP; by examining their signatures in the detector, we can reconstruct the event. Hence, jet reconstruction is a source of systematic uncertainty for our measurement (see section 3.3). An exception to this is the top quark; because of its exceptionally large mass, it decays before hadronization takes place (see section 1.4.3).

Parton Distribution Functions

Hadron colliders produce inclusive collisions; the partons of the hadrons are what interact and hence the initial state of the actual hard scatter interaction is complicated. On the other hand, and thanks to asymptotic freedom, if the momentum of the incoming particles is high enough ($\gg \Lambda_{\mathrm{QCD}}$), it is possible to make the approximation using perturbative QCD (pQCD), that the interaction occurs between two and only two partons, one in each incoming particle. However, in the collisions, the component of the initial momenta parallel to the beamline (the z momenta in CDF) of the partons are unknown. Each parton carries a fraction (x) of the total momentum of the hadron according to a statistical distribution that depends on its type (the gluon or quark flavor) and on the momentum scale (Q); this distribution is known as a Parton Distribution Function (PDF). These functions can not at present be calculated from

[9] Quarks and gluons are referred to as partons, being a "part" of the proton.



(a) $Q = 2$ GeV (b) $Q = 100$ GeV

Figure 1.4: Overview of the CTEQ6M PDFs, from [7].

first principles but must be determined by data. Many data sets from a wide variety of experiments are used for an improved determination of PDFs.

There are several collaborations dedicated to this endeavor, among them the Coordinated Theoretical Experimental Project on QCD (CTEQ) collaboration [10]. The MSU HEP faculty members are pioneer players in this collaboration [11]. Figure 1.4 shows the CTEQ6M PDFs at $Q = 2$ and 100 GeV. The manner in which PDFs are involved in calculating the cross section will be discussed in detail in section 1.5.1. The systematic uncertainties in the PDFs are evaluated in section 4.2.2 .

Initial & Final State Radiation

At the Tevatron, $p\bar{p}$ pairs collide occasionally with a large momentum transfer that is sufficient to produce a $t\bar{t}$ pair. In addition to the $t\bar{t}$ pair, gluons can be radiated off of the incoming or outgoing partons. Given sufficient momentum, these gluons can hadronize and form additional jets. This process is labeled Initial or Final State Radiation (ISR/FSR). It is not yet well-modeled and hence is a source of systematic

[10] Martin-Stirling-Thorne-Watt (MSTW) is another one [6].
[11] As a matter of fact, Joey Huston is the Co-Spokeperson of the CTEQ collaboration.

uncertainty (section 4.2.2). A very interesting aspect of our cross section measurement is that in a way it is a measure of ISR (see section 3.7.2).

1.3.3 The Weak Interaction

Enrico Fermi proposed the first theory of β-decay in 1934 [8]. He proposed that a neutrino is emitted along with an electron in every beta decay. Of course, there was no W boson in Fermi's theory. The weak interaction vertices were given by direct 4-fermion couplings, but from the modern perspective we understand this to be an approximation: that is, Fermi's theory is reproduced whenever the W-mass is much larger than any other energy in the problem (this anyway applies to all the particle except, interestingly, the top quark).

The value of the Fermi constant is set by an accurate measurements of the lifetime of the muon. The lifetime for the muon is given by

$$\tau_\mu = \frac{192\pi^3}{G_F^2 M_\mu^5} \tag{1.12}$$

where G_F is defined by

$$G_F = \frac{\sqrt{2}}{8}\left(\frac{g_{\rm w}}{M_{\rm w}}\right)^2 \tag{1.13}$$

The coupling constant $g_{\rm w}$ is the weak charge $\sqrt{4\pi\alpha_w}$, analogous to $g_{\rm e}$ and $g_{\rm s}$ in QED and QCD respectively. Experimentally, τ_μ is measured to be equal to $(2.197019 \pm 0.000021) \times 10^{-6}$ sec [2]. This then implies:

$$G_F = 1.16637(1) \times 10^{-5}\text{GeV}^{-2} \tag{1.14}$$

which was the early result from Fermi's theory! [12]

Currently, it is possible to independently measure the mass of the W boson, and

[12] In [8], Fermi calculated the constant to be $4 \times 10^{-50}\,\text{cm}^3$.erg which is $\cong 1.0 \times 10^{-5}$ GeV^{-2}.

so we can deduce the strength of the weak coupling constant g_{w}. Equation (1.13) finds $g_{\mathrm{w}} = 0.653 \Rightarrow \alpha_w = 0.034 \simeq \frac{1}{29}$ which is a surprise - the weak coupling strength is $\sim$ five times larger than the electromagnetic strength $\left(\alpha = \frac{1}{137}\right)$!

Weak interactions are indeed weak - but not because the coupling is small. Instead, they are feeble because the gauge bosons that mediate the interactions are very massive compared to all the particles (expect for the top). At low energies, much less than the mass of the weak bosons, the electromagnetic force dominates. For example, the π° can decay electromagnetically into two photons; a $\pi^{\pm}$ decays into lepton and its corresponding anti-neutrino. The longer lifetimes that are observed for the $\pi^{\pm}$ with respect to the π° indicate that the force responsible for the decay is weaker than the electromagnetic force. On the other hand, in the Large Electron-Positron Collider (LEP) experiments for instance - which operated at energies comparable to the mass of the Z Boson - weak interactions were observed to be stronger than the electromagnetic ones.

Today, Fermi's theory has been superseded by the modern-day theory of Electroweak Theory. This model unifies weak and electromagnetic interactions into a more unified whole. This SU(2)×U(1) model undergoes a spontaneous symmetry breaking that gives mass to the particles. This is covered next.

1.3.4 The Electroweak Theory

It was observed experimentally that the weak force violates parity i.e. one could tell whether or not one was observing weak interactions in a mirror. This means that there are two types for each particle that we present in Figure 1.2: a left-handed [13] and a right-handed particle, and the weak interaction only couples to the left-handed particles. Additionally, it was observed that aside from coupling quarks of different flavors (up with down for instance), the weak force also couples quarks of

[13] A particle is said to be left-handed if its spin is opposite to its momentum.

different families or generations (up with strange for example) (see the CKM section below). The simplest gauge group that fulfills such observations is SU(2), where the interaction is only between the left-handed component of the particle and the right-handed component of the antiparticle.

As James Maxwell showed successfully that electricity and magnetism could be unified into a single theory involving a vector field (the electromagnetic field); Glashow, Weinberg and Salam described how it is possible to unify the weak and the electromagnetic forces through the SU(2)×U(1) symmetry. In 1961 Glashow [9] suggested that the two interactions were unified, and in 1967-8 Glashow's model was modified by Weinberg [10] and Salam [11] to make the weak bosons massive. In 1971 't Hooft and Veltman [12] showed that this theory is renormalizable and in 1983 the W and Z bosons were discovered [13].

In their theory, they assumed an $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry which would generate four vector bosons. $\mathrm{U}(1)_Y$ is a symmetry analogous to that of the QED and generates one vector boson B° with a new quantum number, the weak hypercharge Y. On the other hand, $\mathrm{SU}(2)_L$ generates three vector bosons: W^1, W^2 and W^3 with a new quantum number, the weak isospin I_3 [14]. The up-type fermions are assigned $I_3 = +\frac{1}{2}$ and the down-type are assigned $I_3 = -\frac{1}{2}$. With electric charge Q as the quantum number of QED, $Y = 2(Q - I_3)$. Table 1.1 summarizes these quantum numbers. In their work, neutrinos are all left-handed and massless. Recent experimental evidence indicates that neutrinos do have a mass (so right-handed neutrinos should exist as well) [14]. This means that SM needs to be extended in this regard.

The electroweak lagrangian $\mathcal{L}$ is:

$$\mathcal{L} = -\frac{1}{4}W^{\mu\nu}W_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + i\bar{\psi}\gamma^\mu D_\mu\psi \tag{1.15}$$

[14] I_3 is the z-component of the isospin I.

Fermions	1^{st}	2^{nd}	3^{rd}	Q	I_3	Y
Leptons	$\binom{\nu_e}{e}_L$	$\binom{\nu_\mu}{\mu}_L$	$\binom{\nu_\tau}{\tau}_L$	$\binom{0}{-1}$	$\binom{+1/2}{-1/2}$	-1
	e_R	μ_R	τ_R	+1	0	+2
Quarks	$\binom{u}{d}_L$	$\binom{c}{s}_L$	$\binom{t}{b}_L$	$\binom{+2/3}{-1/3}$	$\binom{+1/2}{-1/2}$	+1/3
	u_R	c_R	t_R	-2/3	0	-4/3
	d_R	s_R	b_R	+1/3	0	+2/3

Table 1.1: Quantum numbers in the electroweak theory.

where the covariant derivative is

$$\begin{aligned} D_\mu &= \partial_\mu + ig_W W_\mu \cdot T + ig_Y Y B_\mu \\ &= \partial_\mu + ig_W(W_\mu^+ T^+ + W_\mu^- T^-) + ig_W W_\mu^3 \cdot T_3 + ig_Y Y B_\mu \end{aligned} \tag{1.16}$$

While $g_Y Y B_\mu$ is just a number, $W \cdot T$ is a 2×2 matrix and T_i are the SU(2) generators ($T_i = \frac{1}{2}\sigma_i$) which satisfy [15]

$$[T_i, T_j] = i\epsilon_{ijk} T_k \tag{1.17}$$

The two bosons W^3 and B_μ mix together to form the electromagnetic field A_μ of equation (1.2) and the neutral Z, while the remaining W bosons will combine into the charged $W^\pm$ bosons:

$$\begin{aligned} W_\mu^\pm &\equiv \frac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2) \\ W_\mu^3 &= cos\theta_W Z_\mu - sin\theta_W A_\mu \\ B_\mu &= sin\theta_W Z_\mu + cos\theta_W A_\mu \end{aligned} \tag{1.18}$$

where θ_W is a free parameter of the theory, called the weak mixing angle or Weinberg angle.

Up to this point, the particles are massless; the Higgs mechanism (next paragraph) gives mass to them. After solving the gauge equations of motion, we obtain the

[15] Analogous to QCD, electroweak interactions have a non-Abelian symmetry.

solution:

$$g_{\rm W}\, sin\theta_{\rm W} = g_{\rm Y}\, cos\theta_{\rm W} = g_{\rm e} = e \tag{1.19}$$

This angle clearly relates both the weak and the elecrtomagnetic coupling constants mentioned in the previous sections. Experimentally [2];

$$sin^2\theta_{\rm W} = 0.23116 \pm .00013 \Rightarrow \theta_{\rm W}^{\rm expt} = 28.74^\circ \tag{1.20}$$

The Higgs Mechanism

The lagrangian $\mathcal{L}$ in Equation (1.15) describes only massless bosons and fermions. The problem is not that we don't know how to give a particle mass - terms such as $\frac{1}{2}M^2 B_\mu B^\mu$ or $m\bar{\psi}\psi$ are indeed mass terms. However, adding $\frac{1}{2}M^2 B_\mu B^\mu$ for example would destroy the U(1) gauge invariance of the theory. This in turn renders the theory unrenormalizable. The resolution to this dilemma was found by Higgs in 1963 [15], and then applied to the weak interactions by Salam and Weinberg as described in the previous section. It involves introducing a scalar field/particle that couples to the electroweak bosons in a particular way that the photon remains massless but the Z does not.

The Higgs field is a weak isospin doublet under SU(2) of complex fields, which I will call $\Phi = (\phi^+, \phi^\circ)$. The term $\mathcal{L}_H$ is added to the lagrangian $\mathcal{L}$

$$\mathcal{L}_H = |D_\mu \Phi|^2 - V(\Phi^+\Phi) \tag{1.21}$$

where the scalar potential takes the form:

$$V(\Phi) = \mu^2 \Phi^+\Phi + \lambda(\Phi^+\Phi)^2 \tag{1.22}$$

where λ and μ are real constants. The symmetry is broken because the minimum of

the potential is not at $\Phi = 0$ (the trivial ground state [16]), but rather at $\Phi^{+}\Phi = \upsilon = \frac{\mu}{\lambda}$. This allows us to write $\Phi = (0, \upsilon + h(x))$ i.e. only one of the Higgs field , the h(x) field, has acquired a mass (the Higgs particle), while the other three fields remain massless. By inserting this form for Φ and solving for the equations, the $W^{\pm}$, Z and the Higgs particle acquire the following mass terms:

$$M_{\mathrm{W}} = \frac{g_{\mathrm{W}}\,\upsilon}{2}, \qquad M_{\mathrm{Z}} = \frac{g_{\mathrm{W}}\,\upsilon}{2\cos\theta_{\mathrm{W}}}, \qquad m_{\mathrm{H}} = \sqrt{2}\mu \tag{1.23}$$

yielding the relationship

$$M_{\mathrm{W}} = M_{\mathrm{Z}}\cos\theta_{\mathrm{W}} \tag{1.24}$$

Having the mass of the W and the value of the g_{W}, it's not hard to show that $\upsilon \cong$ 246 GeV. However, the SM makes no prediction as to what m_{H} is - the parameter μ must be input into the model besides the other free parameters. Saying this, it is possible to obtain bounds on m_{H} indirectly from the SM using different precision measurements, such as the W boson mass and the top quark mass. The Higgs boson itself is the subject of extremely intense searches at the Tevatron and the LHC. How the top quark can be used to constrain the Higgs mass is what we shall discuss in section 1.4.3.

Yukawa Coupling

The Higgs mechanism allows the bosons to acquire mass. After all, this was the aim of the whole procedure. On the other hand, fermions acquire their masses directly (we add the mass-term into the equations), but why not make use of the Higgs mechanism to generate masses for them as well? This would provide a unified picture of the origin of mass, in which all particles attain their internal masses through their interactions with the Higgs particle. This is done by adding, for each fermion, a term to the

[16] In Particle Physics, it is called the vacuum expectation value.

lagrangian $\mathcal{L}$ of the form, $\mathcal{L}_{\text{Yukawa}} = -Y_{\text{f}}\bar{f}_{\text{L}}\phi f_{\text{R}}$ where Y_{f} is the Yukawa coupling of the fermion and $\bar{f}_{\text{L}}$ and f_{R} are the left-handed and right-handed components of the fermion field f respectively. After symmetry breaking, each fermion would be coupled to the Higgs; a large coupling of the Higgs to the top becomes the large top quark mass, and a small coupling of the Higgs to the electron becomes the small electron mass.

For the top quark, the Yukawa coupling turns to be very large:

$$Y_{\text{top}} = \frac{\sqrt{2}}{\upsilon} m_{\text{top}} \Rightarrow Y_{\text{top}} \cong 0.99 \tag{1.25}$$

This is a theoretically interesting value, leading to speculation that important new physics may be accessed via top quark by probing the Higgs-top interaction; another reason that triggers us to study its phenomenology!

CKM Matrix

As mentioned before, besides coupling quarks of different flavors, the weak force also couples, with lower probability, quarks of different generations. The amplitudes involved in calculating this probability is given by a 3×3 unitary matrix known as the CKM matrix (for Cabbibo-Kobayashi-Maskawa)[16]:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.26}$$

The current measured values of these entries is [2]:

$$\begin{pmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347^{+0.00016}_{-0.00012} \\ 0.22520 \pm 0.0007 & 0.97345^{+0.00015}_{-0.00016} & 0.0410^{+0.0011}_{-0.0007} \\ 0.00862^{+0.00026}_{-0.00020} & 0.0403^{+0.0011}_{-0.0007} & 0.999152^{+0.000030}_{-0.000045} \end{pmatrix} \tag{1.27}$$

We see empirically that the third generation (t,b) hardly mixes at all with the other two generations. The SM doesn't explain the origin of the values of the entries in the CKM matrix. The best we can currently do is to measure experimentally these values for consistency using the constraints imposed by the SM assumption of the unitary of the CKM matrix [17]. Having V_{tb} much much larger than V_{td} and V_{ts}, we will look only for top decays into bottom quarks as we will explain this in detail later.

1.4 Top Quark Physics Review

Though predicted to exist almost immediately after the discovery of the bottom quark at Fermilab in 1977 [17], the top quark wasn't discovered for 17 more years. The CDF and D∅ collaborations jointly announced discovery of the top quark in 1995 [18], but it wasn't until the recent high luminosity and large data sets of Run II at the Tevatron that precision measurements of the top quark were possible.

In the subsequent subsections I will discuss the top production mechanism, the decay mechanism and the intrinsic properties (such as it's mass, charge, width/V_{tb} and spin correlations). There are many review articles summarizing the latest measurements and studies of top quark. I have particularly consulted these [19, 20, 21, 22].

[17] A rule of thumb definition for the CKM unitary is that quarks do not turn into something else.

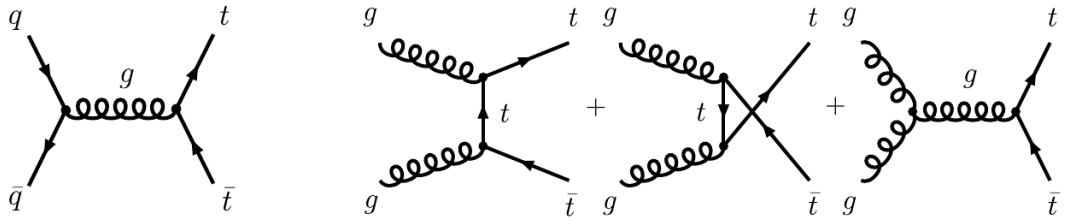


Figure 1.5: LO Feynman diagrams for top quark pair production: (a) quark–anti-quark annihilation and (b) gluon-gluon fusion, from [22].

1.4.1 Top Production Mechanism

The first thing we may ask ourselves about is how top quarks are produced and how they afterword do decay. The production of top quarks is only possible at highest center-of-mass energies ($\sqrt{s}$), set by the scale of top quark mass (m_t). The energies needed for production of top quarks in the SM are currently only accessible at hadron colliders. The Tevatron is operating at$\sqrt{s} = 1.96$ TeV while the LHC currently provides proton-proton collisions at $\sqrt{s} = 7$ TeV.

In the framework of the SM, top quarks can be produced in pairs, $t\bar{t}$, predominantly via the strong interaction, and singly via the electroweak interaction. The Leading Order (LO) Feynmann diagrams [18] for $t\bar{t}$ production are shown in 1.5 .

To produce a $t\bar{t}$, the energy needed at the tree level is $x_1 x_2 \geq 4m_t^2/s$. Assuming $x_1 \approx x_2 = x$ yields as threshold for $t\bar{t}$ production:

$$\langle x \rangle = \frac{2m_t}{\sqrt{s}} \approx \begin{cases} 0.176 & (\text{Tevatron RunII}), \sqrt{s} = 1.96\,\text{TeV} \\ 0.025 & (\text{LHC}), \sqrt{s} = 14\,\text{TeV} \end{cases} \tag{1.28}$$

Since large momentum fractions are required for $t\bar{t}$ production at the Tevatron, and by examining Figure 1.4, the process is dominated by quark–anti-quark annihilation (Figure 1.5(a)) of the valence quarks. This subprocess accounts for $\sim 85\%$ of all

[18] In literature, LO is called the Born level or the Tree level.

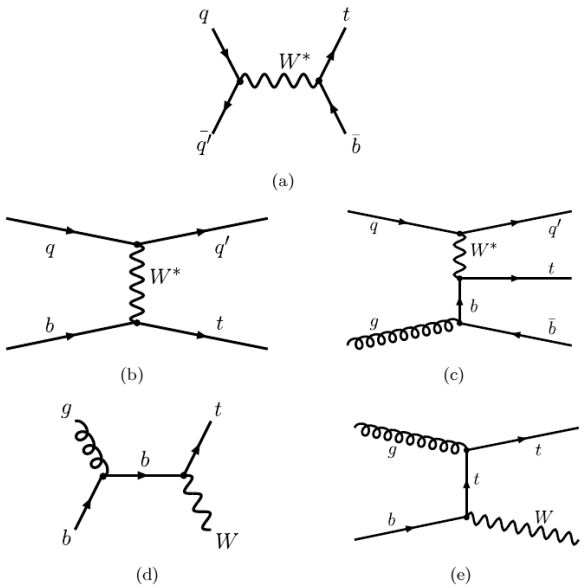


Figure 1.6: Representative Feynman diagrams for electroweak single top quark pair production: (a) s-channel, (b,c) t-channel and (d,e) associated production. The diagrams for single anti-top quark can be obtained by interchanging quarks with anti-quarks, from [22].

$t\bar{t}$ events, with gg being responsible for 15% [23]. At the LHC, gluon-gluon fusion dominates with a contribution of about 90% because a small momentum fraction suffices for $t\bar{t}$ production [24]. This can be verified by recalling that the probability for colliding a quark and its anti-quark is small because the anti-quark can't be a valence quark and thus the annihilation mechanism is suppressed.

Single top quark production can only take place via the weak interaction involving the Wtb vertex (Wts and Wtd vertices are strongly CKM suppressed - as discussed before). The LO Feynman diagrams are shown in Figure 1.6. The t-channel and s-channel modes are the dominant at the Tevatron ($\sim 65\%$ and $\sim 30\%$ respectively); associated W production is not expected to be observable at the Tevatron with the amount of data collected to date.

Electroweak single top production was finally observed, simultaneously, by the CDF collaboration using 3.2 fb^{-1} of data [25], and by the D$\varnothing$ collaboration using

2.3 fb^{-1} of data [26] [19]. The Tevatron has combined these two measurements and reported the first inclusive single top cross section in both the s- and t-channels of $2.76^{+0.58}_{-0.47}$ pb for a top quark of mass 170 GeV [27].

1.4.2 Top Decay Mechanism

After production, top quarks decay very rapidly through the weak interaction into $t \rightarrow Wq$, with q being one of the down-type quarks. The contribution of each quark flavor to the total decay width is proportional to the square of the respective CKM matrix element V_{tq}. Consequently, the decay $t \rightarrow Wb$ is absolutely dominant. The b quark will convert into a B hadron, which will in turn decay into a jet of particles, tagged as a b-jet (see section 3.4).

W bosons decay into two fermions, either leptons (a charged lepton-neutrino pair $l\nu_l$, $l = e, \mu, \tau$, with equal probability per lepton flavor at Born level), or into quark–anti-quark pairs $q\bar{q}$ with $q = u,c$ and $\bar{q} = \bar{d},\bar{s}$ ($W \rightarrow t\bar{b}$ is not kinematically allowed). At LO level, the hadronic decay widths are enhanced over the leptonic modes by a color factor of three (taking the three possible quark color into account), and scaled by the appropriate squared CKM matrix element $|V_{qq}|^2$. In summary, W bosons decay leptonically with a Branching Ratio (BR) of $\approx 1/9$ per lepton flavor and $\approx 1/3$ for each of the hadronic decays ($u,\bar{d}$ and $c,\bar{s}$). For $t\bar{t}$, there are two W bosons in each event. Thus, there are 81 (9×9) possible combinatorics of decays. At higher orders, the QCD corrections to the partial widths (see section 1.4.3 page 29) to quarks enhance the BRs to quarks, causing the BRs to leptons to decrease. In agreement with this prediction, the measured value of the BR for $W \rightarrow l\nu_l$ is (10.80±0.09)% which is less than 1/9, [2]. The resulting possibilities are illustrated in Figure 1.7 .

The four $t\bar{t}$ decay channels are:

- Dilepton channels: Both W bosons decay leptonically ($l\nu_l$, $l = e,\mu$), resulting

[19] 1 barn (b) is equal to 10^{-28} m^2, see section 1.6 .

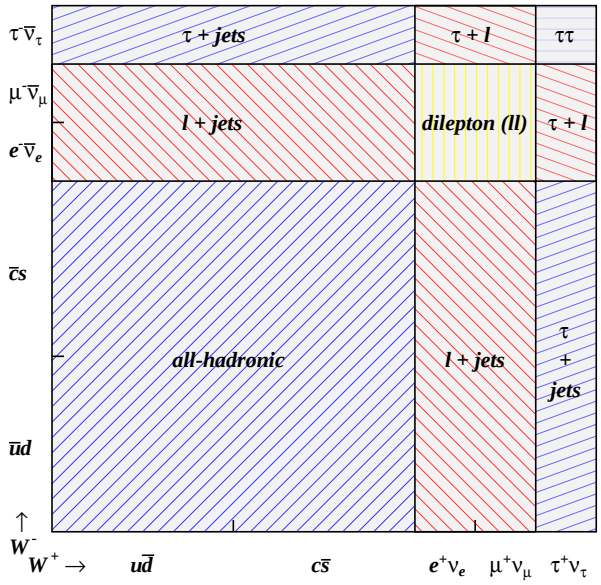


Figure 1.7: The $t\bar{t}$ decay channels via the possible W decays, from [22].

in a final state comprised of two isolated high E_{T} leptons, missing energy ($\not{E}_{\mathrm{T}}$) corresponding to the two neutrinos, and two b-jets. This channel has a BR of approximately 5%. While these channels give samples of the highest $t\bar{t}$ signal purity, they suffer from limited statistics due to the small BR.

- All-hadronic channels: Both W bosons decay to $q\bar{q}$ pairs, resulting in a six-jet final state. With a BR of $\approx 45\%$, this channel yields the highest statistics of $t\bar{t}$ events but also suffers from large background (will be defined later, in section 3.8) from multijet production.

- Lepton + jets channels: One W boson decays leptonically and the other one hadronically, yielding a final state topology of one isolated high E_{T} lepton, $\not{E}_{\mathrm{T}}$ and four jets. These channels which exhibit a BR of approximately 30%, offer (for many analyses) the best trade off between purity of sample and available statistics. For our analysis, we use a part of this channel, using only electrons and muons. This BR should then be encountered in the cross section measurement as will be explained in (Equation 1.53). Figure 1.8 shows a Feynman

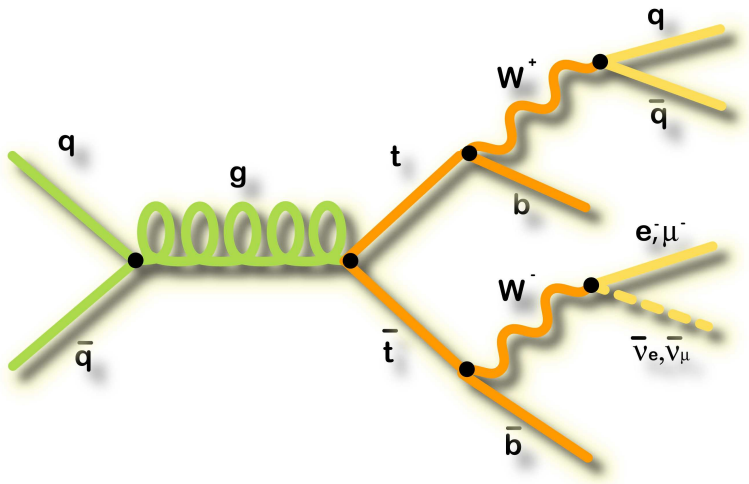


Figure 1.8: A tree level Feynman diagram of $t\bar{t}$ production and subsequent lepton + jets decay, from [28].

diagram for this channel.

- Hadronic τ channels: Final states where at least one W boson yields a charged τ lepton that in turn decays into hadrons (and a neutrino) are called hadronic τ channels, and together comprise a BR of ≈13%. The identification of $\tau \rightarrow$ hadrons makes the final states challenging to reconstruct.

In all of the above final states, two of the jets are b-jets.

From an experimental viewpoint, the issue becomes finding an appropriate set of selection cuts to apply to the decay products in order to reconstruct the $t\bar{t}$ event (identify the jet's E_{T}, the lepton, $\not\!\!E_{\mathrm{T}}$, and the b-tagging). These along with other selection cuts are discussed in detail in chapter 3. The $\tau\tau$ final state remains the only channel that has not yet been explicitly analyzed. The properties of top quarks have been extracted mainly form the first three of the above channels, and especially from the lepton + jet channel.

1.4.3 Intrinsic Properties

To establish that the top quark discovered at the Tevatron is indeed the SM top quark, it is important to confirm its fundamental quantum numbers and to set limits on possible deviations from the SM. In this section, measurements of the top quark's mass, width, spin and electric charge performed thus far at the Tevatron are described. Where appropriate, I comment also on some production and decay properties of the top quark in both the strong and electroweak interactions.

Mass

For a free particle, the physical mass is usually taken to correspond to the pole of their propagator, i.e. the value of the four-momentum squared, $p^2 = E^2 - \vec{p}^2$. Because of confinement, quarks can't exist as free particles and this definition becomes ambiguous. The definition of the top quark pole mass is still possible on an order by order basis in the perturbation theory, but is considered to be intrinsically ambiguous on the order of the confinement scale $\mathcal{O}(\Lambda_{\mathrm{QCD}})$. To determine the top quark mass m_t, one has to find observables measurements of which can be compared to theory predictions which in turn depend on the top quark mass. In HEP, there are two fundamentally different approaches:

(1) Direct reconstruction of top quark decay products: The momenta of the decay products are related to m_t according to

$$m_t(i)^2 = p_t(i)^2 = \left(\sum_j p_j(i) \right)^2 \tag{1.29}$$

where the sum is over all decay products j of the top quark t in a specific event i. Since the squared sum of the four-momenta as given in Equation (1.29) enters in the

denominator

$$p_t^2 - m_t^2 + i m_t \Gamma_t \tag{1.30}$$

of the propagator term, the measured m_t at the Tevatron and LHC taken to be the pole mass.

(2) The indirect constraints are obtained from fits of the SM prediction as a function of the top quark mass to precision measurements of electroweak observables (i.e. non-colored particles). This was the technique used at the LEP.

In perturbation theory, predictions for observables receive contributions (corrections) from loop diagrams (see section 1.5.1). The size of these corrections depends on the values of the mass of the particles in the loops. Of particular importance for SM fits is the dependence of the W boson mass on the top quark and Higgs boson masses. The W boson mass of Equation (1.24) can now be expressed as:

$$\left(\frac{M_{\mathrm{W}}}{M_{\mathrm{Z}}}\right)^2 = \left(1 - \sin^2\theta_{\mathrm{W}}\right)(1 + \Delta r) \tag{1.31}$$

where Δr contains all the one-loop corrections (is 0 at tree level). Contributions to Δr originate from the top quark by the one-loop diagrams shown in Figure 1.9 (a), which contribute to the W and Z masses via:

$$(\Delta r)_{\mathrm{top}} \simeq -\frac{3\,G_F}{8\sqrt{2}\pi^2 \tan^2\theta_{\mathrm{W}}}\, m_{\mathrm{t}}^2 \tag{1.32}$$

Also the Higgs boson contributes to Δr via the one-loop diagrams, shown in Figure 1.9 (b):

$$(\Delta r)_{\mathrm{Higgs}} \simeq \frac{3\,G_F M_{\mathrm{W}}^2}{8\sqrt{2}\pi^2}\left(\ln\frac{m_{\mathrm{H}}^2}{m_{\mathrm{Z}}^2} - \frac{5}{6}\right) \tag{1.33}$$

While the leading m_{t} dependence is quadratic, i.e. very strong, the leading m_{H} dependence is only logarithmic, i.e. rather weak. Therefore the inferred constraints

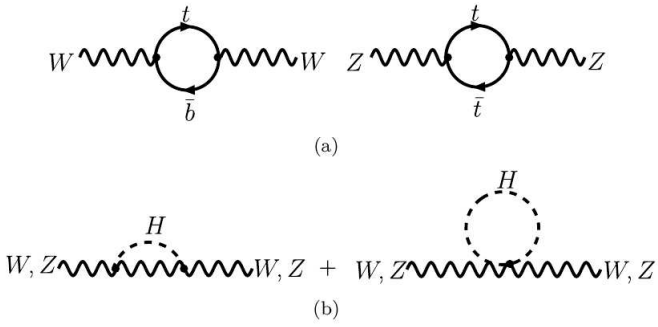


Figure 1.9: (a): Virtual top quark loops contributing to the W and Z boson masses. (b): Virtual Higgs boson loops contributing to the W and Z boson masses, from [22].

on m_{H} are much weaker than those on m_{t}. This was used to successfully predict the top quark mass from the electroweak precision data before it was discovered. This provided confidence in the precision and predictive power of radiative corrections in the SM. Therefore, the SM fit to the electroweak precision data including the direct measurements of the top quark and the W boson masses is used to infer on the mass of the SM Higgs boson. This will also be a service to the LHC experiments which optimize their analysis techniques and strategies for the search for the yet elusive SM Higgs boson in the lower mass range, preferred by the SM electroweak fit.

With up to 5.9 fb^{-1} of data analyzed at CDF, and up to 6.7 fb^{-1} at D$\varnothing$, the Tevatron has recently excluded, at the 95% confidence level, a Higgs boson with a mass between 158 and 175 GeV and between 100 and 109 GeV and showed that the Tevatron experiments are beginning to be sensitive to a low-mass Higgs boson [29]. The 95% confidence level on Higgs boson production is shown in Figure 1.10.

Two main techniques are used in top quark mass analyses at the Tevatron: the Template Method and Matrix Element Method. The first method begins by choosing an observable correlated with the top quark mass. The distribution of this observable is reconstructed in the dataset under consideration. The resulting distribution is then compared in a fit with template contributions from signal (with varying top quark

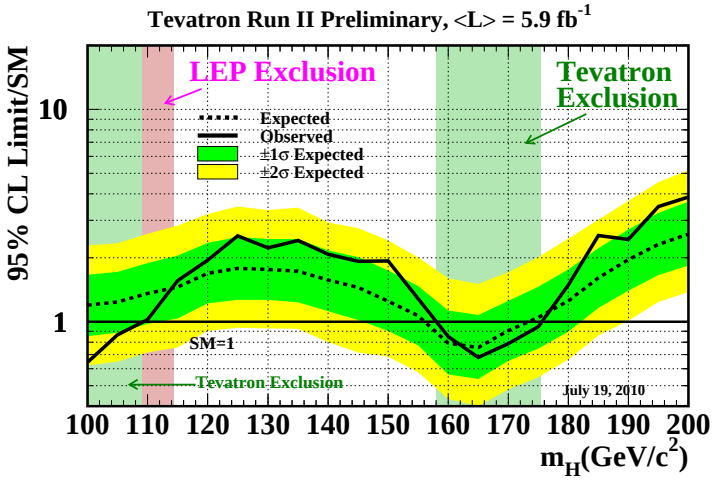


Figure 1.10: Observed and expected exclusion limits for a SM Higgs boson at the 95% confidence level for the Tevatron combined analyses. The yellow and green bands indicate the 68% and 95% probability regions, in the absence of a signal. The Tevatron data exclude a Higgs boson between 158 and 175 GeV and between 100 and 109 GeV at the 95% confidence level, from [30].

masses) and background simulations. The observable most correlated with m_{t} is the reconstructed invariant mass of the $t\bar{t}$ decay products. In the Matrix Element Method and based on the LO matrix elements of contributing signal and background processes, the four-vectors of the reconstructed objects in each event define a probability density as a function of the top quark mass. The total likelihood for the event sample is given as the product of the individual event likelihoods. This method is also referred to as the Dynamical Likelihood Method [31].

The Tevatron updated average mass of the top quark is $m_{\mathrm{t}} = 173.32 \pm 0.56(\mathrm{stat}) \pm 0.89(\mathrm{syst})$ GeV. Adding the statistical and systematic uncertainties in quadratic yields a total uncertainty of 1.06 GeV, corresponding to a relative precision of 0.61% on the top quark mass. Rounding off to two significant digits in the uncertainty, the combination provides $m_{\mathrm{t}} = 173.3 \pm 1.1$ GeV [32]. The input measurements and the resulting Tevatron average mass of the top quark, using up to 5.6 fb^{-1}, are

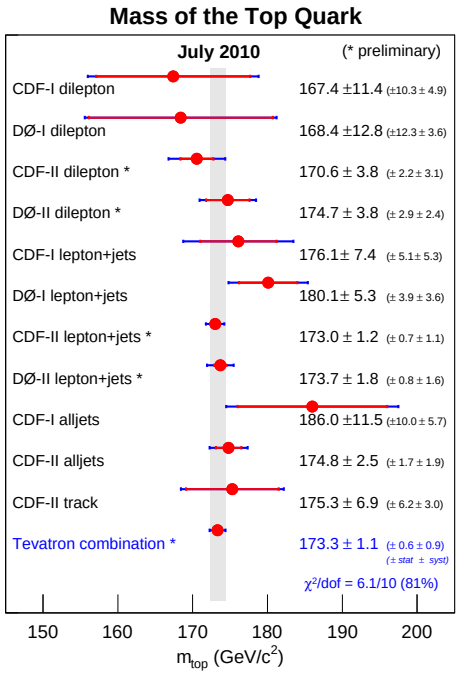


Figure 1.11: Summary of the input measurements and resulting Tevatron average mass of the top quark, from [33].

summarized in Figure 1.11. The combination includes Run I & II measurements using the first three decay channels presented in section (1.4.2).

The overall conclusion is that the top quark mass measurement is in good agreement with indirect constraints valid within the SM and thus there is no significant evidence for phenomena beyond the SM.

Mass Difference

There are three discrete symmetries of interests in HEP. Charge conjugation (C): whenever a particle could undergo certain interactions, so could its anti-particle. Parity (P): if a system is described by $\psi(\vec{x})$, then the transformed wavefunction P$\psi(\vec{x})$ has all coordinates inverted through the origin. Time reversal (T): which transforms the time coordinate (t) into (-t).

It is known that the weak interactions were found to violate C and P maximally.

While the product CP seemed to be conserved, evidence for CP-violation was later found in Kaon decays. On the other hand, the triplet product CPT has never been seen to be violated in any experiment, and in fact, there are several tests that constrain the magnitude of CPT violation in various sectors of the SM [34]. This requires that the particle and its corresponding anti-particle masses be identical. Thus top quark provides a unique opportunity to measure a $(q-\bar{q})$ mass difference free from the QCD effects and to probe CPT-violation to a greater degree than the other sectors of the SM.

CDF presented a preliminary measurement in the lepton plus jet channel using a data sample of 5.6 fb^{-1} and found [35]: $\Delta m_{\rm top} = m_{\rm t} - m_{\bar{\rm t}} = -3.3 \pm 1.4\,{\rm stat} \pm 1.0\,{\rm syst\,GeV} = -3.3 \pm 1.7\,{\rm GeV}$, consistent with CPT invariance.

D$\varnothing$ performed the measurement in the same channel using ~ 1.0 fb^{-1} and found [36]: $\Delta m_{\rm top} = m_{\rm t} - m_{\bar{\rm t}} = +3.8 \pm 3.4\,{\rm stat} \pm 1.22\,{\rm syst\,GeV} = +3.8 \pm 3.8\,{\rm GeV}$, also consistent with CPT invariance. These measurements are statistically limited though.

Width

The total width, Γ_t, (also referred to as decay rate) of the top quark is determined in the SM once its mass is specified. It is defined as

$$\Gamma_t = \frac{\#\text{ of decays per unit time}}{\#\text{ of particles}} \tag{1.34}$$

giving it the units of sec^{-1} as expected. Because the top quark mass is large, so is Γ_t, and consequently its lifetime τ_t, $\sim 1/\Gamma_t$, is extremely short. The partial width is defined for only a specific decay channel, for example $t \longrightarrow Wb$.

The Γ_t at the LO is given by:

$$\Gamma_t^o = \frac{G_F\, m_t^3}{8\pi\sqrt{2}} \tag{1.35}$$

The Γ_t in the SM, including first-order QCD corrections, can be expressed as follows [22]:

$$\Gamma_t = \Gamma_t^o |V_{tb}|^2 \left(1 - \frac{m_{\rm W}^2}{m_t^2}\right)^2 \left(1 + 2\frac{m_{\rm W}^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right], \qquad (1.36)$$

where the above formula assumes $m_b^2/m_t^2 \to 0$, $m_t^2 \gg m_{\rm W}^2$ and ignore corrections of order $\mathcal{O}(\alpha_s^2)$. The other SM decays, $t \longrightarrow Wd$ and $t \longrightarrow Ws$, contribute negligibly to the total decay width $\Gamma_t = \sum_q \Gamma_{tq}$ because of proportionality to $|V_{\rm td}|^2$ and $|V_{\rm ts}|^2$. Γ_t is 1.02, 1.26, 1.54 GeV for the top quark masses of 160, 170, 180 GeV respectively.

The resulting lifetime of the top quark $\tau_{\rm t} = \hbar\, \Gamma_t^{-1} \approx (1.3\ {\rm GeV})^{-1}$ is approximately 5.10^{-25} s, and significantly shorter than the hadronization time-scale $\tau_{\rm had} \approx \hbar\, \Lambda_{\rm QCD}^{-1} \approx (0.2\ {\rm GeV})^{-1} \approx 3.10^{-24}$ s. This means that, unlike lighter quarks, the top quark does not hadronize before decaying but rather decay as a free particle. In particular, there can be no $t\bar{t}$ bound states, toponium, ($\Gamma_{t\bar{t}} \sim 2\ \Gamma_t$). Consequently, this means that the intrinsic properties of the top quark (spin, charge, etc) can be measured directly through reconstruction of its decay products before color confinement occurs.

Nevertheless, although the top quark can generally be considered as a free quark, residual non-perturbative effects associated with hadronization should still be present in the top quark events, and the fragmentation and hadronization processes will be influenced by the color structure of the hard interaction.

CDF used 955 pb^{-1} of data in the lepton+jet channel and employed a kinematic fit to obtain the $m_t^{\rm recon}$ distribution. The shape of the $m_t^{\rm recon}$ distribution is sensitive to the Γ_t and a template method of a nominal $m_{\rm t}$ of 175 GeV was employed to measure for the first time the Γ_t. Figure 1.12 shows 95% C.L in $\Gamma_t^{\rm fit}$; the fitted value from data. CDF allowed negative $\Gamma_t^{\rm fit}$ values that represent $m_t^{\rm recon}$ narrower than the nominal due to statistical fluctuations. The $\Gamma_t^{\rm fit}$ = -4.8 GeV corresponds to a limit of the $\Gamma_{\rm t} < 13.1$ GeV at 95% C.L. which corresponded to $\tau_{\rm t} > 5.2\ 10^{-26}$ s [37]. These

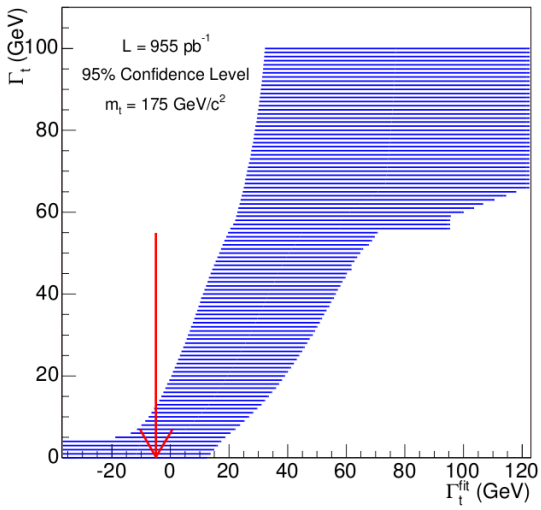


Figure 1.12: The confidence band in $\Gamma_t^{\rm fit}$ for a 95% C.L. The fitted value from the data is indicated by an arrow which corresponds to a limit of $\Gamma_t < 13.1$ GeV, from [37].

limits would improve if the nominal $m_{\rm t}$ used in the template was closer to the current Tevatron average recorded in Figure 1.11.

Since the SM Γ_t is of the order mentioned above, the production and decay vertices of the top are separated by $\mathcal{O}(10^{-16})$m, which is orders of magnitude below the spatial-experimental resolutions at both the Tevatron and the LHC. Consequently, a direct measurement technique of the Γ_t as the one described above will be limited by detector resolution.

However, it is possible to extract Γ_t using the measured cross section for the electroweak production of the single top quarks, as addressed before in [25, 26] [20] . By using the 2.3 fb^{-1} of data in the t-channel [38], D$\varnothing$ measured $\Gamma_t = 1.99^{+0.69}_{-0.55}$ GeV which is equivalent to a lifetime of $\tau_{\rm t} = 3.3^{+1.3}_{-0.9}\,10^{-25}$ s in agreement with the SM prediction. This is displayed in Figure 1.13

[20] In fact, in both literature, and since the single top production rate is proportional to the CKM element $|V_{tb}|^2$, the $|V_{tb}|$ was determined using the measured cross section. CDF obtained $|V_{tb}| = 0.91 \pm 0.11$ and $|V_{tb}| > 0.71$ at the 95% C.L. , while D$\varnothing$ found $|V_{tb}| = 1.07 \pm 0.12$ and $|V_{tb}|$ 0.78 at the 95% C.L. .

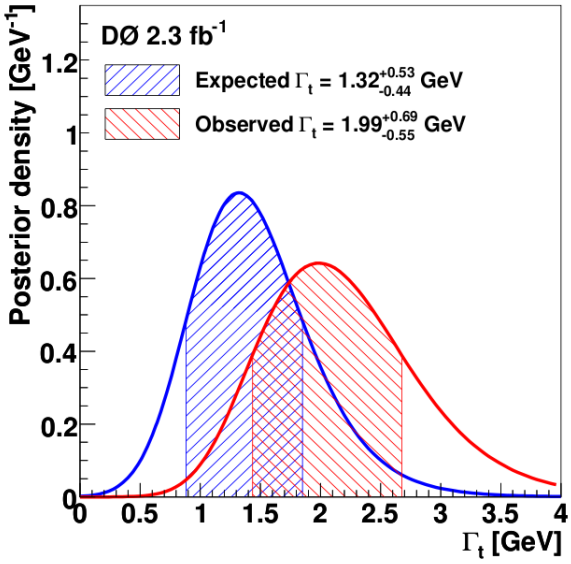


Figure 1.13: The expected and observed Bayesian probability density for the Γ_t. The most probable value for Γ_t is defined by the peak of the probability density function. The hatched area represent one standard deviation around the most probable value, from[38].

Helicity of the W boson

The top quark decays, before forming a bound state, via the left-handed charged current weak interaction $t \longrightarrow Wb$. This is reflected in the observed helicity states of the W boson, which can be exploited to directly probe the couplings at the tWb interaction vetrex.

The emitted b quark can be regarded as massless compared to the top quark, and hence expected to be predominantly of negative helicity (left-handed). The emitted W boson, being a massive spin-1 particle, can assume any of the three helicities: one longitudinal (W_0) and two transverse states (W_-, left-handed and W_+, right-handed). To conserve angular momentum in the $t \longrightarrow W^+ b$ decay, the spin projection of the W boson onto its momentum must vanish if the b quark's spin points along the spin of the top quark, while a left-handed W boson is needed if the b quark's spin points opposite to the spin of the top quark. In the limit of a massless b quark, a

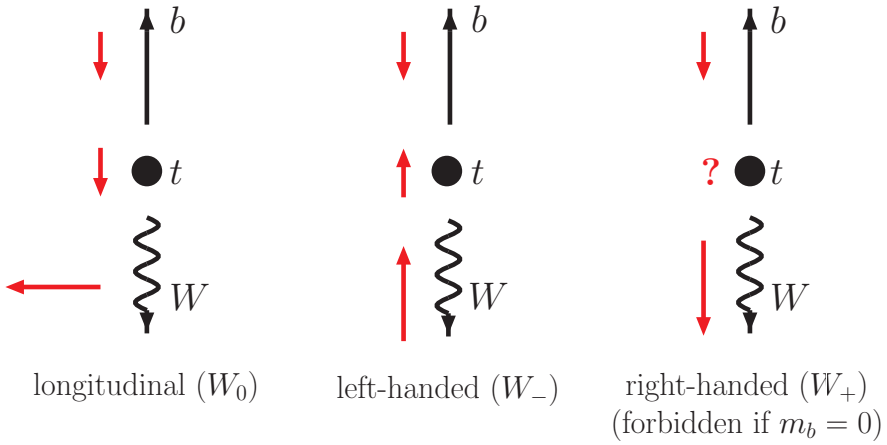


Figure 1.14: Angular momentum conservation in top quark decay does not allow right-handed W bosons when b quarks are assumed massless, from[22].

right-handed W boson can not contribute to the decay, as illustrated in Figure 1.14. For the decay of an anti-top quark, a left-handed W boson is forbidden.

At LO, the expected fractions of decays with different W boson helicities, taking the finite b quark mass into account, are given by:

$$f_0 = \frac{\Gamma_0}{\Gamma_t} \approx \frac{1}{1+2x^2} \tag{1.37}$$

$$f_- = \frac{\Gamma_-}{\Gamma_t} \approx \frac{2x^2}{1+2x^2} \tag{1.38}$$

$$f_+ = \frac{\Gamma_+}{\Gamma_t} \approx y^2 \frac{2x^2}{(1-x^2)^2(1+2x^2)} \tag{1.39}$$

where the scaled masses $x = m_{\mathrm{W}}/m_t, y = m_b/m_t$ were used. Inserting $m_{\mathrm{t}} = 175$ GeV, $m_W = 80.4$ GeV, and $m_b = 4.8$ GeV, the partial helicity rates are found to be [22]: $f_0 = 0.703, f_- = 0.297, f_+ = 0.00036$. NLO corrections lower the f_0 by 1.1% and increase the f_- by 2.2% while the f_+ remains unchanged at the per mill level. Consequently, any observation of f_+ at the percent level would signal the presence of physics beyond the SM.

Thus far, four analysis techniques have been deployed at the Tevatron in $t\bar{t}$ events

to extract the W boson helicity fractions. This is explained in detail in [22]. All the measurements are compatible with the SM expectation. The results of a couple of analyses combined with 1.9 fb^{-1} of data at CDF found [39]:

$f_0 = 0.66 \pm 0.16(\text{stat}) \pm 0.05(\text{syst})$, $f_+ = -0.03 \pm 0.06(\text{stat}) \pm 0.03(\text{syst})$.

D$\varnothing$ has investigated the datasets of RunII with a total luminosity of 2.2–2.7 fb^{-1} and found [40]:

$f_0 = 0.490 \pm 0.106(\text{stat}) \pm 0.085(\text{syst})$, $f_+ = -0.110 \pm 0.059(\text{stat}) \pm 0.052(\text{syst})$.

$t\bar{t}$ Spin Correlations

The SM predicts that the top quark decays before its spin flips. This is in contrast with the lighter quarks, which are depolarized by QCD interactions long before they decay. The spin of a heavy quark is flipped by its chromomagnetic moment, which is inversely proportional to its mass, m_Q. The spin-flip time is therefore proportional to $m_{\text{Q}}/\Lambda^2_{\text{QCD}}$. The spin-flip time of the top quark is estimated [41] to be $\approx (1.3\text{ MeV})^{-1}$. This is much longer than the anticipated top quark lifetime given by Equation (1.36) $\approx (1.3\text{ GeV})^{-1}$. Because of this, the top quark assures that its spin information is passed on to its decays products, and is reflected in their corresponding angular distributions. This provides experimental confirm that the top can indeed be considered as a free quark, thereby setting an upper bound on its width, which is proportional to the combination of the CKM elements $|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2$.

Although Tevatron and LHC use unpolarized beams and therefore produced unpolarized $t\bar{t}$ pairs, their spins are expected to have a strong correlation and point along the same axis [21] in the $t\bar{t}$ rest frame on an event-by-event basis. This is what called the Helicity Basis [42]. In this basis, when a top quark's spin is oriented in the same direction as the quark's direction of motion in the $t\bar{t}$ rest frame, it is said

[21] The spin is measured along a particular axis called the spin quantization axis. If the top quark's spin is oriented in the same (opposite) direction as the quantization axis, it has a spin + $\frac{1}{2}$ (- $\frac{1}{2}$). The choice of the quantization axis is referred to as a choice of basis.

to be a right-handed quark (t_R), while when the spin is oriented against the quark's direction of motion, it is called a left-handed top quark (t_L).

CDF has recently performed , in the helicity basis, the first measurement of the $t\bar{t}$ spin correlation [43]. I'll briefly discuss the theory and summarize the results in turn.

The spin correlation is dependent on the initial particles in the collision which created the $t\bar{t}$ pair and on the energy of the $t\bar{t}$ pair. Because the initial particles will have a particular angular momentum, which is conserved, there will be a correlation between the helicities of the $t\bar{t}$ pair in the event. There are four possible helicity state for $t\bar{t}$ pair: $t_R\bar{t}_R, t_L\bar{t}_L, t_L\bar{t}_R, t_R\bar{t}_L$. Two of these states have $t\bar{t}$ pairs with the same helicity and two have $t\bar{t}$ pairs with opposite helicites. In [43], $F_{\rm OH}$ is defined to be the fraction of opposite helicity $t\bar{t}$ pairs:

$$F_{\rm OH} = \frac{\sigma(t_L\bar{t}_R) + \sigma(t_R\bar{t}_L)}{\sigma(t_R\bar{t}_R) + \sigma(t_L\bar{t}_L) + \sigma(t_L\bar{t}_R) + \sigma(t_R\bar{t}_L)} = \frac{N_o}{N_o + N_s} \tag{1.40}$$

where N_o and N_s are the numbers of $t\bar{t}$ with opposite and same helicity respectively. The same helicity fraction $F_{\rm SH}$ can be defined similarly, and it is apparant that $F_{\rm OH}$ + $F_{\rm SH} = 1$. The $t\bar{t}$ spin correlation coefficient κ is defined as:

$$\kappa = \frac{N_o - N_s}{N_o + N_s} \tag{1.41}$$

The coefficient κ can vary between +1, a fully negative correlation, and κ = -1 for a fully positive correlation. If there were no $t\bar{t}$ spin correlation, $F_{\rm OH} = 0.5$, then $\kappa = 0$.

As described in Section (1.4.1), $t\bar{t}$ production is dominated by $q\bar{q}$ annihilation, where the interaction is mediated by a gluon with spin 1 (J=1), so that the total angular momentum of the $t\bar{t}$ pair must be 1. Close to the energy threshold [22] for $t\bar{t}$ production, the $t\bar{t}$ pair does not have an orbital angular momentum, then conservation of momentum requires that $t\bar{t}$ must spin in the same direction. Thus, $t\bar{t}$ pair will be

[22] The kinematic threshold = $2m_{\rm t} \approx 350$ GeV.

in a $^3S_1$ spin state with three eigenstates: $|++\rangle, \frac{1}{\sqrt{2}}[|+-\rangle + |-+\rangle], |--\rangle$. In the helicity basis, since the two quarks move oppositely in the $t\bar{t}$ rest frame, they will have opposite helicities. Thus, at the threshold, two of the three possible states for the $t\bar{t}$ pair will have opposite helicities (67%). The $t\bar{t}$ pair produced by the gg fusion tends to have the same helicity, since in this case the total angular momentum of the $t\bar{t}$ is (J=0, J_z=0). Hence, and at the threshold, the $t\bar{t}$ will be in a $^1S_0$ spin state with only one eigenstate: $\frac{1}{\sqrt{2}}[|+-\rangle - |-+\rangle]$. Since $t\bar{t}$ production at the Tevatron proceeds manily via $q\bar{q}$ annihilation, as opposed to the LHC where main contribution comes from gg fusion, the observable correlation will have opposite signs at the two colliders.

As a spin-$\frac{1}{2}$ fermion, and at LO calculation, the SM predicts an opposite helicitiy fraction $F_{\rm OH} = 0.70$, corresponding to a $t\bar{t}$ spin correlation coefficient $\kappa = 0.40$. NLO calculations predict $\kappa = 0.352$. Any deviation of the measured κ from the theoretical predictions would indicate physics beyond the SM.

Utilizing 5.3 fb^{-1} of CDF data, a measurement of the $t\bar{t}$ spin correlation coefficient κ has been performed in the lepton plus jet channel by studying the distributions of the helicity angle [23] for the decay product of the $t\bar{t}$. This is shown in Figure 1.15. The opposite helicity fraction was measured to be $F_{\rm OH} = 0.74 \pm 0.24_{\rm stat} \pm 0.11_{\rm syst}$ which yielded a value for the $t\bar{t}$ spin correlation coefficient $\kappa = 0.48 \pm 0.48_{\rm stat} \pm 0.22_{\rm syst}$, [44]. The experimental difficulties of distinguishing the down-type jets from other light jets can be avoided by focusing on the dilepton channel, where the charged down-type leptons are clearly identified. In the off-diagonal basis [24], and using 2.8 fm^{-1} of data, CDF found $\kappa = 0.320 \pm 0.545_{\rm stat} \pm 0.775_{\rm syst}$ using the dilepton channel [46]. These results are consistent with the SM predictions though they are statistically limited.

As opposed to the unpolarized $t\bar{t}$ pair produced via the strong interaction, single

[23] It is the angle between the direction of motion of that decay product (in the top quark rest frame) and the direction of motion of the top quark (in the $t\bar{t}$ rest frame).

[24] The spin quantization axis is the direction of motion of the incident colliding particles.

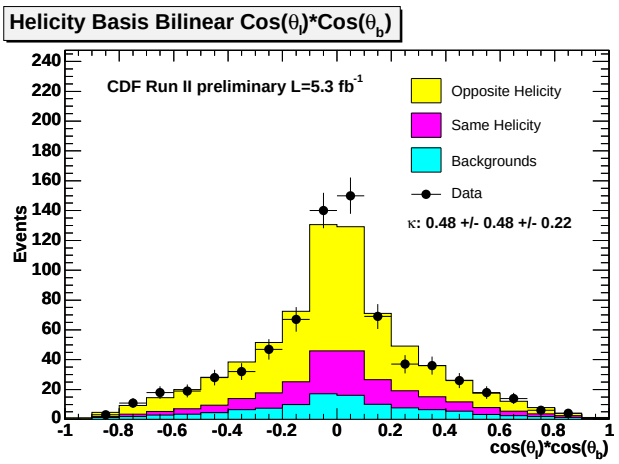


Figure 1.15: The distribution of the helicity basis angles for the decay product of the $t\bar{t}$, from [45].

top quarks produced via the weak interaction are expected to be highly left-handed polarized. This polarization will be reflected in the kinematic distributions of the top quark decay products. Since the W boson couples solely to left-handed fermions, single top quarks are $\approx 100\%$ polarized along the direction of the down-type quark i.e. the spin of it is always correlated with the direction of the down-type quark in the event [22]. Unfortunately, there has been no spin polarization measurement as yet at the Tevatron, although it would greatly benefit from the increase in data already available.

Gluon Fusion Fraction in $t\bar{t}$ Production

As mentioned before in Section (1.4.1), $t\bar{t}$ pairs at the Tevatron are produced dominantly through $q\bar{q}$ annihilation, while about 15% of $t\bar{t}$ pairs are predicted to be produced via gg fusion. Because of uncertainties in the gluon PDF, the prediction of this fraction has a large ambiguity and will change by up to a factor of 2 (from 10% to 20%) [47]. Hence, measurement of the gg fraction can assist in reducing the uncertainties in the gluon's PDF within protons as well as tests the QCD prediction.

There are two ways allow us to separate the two different production mechanism

and to measure their relative contributions.

As mentioned in the previous section, the $t\bar{t}$ pair of the gg fusion will be in a $^1S_0$ spin state, while the $q\bar{q}$ annihilation yields $^3S_1$ spin state. This difference manifests itself efficiently as an azimuthal correlation of charged leptons in the $t\bar{t}$ dilepton channel. Utilizing the difference in these azimuthal angles between the two opposite charged leptons in the dilepton channel would enable us to distinguish $t\bar{t}$ pair produced via gluon-fusion from $q\bar{q}$ annihilation. With a total integrated luminosity of 2.0 fb^{-1} of dilepton events [48], CDF found the fraction of gg fusion $\mathcal{F}_{gg} = 0.53^{+0.36}_{-0.38}$. This is statistically limited and can not be used to restrict the theoretical uncertainties mentioned above.

Alternatively one can discriminate between the similar final state signatures of $gg \rightarrow t\bar{t}$ and $q\bar{q} \rightarrow t\bar{t}$ by taking advantage of the characteristic differences between soft gluons radiated by quarks and gluons. QCD predicts that the average multiplicity of any type of object in a gluon jet should become $\frac{9}{4}$ times greater than that in a quark jet. As a consequence of the higher multiplicity, the fragmentation function of a gluon jet would be expected to be softer, i.e. more concentrated at low values of the momentum fraction x, than that of a quark jet. Another characteristic difference is that gluon jets should be broader than quark jets [49]. In [50], CDF used the low momentum track analysis and found a gg fusion fraction in lepton plus jet channel of $\mathcal{F}_{gg} = 0.07 \pm 0.14(\text{stat}) \pm 0.07(\text{syst})$, corresponding to an upper limit of 0.33 at 95% C.L., in 0.955 fb^{-1} of data. This is in agreement with the SM prediction though it is still statistically limited.

A third analysis done by CDF had used the kinematics of the production and decay of the $t\bar{t}$ to differentiate the two production mechanisms. The kinematic variables are then used to train a neural network (NN) to increase the sensitivity of the method. This analysis relies on Monte Carlo (MC). A full description of the NN method can be found here [51]. The low-momentum tracks and the NN methods are very

complementary. Firstly, the physics is complementary in that the number of tracks is an indirect measurement versus the NN analysis which use very precise kinematics of the different processes. Secondly, the experimental techniques complement each other in that one is data driven and one is MC driven (the NN). CDF combined the two measurement to increase the sensitivity of both. Using a total integrated luminosity of 955 pb^{-1} CDF found $\mathcal{F}_{gg} = 0.07^{+0.15}_{-0.07}$ [52].
D$\varnothing$ did not measure the fraction up to date.

The Electric Charge

The electric charge of the top quark predicted in the SM to be $+\frac{2}{3}$ in units of e. An experimental determination of the top quark electric charge serves as a check of SM consistency. So far the assumption is made that $t \to W^+b$ and $\bar{t} \to W^-\bar{b}$. However, it is possible that what is being observed at the Tevatron is an exotic top-like quark that decays to $t \to W^-b$ and $\bar{t} \to W^+\bar{b}$. Such a possibility has been put forward in [53]. In the SM, t is the SU(2) partner of the left-handed b quark and the right-handed b is a singlet. In [53] such a particle can be found in a four generation model, Exotic Model (EM), and the b is allowed to mix with a heavy quark (Q_1) of charge -$\frac{1}{3}$ whose doublet partner (Q_4) has charge of -$\frac{4}{3}$. In the EM, (Q_4) has the reported mass at the Tevatron while the left-handed real top quark has a mass around 274 GeV and would thus far have escaped detection. It is worth noting that the precision electroweak data can be satisfactorily described by this model.

In order to determine the charge of the top quark, one can either measure the charge of its decay products, in particular of the b-jet via jet charge techniques, or investigate photon radiation in $t\bar{t}$ events. The later method is based on measuring the cross section of $p\bar{p} \to t\bar{t}\gamma$, which is proportional to the square of the top charge. However, studies show that Tevatron needs around 20 fb^{-1} of data to measure the top charge at 95% C.L. using $t\bar{t}\gamma$ events [54]. For the former technique, the three

ingredients needed to establish the top charge are the charge of the W, the flavor of the b-jet and the most important the Wb pairing fraction: W^+b (corresponds to SM) or $W^+\bar{b}$ (corresponds to EM).

Likelihood methods were used at the Tevatron to test the consistency of the data with the SM and the EM scenario. Using 2.7 fb^{-1} [55], the CDF data strongly favored the SM over the EM by excluding the exotic -$\frac{4}{3}$ charged quark interpretation at 95 % C.L. . Using 370 pb^{-1} [56], D$\varnothing$ data confirmed the SM and and excluded the EM at the 92 % C.L. .

Forward-Backward Charge Asymmetry

At LO in QCD, $t\bar{t}$ production is symmetric under charge conjugation, implying it does not discriminate between top and anti-top quarks. Considering that the initial $p\bar{p}$ state at the Tevatron is not an eigenstate of charge conjugation, this symmetry is a coincidence. However, at NLO, radiative corrections involving either virtual or real gluon emissions lead to a small charge asymmetry; an excess of top over anti-top quarks. This asymmetry has its origin in $q\bar{q}$ mechanism from interference between amplitudes that are symmetric and antisymmetric under the exchange of top and anti-top quarks while the gg mechanism remains charge symmetric. The real emission corrections tend to push the top quark backward, opposite to the proton beam direction, while the virtual corrections push it forward, in the direction of the proton beam [19]. Due to this, the charge asymmetry can be interpreted as a top quark forward-backward charge asymmetry; A_{fb} .

A_{fb} is frame dependent and can be defined either in the laboratory frame $p\bar{p}$ or in the parton rest frame $t\bar{t}$. The latter frame is more dificult to experimentally reconstruct but the asymmetry in the former is predicted to be less. At NLO, $A_{fb}^{p\bar{p}}$ and $A_{fb}^{t\bar{t}}$ are expected to be 0.051(6) and 0.078(9) respectively [57].

In the $p\bar{p}$ frame, we look at the angle between the top quark and the proton

direction and measure the rapidity $y_{\rm t}$. However, both CDF and D$\varnothing$ have measured A_{fb} using the lepton plus jet channel. This makes it problematic to use the $y_{\rm t}$ distribution since acceptances are different for $y_{\rm had}$ and $y_{\rm lep}$. We know on the other hand which quark (top or anti-top) decayed hadronically by the charge of the lepton in the event. This means that $y_{\rm had}$ is either $y_{\rm t}$ or $-y_{\bar{\rm t}}$. Because top quark and anti-top quark differs by their charge, a top quark moving forward is equivalent to an anti-top quark moving backward. This allows us to write $y_{\rm t} = -q.y_{\rm had}$, where q is the lepton charge. With N as the number of events with a given rapidity, A_{fb} can hence be written as [58]:

$$A_{fb}^{p\bar{p}} = \frac{N(-q.y_{\rm had} > 0) - N(-q.y_{\rm had} < 0)}{N(-q.y_{\rm had} > 0) + N(-q.y_{\rm had} < 0)} \tag{1.42}$$

In the $t\bar{t}$ rest frame [58]:

$$\begin{aligned} y_{\rm top}^{\rm rest} &= y_{\rm top}^{\rm lab} - y^{\rm t\bar{t}} \\ &= y_{\rm t} - \frac{1}{2}(y_{\rm t} + y_{\bar{\rm t}}) \\ &= \frac{1}{2}(y_{\rm t} - y_{\bar{\rm t}}) \\ &= \frac{1}{2}\Delta y_{\rm top} \\ &= \frac{1}{2}q\Delta y \end{aligned} \tag{1.43}$$

$$A_{fb}^{t\bar{t}} = \frac{N(q\Delta y > 0) - N(q\Delta y < 0)}{N(q\Delta y > 0) + N(q\Delta y < 0)} \tag{1.44}$$

CDF has measured recently the asymmetry in both frames using 5.3 fb^{-1} and found [59] $A_{fb}^{p\bar{p}} = 0.150 \pm 0.050_{\rm stat} \pm 0.024_{\rm syst}$ and $A_{fb}^{t\bar{t}} = 0.158 \pm 0.072_{\rm stat} \pm 0.017_{\rm syst}$. This result is discrepant at the level of three standard deviations (3 σ) for a Gaussian distribution; it neither confirms nor rejects the SM, nor beyond SM, either. In 4.3 fb^{-1} of data collected by D$\varnothing$, the asymmetry in the $t\bar{t}$ frame was found to be [60] $A_{fb}^{t\bar{t}} = 0.08 \pm 0.04_{\rm stat} \pm 0.01_{\rm syst}$. In contrast to the CDF results, this number presents the raw measured value without being corrected for acceptance and resolution effects.

In the dilepton channel, CDF aimed very recently to measure the equivalent quantity in the $t\bar{t}$ frame. Using 5.1 fb^{-1}, CDF found the asymmetry to be [61] $A_{fb}^{t\bar{t}} = 0.417 \pm 0.148_{\text{stat}} \pm 0.053_{\text{syst}}$.

It should be noted that $A_{fb}^{p\bar{p}}$ vanishes at the LHC due to the symmetric initial state, in contrast to the Tevatron. The dominance of $t\bar{t}$ production via the charge symmetric gg fusion reduces the observable charge asymmetry at the LHC.

1.5 Top quark Cross Section

What quantity should the experimentalist measure and the theorist calculate in a scattering process? Is it the exclusive scattering cross section or the inclusive (total) scattering cross section or the differential scattering cross section? Well, the cross section σ for the former is a mere number (the number of events scattered given a certain amount of data), while for a theorist, it is a mere number multiplied by a phase factor; this number is related to the invariant amplitude of a given initial state scattering to a given final state (exclusively). The amplitude can be imagined as one element in a matrix, usually called the scattering matrix (S-matrix), which includes the amplitudes of all possible initial and final states (inclusively), and so is usually called the matrix element $\mathcal{M}$. The matrix element contains only the dynamical information; we calculate it by evaluating the relevant Feynman diagrams of the process. The phase space factor contains the kinematical information; it depends on the masses, energies and momenta of the particles.

1.5.1 Perturbative QCD (pQCD)

Current QCD is unable to calculate the cross section exactly; instead, a perturbative expansion is made in powers of the running α_s. These perturbative terms are conveniently represented by Feynman diagrams. To make optimal use of the experimental

measurements requires theoretical predictions of similar precision. The diagrams with the fewest possible number of vertices for a certain process represent the LO term of the expansion ($\mathcal{O}(\alpha_s^2)$). At hadron colliders, these are often insufficient to describe the process, and including higher-order corrections in QCD is mandatory.

To date, theoretical predictions for $t\bar{t}$ production are based on NLO calculations of the total cross section (inclusive), differential cross section and the forward-backward asymmetry (see previous section). However, the NLO computations suffer from theoretical uncertainties larger than 10% [62], both for Tevatron and LHC. These uncertainties are partly due to our imperfect knowledge of the PDFs and partly to the truncation of the perturbative series in α_s, which introduces a dependence on the unphysical renormalization (μ_R) and factorization (μ_F) scales into the physical predictions. The later uncertainty is typically reduced by including more terms in the perturbative series, and for this reason the calculation of the cross section to next-to-next-to-leading order (NNLO) has been an area of active research. However, due to the complexity of the calculations, complete results for the cross section at NNLO are not yet available. I discuss the systematics of these scales briefly.

One of the unpleasant feature of the NLO diagrams is that they contain closed loops. The calculations require an integral to be performed over the momenta of the particles in the loop, and the integral often diverges. This divergence [25] can be dealt with by a process called renormalization, in which a renormalization scale (μ_R) is introduced ad hoc to truncate the integral before it diverges. This gives a finite result to the calculation that agrees well with experiment for many processes. The running α_s introduces a problem for practical calculations because the energy at which the α_s is calculated must be chosen. This causes the uncertainty in the cross section calculations and the generation of simulated events as well.

Another kind of divergence is called an infrared divergence. In QCD, and at low

[25] Usually called ultraviolet divergence because they occur for very large momentum scales.

momentum, the α_s becomes larger than one. In this case, each successive term in the perturbation is larger than the one before it, making perturbation diverges without giving a valid result. Such divergence is dealt with by introduction of a factorization scale (μ_F) which truncates the integral before it reaches the nonperturbative region. Thus, particles with momentum scale above μ_F are treated perturbatively and below which are treated as a part of modified PDF.

In practice, an appropriate choice of these scales should minimize the dependence. For the $t\bar{t}$ cross section, the μ_R and μ_F scales are set to be the same scale at which α_s is evaluated. Fortunately, the energies of the $t\bar{t}$ are well above the cutoff for valid perturbation (Λ_{QCD}); α_s is set to m_{t} and varied from $m_{\mathrm{t}}/2$ to $2m_{\mathrm{t}}$ to determine the uncertainty associated with this choice of scale.

We consider the process

$$N_1(P_1) + N_2(P_2) \rightarrow t(p_3) + \bar{t}(p_4) + X(p_X) \tag{1.45}$$

where X is an inclusive hadronic final state. At Born level this proceeds only through the $q\bar{q}$ annihilation and gg fusion channels as shown in Figure 1.5

$$q(p_1) + \bar{q}(p_2) \rightarrow t(p_3) + \bar{t}(p_4),$$

$$g(p_1) + g(p_2) \rightarrow t(p_3) + \bar{t}(p_4),$$

where $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$. We define the kinematic invariants

$$S = (P_1 + P_2)^2 \ \ , s = (p_1 + p_2)^2$$

S and s denote the hadronic center-of-mass energy squared and the partonic center-of-mass energy squared. The total hadronic cross section for $t\bar{t}$ production is obtained from the convolution of the factorized partonic cross section $\hat{\sigma}_{ij}$ with the parton

luminosities $\mathcal{L}_{ij}$ [63] [26]:

$$\sigma_{p\bar{p}\to t\bar{t}X}(S, m_t) = \sum_{i,j=q,\bar{q},g} \int_{4m_t^2}^{S} ds\, \mathcal{L}_{ij}(s, S, \mu_F)\, \hat{\sigma}_{ij}(s, m_t, \alpha_s(\mu_R), \mu_F), \quad (1.46)$$

$$\mathcal{L}_{ij}(s, S, \mu_F) = \frac{1}{S} \int_s^S \frac{d\hat{s}}{\hat{s}}\, f_{i/p}(\frac{\hat{s}}{S}, \mu_F)\, f_{j/\bar{p}}(\frac{s}{\hat{s}}, \mu_F) \quad (1.47)$$

Here $f_{i/p}(x, \mu_F)$ is the PDF describing the density of partons of flavor i in the proton carrying a fraction x of the initial proton momentum, at factorrization scale μ_F. The QCD coupling constant $\alpha_s(\mu_R)$ is evaluated at the renormalization scale μ_R. The top quark mass $m_{\rm t}$ appearing in Equation (1.46) is the mass renormalized in the on-shell (pole-mass) scheme while the sum runs over all the light flavors *u,d,c,s,b* which considered massless.

In perturbative QCD the partonic cross section $\hat{\sigma}_{ij}(s, m_t, \alpha_s(\mu_R), \mu_F)$ can be expanded in the QCD coupling constant up to any order. Up to NNLO with $m_t = \mu_R = \mu_F$ we read :

$$\hat{\sigma}_{ij} = a_s^2 \hat{\sigma}_{ij}^{(0)}(s, m_t) + a_s^3 \hat{\sigma}_{ij}^{(1)}(s, m_t, \mu_R, \mu_F) + a_s^4 \hat{\sigma}_{ij}^{(2)}(s, m_t, \mu_R, \mu_F) + \mathcal{O}(a_s^5), \quad (1.48)$$

with $a_s = \alpha_s/\pi$. In LO only the parton channels $q\bar{q}$ and gg contribute and the respective Born cross sections are given by:

$$\hat{\sigma}_{q\bar{q}}^{(0)} = \frac{4\pi^3}{27} \frac{1}{s} \beta(3 - \beta^2), \quad (1.49)$$

$$\hat{\sigma}_{g\bar{q}}^{(0)} = \hat{\sigma}_{gq}^{(0)} = 0, \quad (1.50)$$

$$\hat{\sigma}_{gg}^{(0)} = \frac{\pi^3}{48} \frac{1}{s} \left[(33 - 18\beta^2 + \beta^4) \ln\left(\frac{1+\beta}{1-\beta}\right) - 59\beta + 31\beta^3 \right], \quad (1.51)$$

[26] The discussion in this section is largely drawn from the description of the program HAdronic Top and Heavy quarks crOss section calculatoR -HATHOR- of this reference. Therefore, the small s denotes the hadronic center-of-mass whenever it is mentioned in the thesis except for the formalism presented here.

Cross Section [pb]	Tevatron	LHC
σ_{LO}	5.820	583.7
σ_{NLO}	7.229	877.4
σ_{NNLO} (approx)	7.814	923.0

Table 1.2: The LO, NLO and approximate NNLO prediction for the total cross section of $t\bar{t}$ events at LHC ($\sqrt{s} = 14$ TeV) and Tevatron, from [64].

Reference	CTEQ	MRST
[64]	$7.34^{+0.24\ +0.41}_{-0.38\ -0.41}$	$7.04^{+0.24\ +0.14}_{-0.36\ -0.14}$
[65]	$6.79^{+0.35\ +0.43}_{-0.74\ -0.34}$	$7.13^{+0.45\ +0.21}_{-0.84\ -0.17}$

Table 1.3: The best predictions for $t\bar{t}$ [pb] production at the Tevatron for a pole mass of $\mathrm{m}_t = 173$ GeV.

with $\beta = \sqrt{1-\rho}$ and $\rho = 4m_t^2/s$. At higher orders in α_s the partonic cross section receives corrections from virtual loop diagrams and real gluon emissions in the $q\bar{q}$ and gg channels, as well as from the other partonic channels $g\bar{q}$ and gq.

1.5.2 The Predicted Inclusive Cross Section of $t\bar{t}$

Table 1.2 compares the LO, NLO and approximate NNLO prediction for the total cross section at LHC ($\sqrt{s} = 14$ TeV) and Tevatron using $\mathrm{m}_t = 171$ GeV, the PDF set CTEQ6.6 and $\mu_R = \mu_F = m_t$.

Recently, different calculations using different PDFs have been reported for the $t\bar{t}$ cross section. Table 1.3 summarizes the results for some of them at the Tevatron using $\mathrm{m}_t = 173$ GeV. The results are presented in the form

$$\sigma = \sigma(\text{central})\,{}^{+\Delta\sigma_{\mu+}\ +\Delta\sigma_{\mathrm{PDF+}}}_{-\Delta\sigma_{\mu-}\ -\Delta\sigma_{\mathrm{PDF\text{-}}}},$$

where σ(central) is the best prediction, and $\Delta\sigma_{\mu\pm}$ and $\Delta\sigma_{\mathrm{PDF}\pm}$ quantify the uncertainties due to the scale and PDF choices respectively. The explicit dependence of the cross section on the top quark mass and the hadronic center-of-mass energy as shown in Equation (1.46) is depicted in Figures 1.16 and 1.17 respectively. We can

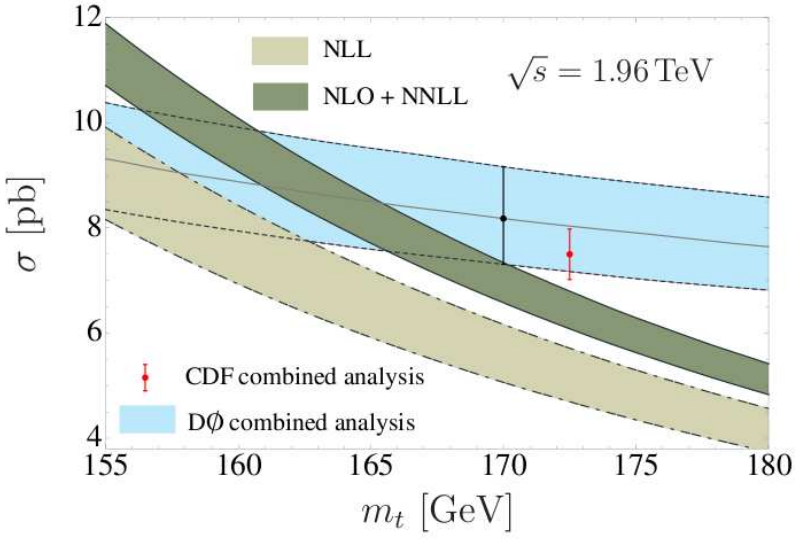


Figure 1.16: Dependence of the total cross section on the top quark mass defined in the pole scheme. The NLL and NLO+NNLL bands indicate the scale uncertainties and PDF uncertainties combined linearly, while the blue band represents the dependence of the DØ measurement of the cross section on m_t. For comparison, a CDF combined measurement is also shown, from [62].

see that the small changes in the center-of-mass energy (by upgrading the Tevatron from RunI to RunII) results in large changes in the cross section. On the other hand, Figure 1.16 shows that the cross section changes by ± 1 pb for each ± 5 GeV change in the top quark mass (a heavier top is harder to produce) [27]. This is actually a manifestation of the fact that $t\bar{t}$ are produced predominantly near the kinematic threshold at the Tevatron. This situation is different at the LHC, where the center-of-mass energy is far above the threshold.

1.5.3 The Measured Inclusive Cross Section of $t\bar{t}$

It is particularly important to measure the $t\bar{t}$ inclusive production in the different top quark SM decay channels since new physics contributions can affect the various $t\bar{t}$ final states, differently. Within uncertainties, all the measured $t\bar{t}$ cross sections in

[27] The next-to-leading-logarithm (NLL) contributions presented in Figure 1.16 have been shown to be small, but they improve the scale stability.

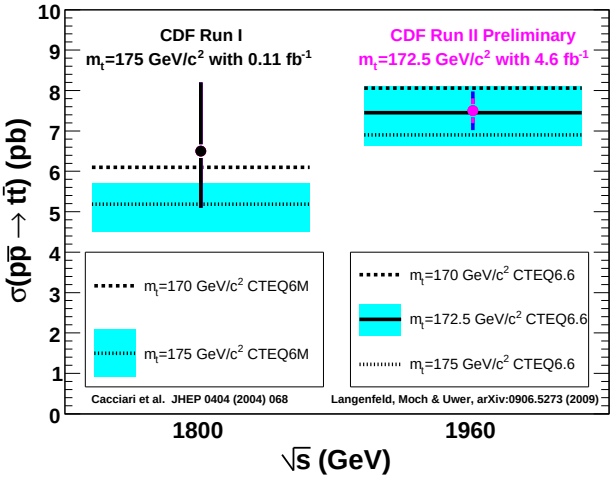
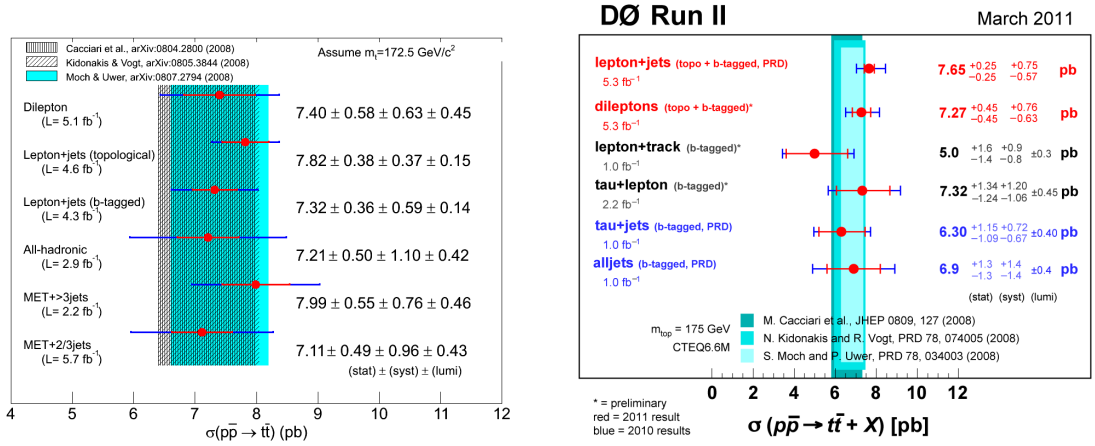


Figure 1.17: Combination of CDF measurements of the $t\bar{t}$ production versus the theoretical predictions as a function of the center-of-mass energy. For RunII, the central values of the predictions from [64] are shown at top quark mass of 170, 172.5 and 175 GeV for CTEQ6.6. The uncertainty of the 172.5 GeV prediction is shown by the filled blue region. For RunI, the predictions are shown at a top quark mass of 170 and 175 GeV fro CTEQ6M, from [66].

almost all possible final states conducted by CDF and D∅ agree with each other. A combination of all these results yields an improved precision on the $t\bar{t}$ cross section.

Combining three measurements using about 1 fb^{-1}, D∅ finds $\sigma_{t\bar{t}} = 8.18^{+0.98}_{-0.87}$ pb for $m_{\rm t}$=170 GeV [67]. CDF performs a combination of four measurements, using up to 4.6 $^{-1}$ of data. For $m_{\rm t}$=172.5 GeV [68], the combination yields $\sigma_{t\bar{t}} = 7.50 \pm$ 0.48 pb. These measurements are in good agreement with the predictions and have uncertainties comparable to those associated with the theoretical expectations. Figure 1.18 shows a summary of the latest cross section measurements performed at both experiments separately. Unfortunately, no combined cross section measurement from both experiments exists to date, unlike for the single top cross section and the top quark mass as shown previously.

In addition to the total inclusive cross section, kinematic distributions are also of interest. The differential cross section for the $t\bar{t}$ invariant mass ($M_{t\bar{t}}$) can be used as



(a) CDF (b) DØ

Figure 1.18: Summary of the latest $t\bar{t}$ total cross section measurements performed at the Tevatron, from [69, 70].

a complementary method for measuring m_{t}. The theoretical interpretation can be drawn from Equation (1.46): all of phase space is integrated over but for a region of the invariant mass ($dM_{t\bar{t}}$). The presence of bumps in the smoothly decreasing $M_{t\bar{t}}$ distribution would be a clear signal of a heavy resonance, which is predicted in many new physics scenarios [28]. If there was a massive particle x that strongly coupled to the top quark it would be seen as a resonance (distortion) in $d\sigma/dM_{t\bar{t}}$ in the region of m_x. It is possible to imagine a scenario where such a process would not change the inclusive cross section to an appreciable degree. In this way the $M_{t\bar{t}}$ distribution provides the window for the new physics models.

Such searches have been pursued at the Tevatron, and results for the $M_{t\bar{t}}$ distributions were recently obtained from 2.7 fb^{-1} of data collected by CDF [71]. The measured $d\sigma/dM_{t\bar{t}}$ is shown in Figure 1.19 which concludes that there is no evidence for non-SM physics in the $M_{t\bar{t}}$ distribution. Limits on KK gravitons of mass 600 GeV which decay to $t\bar{t}$ are set. DØ has measured the differential cross section as a function of the transverse momentum (P_{T}) of the top quark, $d\sigma/dP_{\mathrm{T}}^{\mathrm{t}}$ using ≈ 1 fb^{-1} of data.

[28] Such as the extended gauge theories with massive Z-like bosons, theories with top color, or with axigluons or more recently signatures for Kaluza-Klein (KK) states of gravitons.

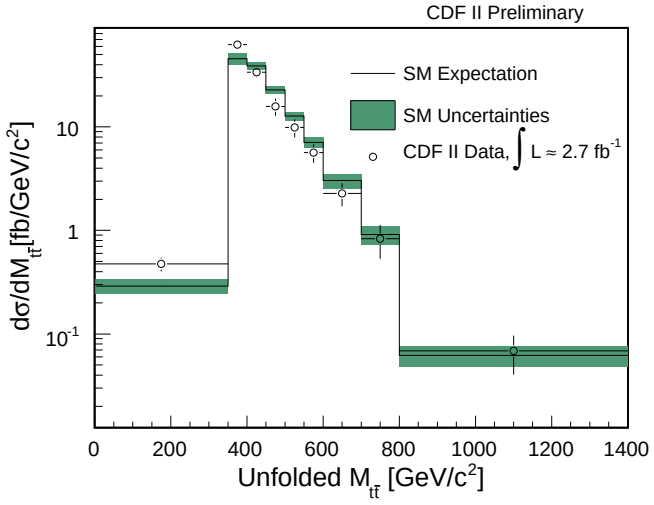


Figure 1.19: The measured differential cross section for the $t\bar{t}$ invariant mass with 2.7 fb^{-1} of data at CDF, from [72].

Results [73] from higher order QCD corrections agree with the measured results as shown in Figure 1.20 .

The ATLAS and CMS experiments at the LHC have both recently reported their first measurement of the total $t\bar{t}$ cross section at center-of-mass energy of $\sqrt{s} = 7$ TeV in a data sample of 2.9 and 3.1±0.3 pb^{-1} respectively. The cross section is measured to be [75, 76] $\sigma_{t\bar{t}} = 145 \pm 31(\text{stat}) \, ^{+42}_{-27}(\text{syst})$ and $\sigma_{t\bar{t}} = 194 \pm 72(\text{stat}) \pm 24(\text{syst}) \pm 21(\text{lumi})$ for ATLAS and CMS respectively. These are in agreement with each other, as well as with NLO QCD prediction ($\sigma_{t\bar{t}}$ increases at $\sqrt{s} = 14$ TeV $\approx$ 100-fold relative to the Tevatron as displayed in Table 1.2). Figure 1.21 shows the ATLAS and CMS measurements together with other relevant Tevatron measurements.

1.5.4 The Predicted Exclusive Cross Section of $t\bar{t}$+jet

As mentioned in Section 1.5.1, the scale dependence of the NLO calculations render the theoretical results quite uncertain which calls for the NNLO calculations of the top quark production. In general, there are four classes of contributions that need to be calculated for the NNLO top quark production [78]. The first class involves the

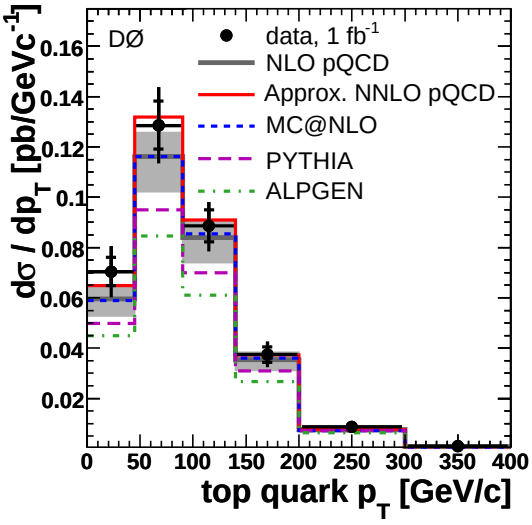


Figure 1.20: The measured differential cross section for the $t\bar{t}$ transverse momentum with 1 fb^{-1} of data at DØ, from [74].

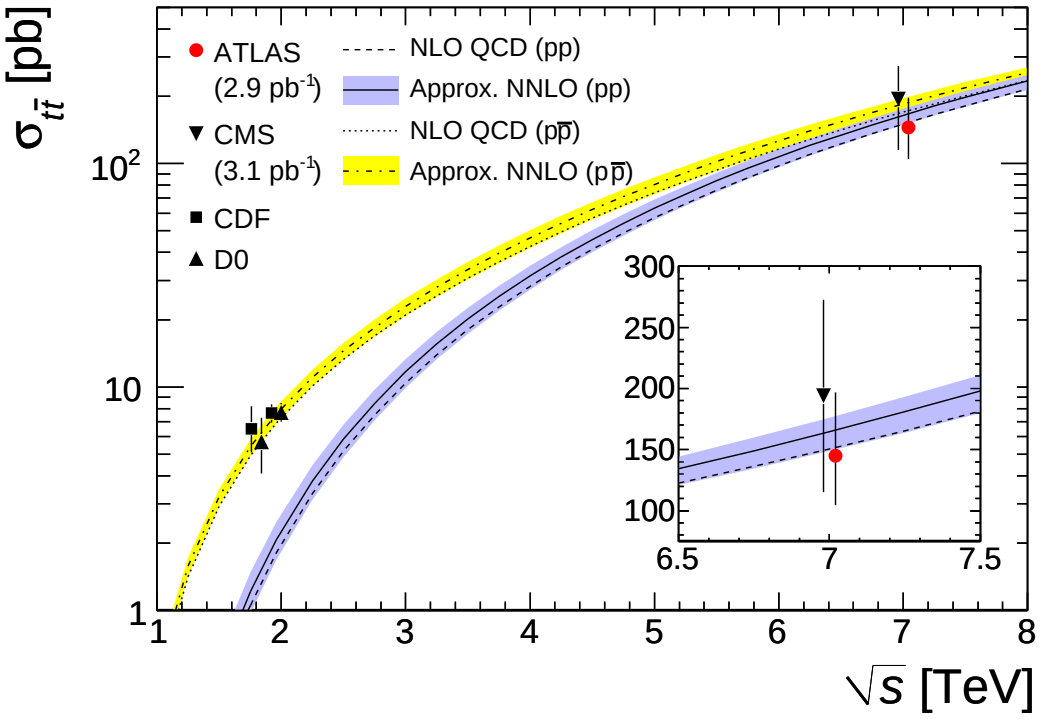


Figure 1.21: The $t\bar{t}$ total cross section as measured by ATLAS [75] and CMS [76] at LHC, CDF and DØ at Tevatron [77]. The theoretical predictions for $p\bar{p}$ and pp collisions [64] obtained using the HATHOR tool [63] and assume a top quatk mass of 172.5 GeV, from [75].

two-loop contributions, which has to be folded with the LO Born term. Second, there are the squared one-loop diagram contributions. Further, the third class of diagrams consists of the one-loop gluon emission contributions that are folded with the one-gluon emission graphs. The later is the topic of present section. Finally, there are the squared two-gluon emission contributions that are of tree-type. Bits and pieces of the NNLO calculation are now being assembled. Some recent studies can be found in [62].

An important prediction of the NLO calculation is the possibility of additional hard partons produced in association with the $t\bar{t}$ pair. If one is produced with sufficiently high P_{T} and large angular separation from other objects in the event, it is detected as a jet.

In the work of [79, 80] the full, exact NLO calculations to $t\bar{t}$+jet were presented for the first time. When integrating over the full phase space of the jet, this calculation can be turned into a NNLO calculation of the third class as aforementioned above. The predicted cross section calculation for 14 TeV is also important as a preview of the LHC, for which [79, 80] has found a cross section of 376 pb for $t\bar{t}$+jet with a P_{T} above 50 GeV (see Table 1.4). This is almost half of the total inclusive $t\bar{t}$ cross section evaluated at NLO in [64] (Table 1.2). This additional jet activity could signal new physics such as top-quark compositeness. It also provides a sensitive tool to search for anomalous top-gluon couplings. As mentioned in section 1.4.3 on page 40, the charge asymmetry for inclusive $t\bar{t}$ appears first at one-loop level, and thus a NLO calculation for the inclusive production provides only a LO calculation of the asymmetry. On the other hand, an asymmetry is present at LO for $t\bar{t}$+jet production, thus the NLO calculation for this cross section is truly a NLO calculation for the $t\bar{t}$+jet asymmetry. Apart from its significance as a signal process, it turns out that $t\bar{t}$+jet production is also an important background to various new physics searches, where the number of jets in the process is often taken as a discriminator against the background [81].

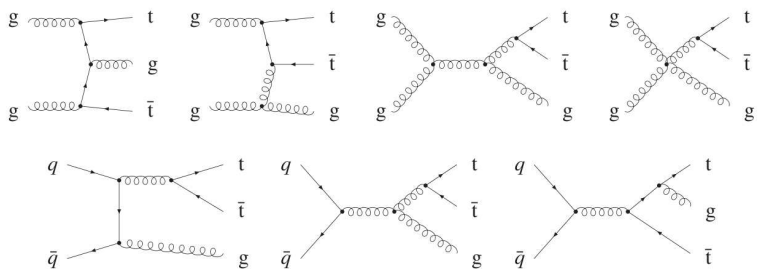


Figure 1.22: Representative sets of LO diagrams for gg fusion and $q\bar{q}$ annihilation in $t\bar{t}$+jet production, from [79].

A prominent example is Higgs production via vector boson fusion. This reaction represents an important discovery channel for a SM Higgs boson with a mass of up to several hundred GeV. The major background to this reaction is due to $t\bar{t}$+jet, again underlying the need for precise theoretical predictions for this process. I will briefly describe the calculation and present the results.

At LO, $t\bar{t}$+jet production receives contributions from the partonic processes $q\bar{q} \to t\bar{t}g$, $qg \to t\bar{t}q$, $\bar{q}g \to t\bar{t}\bar{q}$, and $gg \to t\bar{t}g$. The first three channels are related by crossing symmetry. Therefore, the required generic matrix elements are $0 \to t\bar{t}q\bar{q}g$ (first three), $0 \to t\bar{t}ggg$, where for convenience all particles are listed as outgoing. Representative sets of Born diagrams for the gg and $q\bar{q}$ channels are depicted in Figure 1.22 . In total, there are 16 LO diagrams for $0 \to t\bar{t}ggg$ and 5 for $0 \to t\bar{t}q\bar{q}g$. Evaluating 2→3 particle processes at NLO level, is not trivial, both in the analytical and numerical parts of the calculation. There is significant complexity due to the presence of one-loop pentagon diagrams and the involved tensor integrals which render computer codes quite lengthy and CPU time consuming. Prototypes of the pentagon graphs are shown in Figure 1.23 . The total number of diagrams (self-energy, vertex, box-type and pentagon-type) is about ∼400. The challenging step in this context is the numerically fast and stable reduction of the tensor integrals to box integrals which in turn are reduced to scalar integrals.

Figure 1.24 shows the dependence of the LO and NLO cross section for $t\bar{t}$+jet

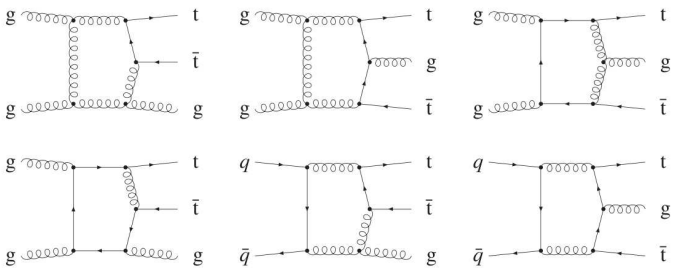
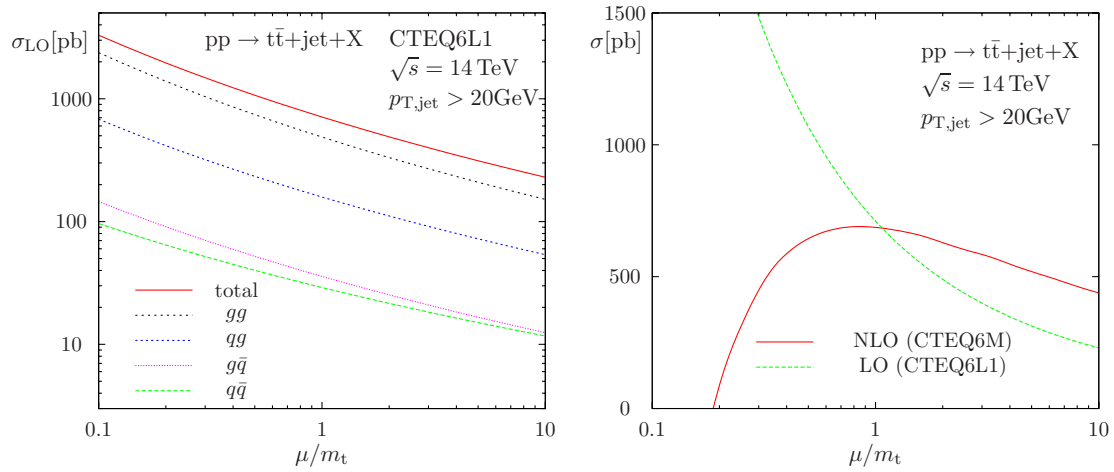
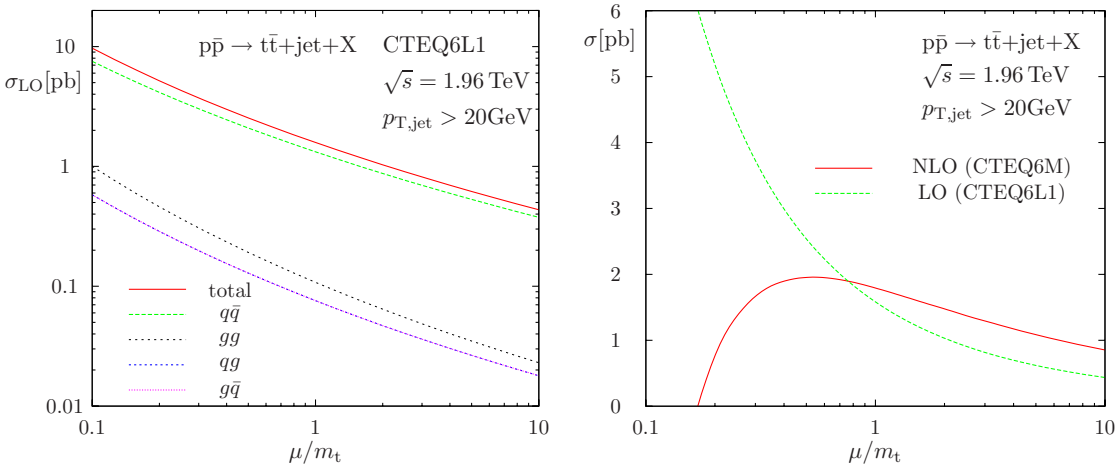


Figure 1.23: Representative sets of pentagon diagrams for gg fusion and $q\bar{q}$ annihilation in $t\bar{t}$+jet production at NLO QCD, from [79].



(a) LO Scale dependence, LHC

(b) NLO Scale dependence, LHC



(c) LO Scale dependence, Tevatron

(d) NLO Scale dependence, Tevatron

Figure 1.24: Scale dependence for the cross section for $t\bar{t}$+jet production, where μ_{R} and μ_{F} scales are set equal to μ. The individual contributions of the various partonic channels are shown also separately, from [80].

on the μ_R and μ_F scales, which are set equal to μ. The dependence for the LO case is rather large, illustrating the well-known fact that the LO predictions can only provide a rough estimate. At the Tevatron (Figure 1.24 (c)) the $q\bar{q}$ channel dominates by about 85%, followed by the gg channel with about 7%. This is rather similar to the inclusive $t\bar{t}$ pair cross section as discussed in Section 1.4.1. In contrast to the inclusive case, the gq and $g\bar{q}$ subprocesses are not suppressed. At the LHC (Figure 1.24 (a)), the gg channel comprises about 70%, followed by gq with about 22%. The scale dependence of the NLO cross section is shown in the right plots. For comparison, the LO results are included as well. As expected, the NLO corrections significantly reduce the scale dependence.

Table 1.4 provides the LO and NLO predictions for the cross section with P_T cuts of 20 and 50 GeV on the hard jet. The values presented are for the central scale $\mu = \mu_R = \mu_F = m_t$. The uncertainty due to the numerical integration is quoted in parentheses. The scale dependence is indicated by the upper and lower indices. The upper (lower) index represents the change when the scale is shifted towards $\mu = m_t/2\,(\mu = 2m_t)$. Rescaling the common scale $\mu = \mu_R = \mu_F$ from the default value m_t up (down) by a factor 2 changes the cross section at Tevatron in LO and NLO by about 60% (35%) and 9% (18%), respectively. This means that the scale uncertainty is reduced considerably through the inclusion of the NLO corrections. Comparing the LO and NLO it is found that the large scale dependence at LHC of about 100% in the LO cross section is considerably reduced after including the NLO corrections. The ratio of the NLO $t\bar{t}$+jet cross section at LHC to the total inclusive $t\bar{t}$ cross section is about 47%, 22% and 7% for a P_T cut of 50 GeV, 100 GeV and 200 GeV, respectively. At Tevatron, the ratio for the small P_T cut of 20 GeV is almost 30%. This fraction is reduced to about 8% when 50 GeV is chosen for the cut.

Having presented the results of [79, 80], and for completeness, it is worthy to note that $t\bar{t}$ pairs were treated as stable particles and their decays, even at LO, were

$\sqrt{s}$	P_{T} [GeV]	LO [pb]	NLO [pb]
Tevatron	20	$1.583(2)^{+0.96}_{-0.55}$	$1.791(1)^{+0.16}_{-0.31}$
	50	$0.4670(6)^{+0.29}_{-0.17}$	$0.5244(4)^{+0.049}_{-0.096}$
LHC	20	$710.8(8)^{+358}_{-221}$	$692.3(3)^{-40}_{-62}$
	50	$326.6(4)^{+168}_{-103}$	$376.2(6)^{+17}_{-48}$

Table 1.4: Cross section for $t\bar{t}$+jet [pb] for different values of P_{T} jet cut for $\mu = \mathrm{m}_t$. As a numerical value, m_t=174 GeV. The upper and lower indices are the shifts towards $\mu = \mathrm{m}_t/2$ and $\mu = 2\mathrm{m}_t$, from [79].

not included. This is not entirely realistic since all cuts designed apply to $t\bar{t}$ decay products. Therefore, it is important to account for decays of $t\bar{t}$ in the calculations of NLO corrections, to have full confidence in the results.

How the results in [79, 80] should be compared to our measured cross section is what we will discuss in section 4.2.3 .

1.6 Measurement of $\sigma(p\bar{p} \rightarrow t\bar{t}$+jet)

The Tevatron experiment is a scattering experiment; $(p\bar{p})$ beams collide and then scatter elastically to $(p\bar{p})$ and inelastically producing new particles as $(t\bar{t}$+jet). The total inclusive cross section is the sum of all the possible elastic and inelastic scattering processes. On the other hand, the exclusive cross section for $p\bar{p} \rightarrow t\bar{t}$+jet can be thought of as

$$\sigma = \frac{\text{transition rate}}{\text{incident flux}}$$

$$\sigma = \frac{\#\text{ of times } (\mathrm{t\bar{t}} + \text{jet}) \text{ produces per unit time}}{\#\text{ of } (\mathrm{p\bar{p}}) \text{ beams per unit area per unit time}}$$

leading σ to have unit of area as expected, to give a number which is independent of the beam flux. Because the length scale accessible to a given probe particle is inversely proportional to its energy, the customary unit of σ is barn. 1 barn (b) is $10^{-28}\mathrm{m}^2$. The barn was originally intended to describe the nuclear physics processes,

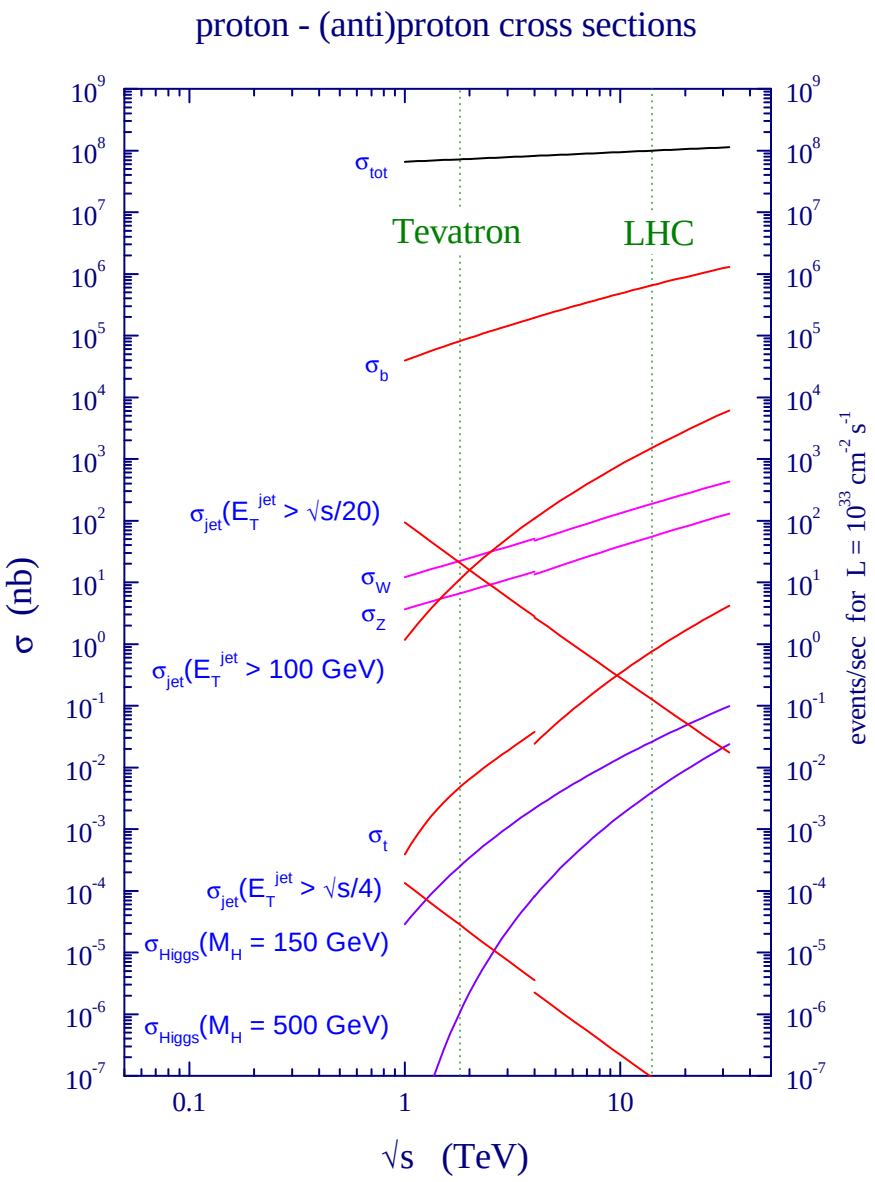


Figure 1.25: SM corss section at the Tevatron and LHC colliders, from [81].

and hence for HEP, the scale for the inclusive σ is of the order of microbarn, for electroweak and top processes is picobarn (pb) and it is of the order of femptobarn (fb) for Higgs boson production. Figure 1.25 shows the predictions for some exclusive processes at $p\bar{p}$ and pp colliders. So, for example, production of a Higgs boson with

mass of 150 GeV is about an order of magnitude less likely than $t\bar{t}$ production, while W boson production is three orders of magnitude more likely than the $t\bar{t}$ production.

The incident flux, expressed as the instantaneous luminosity ($\mathcal{L}$), measures the amount of data collected. The cross section for $t\bar{t}$+jet then can read:

$$\sigma = \frac{N_{t\bar{t}+\text{jet}}}{\int \mathcal{L} dt} \tag{1.52}$$

Note that the integrated luminosity $\int \mathcal{L} dt$ has the unit of inverse barn, so that it can readily be converted to σ given the number of the events, and vice versa. In our measurement, we use 4.1 fb^{-1} of integrated luminosity. On the other hand, counting the number of $t\bar{t}$+jet events is not simple due to backgrounds. Backgrounds are other processes which mimic the signature of the $t\bar{t}$+jet events in the detector and hence pass the event selection. Therefore, the backgrounds events (N_{bkg}) are estimated using MC simulation and data and subtracted from the $t\bar{t}$+jet observed events. However, in practice, detectors do not offer ideal coverage and efficiency and hence this should be taken into account. Having the factor ϵ to encompass for the acceptance selections, lepton identification, b-jet tagging and BR into lepton plus jet channel, Equation (1.52) is modified to:

$$\sigma = \frac{N_{t\bar{t}+\text{jet}} - N_{bkg}}{\epsilon \int \mathcal{L} dt} \tag{1.53}$$

The two subsequent chapters are presented to pave the way for the measurement and to describe all the aspects revealed from this equation. Chapter 4 will conclude with the ultimate results.

Chapter 2

Experimental Apparatus

Figure 2.1: An aerial view of Fermilab. The top ring in the photograph is the Tevatron. Protons circle Tevatron clockwise as anti-protons circle counter-clockwise. The beams are steered to collide at the top-left of the ring (CDF) and at the top-right (D∅). The small ring at the bottom is the main injector, from [82]

2.1 Introduction

The Tevatron is the last stage of an entire accelerator chain at Fermilab. Collisions occur at two points at the Tevatron (the ring), where the two multipurpose detectors, CDF and D∅, are placed. In this chapter, I will discuss the accelerator complex and CDF, focusing on the aspects of CDF most used in the analysis.

The Tevatron began operations in 1985 with a center-of-mass energy of $\sqrt{s} = 1.8$ TeV. RunI occurred from 1992-1996, during which the top quark was discovered. The current physics run, RunII, began operation in 2001, after substantial upgrades to the accelerator and both detectors, and is smoothly running with $\sqrt{s} = 1.96$ TeV. However, the U.S. Department of Energy has decided to shut down the Tevatron by the end of September of this year for financial deficit. What a great loss!

2.2 The Accelerator Complex

One would wonder, as electrons are lighter and so much easier to accelerate, why not build electron synchrotrons? This can be done, but there is a significant cost because all charged particles radiate energy as they accelerate. Circularly moving particles at higher speeds accelerate more and more, losing energy (synchrotron radiation) that varies inversely with the fourth power of their mass. This makes protons radiate much less.

On the other hand, designing a $p\bar{p}$ collider is more easy than pp collider because for the former case, the particles can travel opposite directions in the same beam pipe and be bent by the same set of magnets without being disturbed by each other. However, a $p\bar{p}$ machine poses unique problems in that obtaining a beam of anti-protons is much harder than obtaining a positron beam; anti-protons must be stored once they are produced. In addition, a single accelerator can not bring particles from rest to TeV scale of energy because no magnets have the dynamic range necessary. Consideration

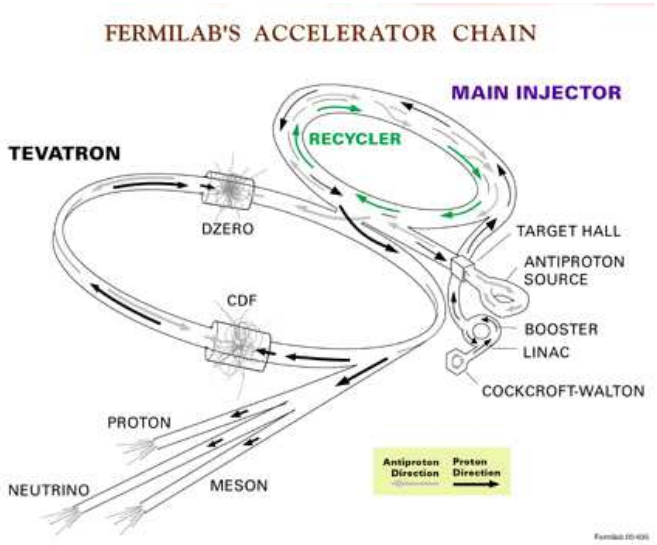


Figure 2.2: A schematic figure of Fermilab accelerator chain, from [82].

of these requirements and others led to the design of a chain of accelerators at Fermilab as shown in Figure 2.2. I will consider each element in turn. For a complete updated review, consult Reference [83].

2.2.1 Proton Production

Pure Hydrogen gas (H) is moved between electrodes and a spark ionizes the atoms into electrons and H^+ ions. The positive ions strike a Cesium cathode which results in the formation of H^- ions and electrons. These H^- ions are sent to the Cockcroft-Walton pre-accelerator (preacc) which will subject the negative ions to a -750 kV potential, thus producing H^- beams with energy of +750 kV.

The linear accelerator (linac) accelerates the H^- ions through drift tubes to 400 MeV. The tubes use an alternating-current electric field (usually referred to as RF, for Radio Frequency). RF cavities are designed so that the H^- are shielded from the field when it points opposite their direction of motion. Thus, H^- only feel a force from the field when it will increase their speed. This method of acceleration causes the H^- to group together into bunches which in turn are sent along a transfer line to

the Booster.

The Booster is the first synchrotron in the accelerator complex, composed of 96 magnets arranged in a circle approximately 150 meters in diameter. With 18 RF cavities, the booster receives the H^- bunches and strips the electrons off, leaving bare proton bunches (H^+) with an energy of 8 GeV. At this point, the bunches go into a transfer line that leads to the main injector, and then on to the Tevatron.

The main injector is a larger proton synchrotron with a diameter of about one km. It uses 344 dipole magnets and 208 focusing quadrupole magnets to both accelerate the 8 GeV protons from the booster to 150 GeV for injection into the Tevatron in two seconds and to steer the bunches as as well. It has another function involving the Tevatron; accelerating the 8 GeV protons to 120 GeV, which are then used to strike the anti-proton source and create anti-proton bunches.

2.2.2 Anti-Proton Production

The 120 GeV protons sent from the main injector strike a nickel target [1], creating a spray of particles. A magnetic spectroscopy procedure is used to select the anti-protons from the bending radii information. The resultant anti-protons have a wide range of energy, averaging 8 GeV. Studies have shown that the 120 GeV is the optimal energy for anti-protons production; at this energy, approximately one anti-proton is produced per 10^5 protons sent to the nickle target.

The Debuncher is a triangular storage ring which accepts the 8 GeV anti-protons bunches. Its main purpose is to "debunch" the anti-proton beam (which have a large energy spread) by removing its RF bunch structure. This results in an anti-proton beam that have no RF bunch structure but have a similar energy. This is referred to as the stochastic cooling process. The cooled anti-protons are then injected into the accumulator at 8 GeV with a uniform energy.

[1] Nickel can absorb more heat without melting, making it a good choice for anti-proton target.

The accumulator lies in the same tunnel as the Debuncher. It is a long-term storage ring, designed to store the anti-protons beams with minimal losses for days. They are manipulated by RF systems to fill a stable region of phase space, known as the core; a smaller core results in a higher luminosity. They are accumulated and cooled here until the storage reaches its maximum capacity. At that point they are passed into the recycler.

The recycler is located in the same ring as the main injector. It is a ring of permanent magnets designed to store anti-protons at 8 GeV. However, the extracted anti-protons from the accumulator should be collected into bunches again. RF activation causes them to be grouped into bunches, which are then transferred back to the main injector, decelerated to 8 GeV, and injected into the recycler.

When the Tevatron is ready for a store [2], anti-protons are injected from the recycler to the main injector, which accelerates them to 150 GeV and passed them to Tevatron.

2.2.3 Tevatron

The Tevatron, which we colloquially use it to represent the entire chain, is actually the main large synchrotron at Fermilab. It is 1 km in radius, accelerates $p\bar{p}$ beams from 150 to 980 GeV and steers them using 774 superconducting dipole and 240 quadrupole magnets with a maximum magnetic field of 4.2 T. The magnets are cooled by liquid Helium to 4.2 K, at which the Niobium-Titanium (NbTi) alloy becomes superconducting.

The process of injecting the beams into the Tevatron, accelerating them, and initiating collisions is referred to as a shot. It starts with loading the protons, one bunch at a time. This is repeated 36 times to obtain the 36 bunches needed for the 36×36 store. Afterword, the anti-protons are injected four bunches at a time,

[2] The Tevatron can sustain both beams in collision for hours at a time, called a store.

repeated nine times. Since they are oppositely charged, $p\bar{p}$ beams circle oppositely in helical orbits in the magnetic field, and hence the Tevatron houses both of them.

The Tevatron is not a perfect circle. There are six sectors (A–F), each sector has five service buildings (0–4). The "0" parts have large straight sections. B0 contains CDF and the D∅ detector is aptly named for its place along the Tevatron. When all the $p\bar{p}$ beams are loaded, RF cavities accelerate them to 980 GeV within around 85 seconds. At B0 and D0, they are then focused (squeezed) using the quadrupole magnets into very narrow beam of width of order 32 μm [3]. The beams are then crossed to induce collisions at the center of each detector, at a collision rate of approximately 2.5 MHz (396 ns). This is done by some electrostatic separators which switch polarity to cause the beams to collide. In reality, the collisions don't take place at the center, but are distributed as a Gaussian around the centers of the detectors. This is interpreted by the concept of the cross section.

The store can go on and both detectors can accumulate more integrated luminosity until the anti-protons are used or lost. At that point, the store is ended and the remaining anti-protons, if any, are sent to the recycler for the next shot.

2.3 Coordinates, Conventions & Kinematics

In performing pQCD calculations for top quarks cross section like in Equation (1.46), the partonic center-of-mass frame was not the same as the hadronic center-of-mass frame (the laboratory frame). Since the former frame for the two colliding partons is a prior undetermined with respect to the later frame, the scattering polar angle from the beamline θ is not a good observable to describe the theory and the experiment as well. Therefore, it would be more desirable to describe the kinematics of interactions in terms of variables transverse to the beam axis i.e. invariant under longitudinal boosts.

[3] This reduces the length of the bunch in the z direction from 1.7 m to 30 cm as well.

CDF is cylindrically symmetric about the beamline and extends through the B0 straight section. Protons enter the detector from the west side, and anti-protons enter from the east side. The z axis lies along the beamline, with +z direction defined as the direction of protons traveling from west to east. Longitudinal refers to the component along the z axis, and transverse refers to components perpendicular to the z axis. θ is the polar angle from the beamline and r is the radial dimension transverse to the beamline; $\tan\theta = r/\mathrm{z}$. Commonly, x and y are not used; +x points to north, while +y points up. The azimuthal angle ϕ, is defined clockwise when looking from west to east; $\tan\phi = \mathrm{y/x}$.

Instead of using θ to describe the kinematics, we use the concept of the rapidity, defined as

$$y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z} \tag{2.1}$$

for a particle with energy E and z-momentum p_z. At the TeV scale, the particle mass is generally negligible compared to its momentum, making the rapidity y be approximated by the pseudorapidity η defined as

$$\eta = -\ln\tan\frac{\theta}{2} \tag{2.2}$$

which has one-to-one correspondence with the polar angle $\pi \geq \theta \geq 0$ for $-\infty < \eta < \infty$. η is invariant under boost in the beam direction. Another consequence of this is that instead of E and p, the transverse momentum $p_T = p\sin\theta$ and transverse energy $E_T = E\sin\theta$ are used. Transverse energy is somewhat ambiguous concept because energy is not a vector, but it is convenient to treat only the fraction of energy which is attributable to the motion transverse to the beam. The sum of the transverse energy of all particles in the event should sum to zero in the absence of any particles escaping detection. Therefore, if a particle has escaped the detector without interactions, the x and y components of its transverse energy can be inferred from the transverse energy

apparently missing from the event. In practice, this is the case for neutrinos, so this missing energy, ($\not\!\!E_{\rm T}$), is the distinctive signature of neutrinos in our $t\bar{t}$+jet samples (see section 3.5).

This concludes that we need to use the kinematical variables (η, ϕ) to describe events in CDF and hadronic collisions in general. Therefore, the separation between objects in the detector is invariant under longitudinal boosts and defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{2.3}$$

A very important consequence for the introduction of separation is that it provides a practical definition of a hadronic jet, and ΔR specifies the cone size of a jet formed by the multiple hadrons within it (see section 3.3).

2.4 The CDF Detector

Broadly speaking, the CDF detector contains three main components arranged coaxially around the beamline. The innermost one is the tracking system, which contains a set of silicon strips and a drift chamber inside a superconducting solenoid to measure the momenta of charged particles. Outside of the solenoid are the calorimeters, which measure the energies of electrons, photons and hadrons. Finally, outside of the calorimeters are the muon chambers, which are used to identify muons. The specific components of CDF are shown in Figure 2.3 and are explained in detail in the following subsections [4].

However, Figure 2.4 shows the tracking volume of CDF on which lines of the pseudorapidity of Equation (2.2) are labeled. We can see that $\theta = 90^\circ$ corresponds to $\eta = 0$, and a value of $\eta = 1$ corresponds to $\theta \sim 40^\circ$. As a quantitative illustration, for

[4] The Technical Design Report (TDR) offers a complete (but outdated) overview of the CDF detector [84].

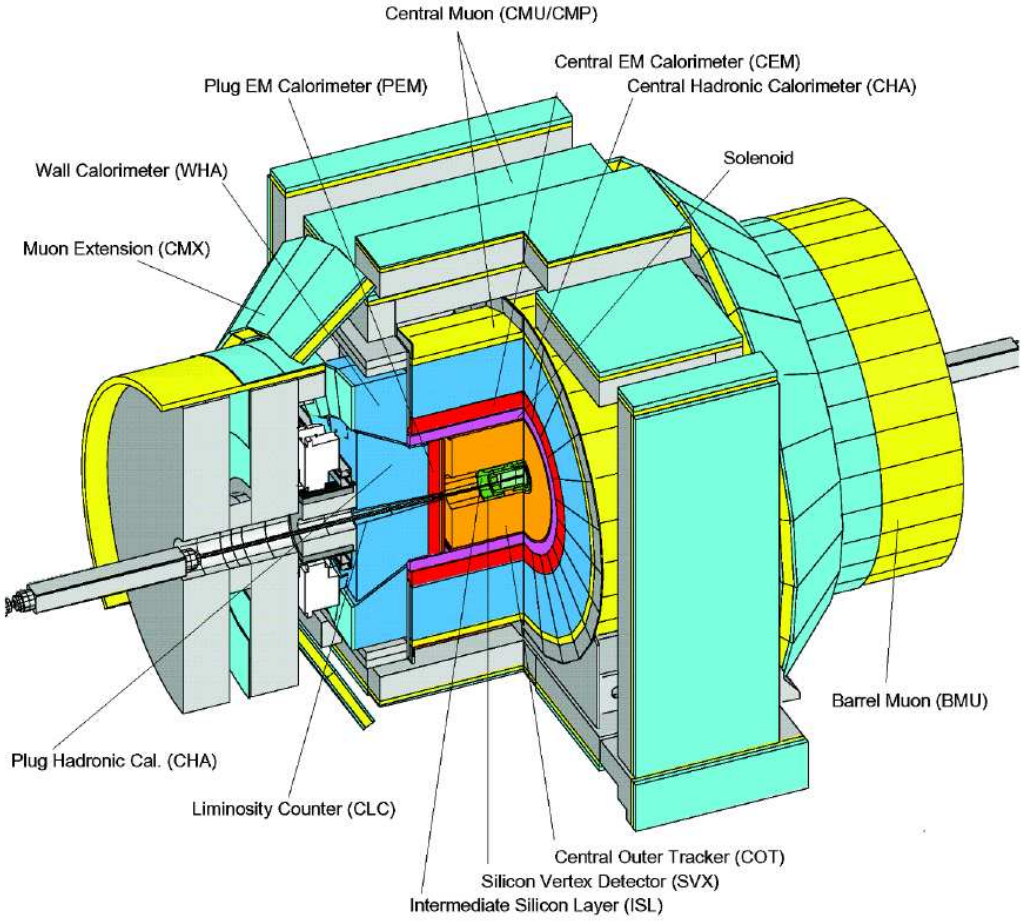


Figure 2.3: An Isometric of CDF with the different subdetectors labeled, from [82].

two objects back-to-back in the central region of the detector, $\Delta\eta < \Delta\phi$ and hence $\Delta R \approx \Delta\phi \sim \pi$. [5]

2.4.1 Tracking System

The tracking system is used to reconstruct the trajectories and momenta of the charged particles. Hence, it is placed inside the solenoid closest to the interaction point. It consists of two major subsystems, the silicon microstrip detectors and the Central Outer Tracker (COT).

[5] The central region of CDF falls in the range of $0 \leq |\eta| \leq 1.0$, while the forward detectors are located in the range $1.0 \leq |\eta| \leq 3.0$.

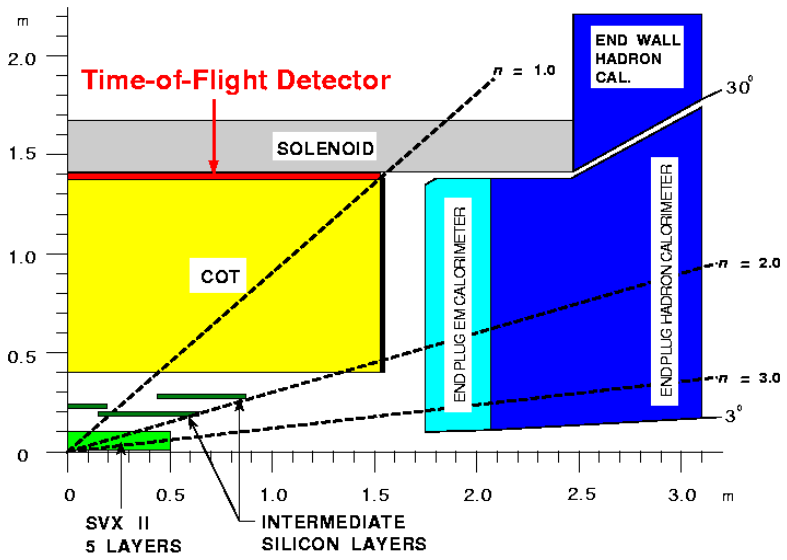


Figure 2.4: A tracking volume in the x-y plane at CDF, from [82].

The best position resolution comes from the silicon detectors which are constructed of thin strips of p-type silicon (100-μm) doped with n-type silicon. When a charged particle strikes the strip, it ionizes the silicon, creating electron-hole pairs. The voltage applied across the silicon draws electrons to one end of the strip, the holes to the other end. The resulting charge is collected by readout chips mounted at the end of the sensors. With double-sided silicon strips, perpendicular to each other, CDF can measure simultaneously the two-dimensional particle's position. Multiple layers of the silicon makes three-dimensional reconstruction of particle tracks possible. However, due to its cost, CDF can not use silicon through out the entire tracking volume depicted in Figure 2.4. In total, the silicon detectors form a cylinder 1.9 m long with a radius of 28 cm and is, in turn, composed of three separate components. An end view of the three components are shown in Figure 2.5.

Layer 00 (L00) [85] is the innermost silicon component and lies only 1.6 cm from the beamline. It consists of single-sided silicon strips with improved resistance to radiation damage and it improves the overall tracking resolution. The main component

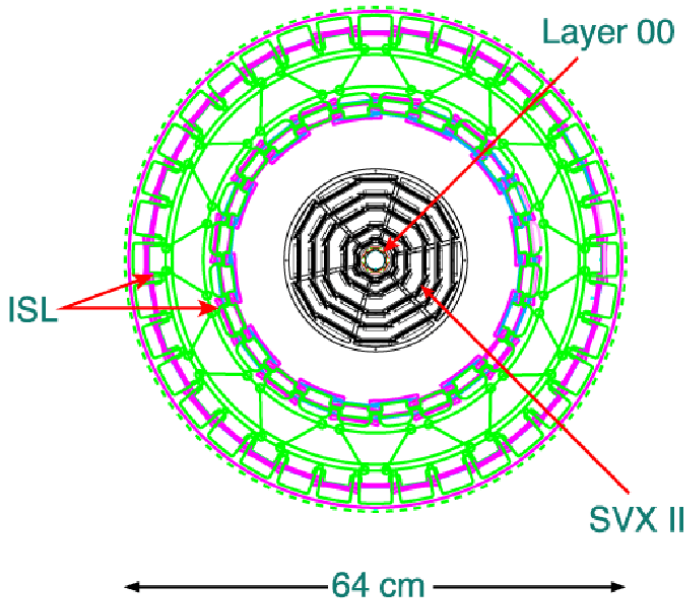


Figure 2.5: A cross-sectional view of silicon detectors in the r-ϕ plane, from [82].

of the Silicon detector is the Silicon Vertex Detector (SVX II) which extends from r = 2.1 cm to r = 17.3 cm around the beamline and covers $\eta < 2.0$. It consists of five layers of doubled-sided silicon strips; each layer consists of twelve ladders arranged radially. Strips on one side are parallel to the z axis (and thus provide $r-\phi$ information and are called axial strips) and those on the other side are 1.2° offset from the z axis (and thus provide r-z information and are called stereo strips). The SVX detector is particularly useful for resolving displaced secondary verticies; the common feature of b-jets (see section 3.4). Lying outside of the SVX II is the Intermediate Silicon Layer (ISL). In the central region ($\eta < 1.0$), there is one layer of silicon at r = 22 cm, and two layers at r = 20 and r = 28 cm for $1.0 < \eta < 2.0$, using double-sided strips. Being further away from the beamline, the ISL is less affected by radiation damage. It serves to improve the linking of tracks between the silicon detector and the COT.

The COT is a 3 m cylindrical drift chamber with a total of 2520 cells organized

into eight radial sections (superlayers), lying outside the silicon detector. It extends to a radius of $r = 137$ cm, covers the range $\eta < 1.0$. The chamber is filled with a nearly equal mixture of Argon and Ethane gases, with a small $\sim 1\%$ component of alcohol. The superlayers are arranged in alternating axial and stereo sections. Each cell contains both sense wires and potential wires. When voltage is applied to the potential wires, a charged particle traveling through the chamber ionizes the gas, and the electrons are attracted to the sense wires. These signals on the sense wires are processed by COT electronics that provide hit time and charge deposition information from each wire [86]. Figure 2.6 shows CDF during the installation of the COT into the center of the tracking system, surrounded by the calorimeter (blue and red), with the plug pulled out (to the left of the photograph).

To determine the time at which the charged particle passes through the COT, a Time Of Flight detector (TOF) made out of scintillating counters is used. TOF sits above the COT just inside the solenoid at $\eta < 1.0$ as can be seen in Figure 2.4 and has a timing resolution of 100 ps. Charged particles passing through the TOF create bursts of light which are collected by photomultiplier tubes (PMTs) attached at each end of the scintillators [87]. The TOF system measures the difference in time between the time of a particle interaction and the arrival of a charged particle. It is used -together with the other information- for particle identification, especially for particles within b-jets.

The whole tracking system is contained within a superconducting solenoidal magnet that provides the magnetic field necessary for measuring charged particle momenta in the central tracking system. It is built from Aluminum-stabilized NbTi conductor and operates with a current of about 4650 A and provides up to 1.5 T magnetic field in the z direction. The solenoid (5 m in length with radius of 1.5 m) is kept inside a cryostat where it is cooled with liquid Helium [88].

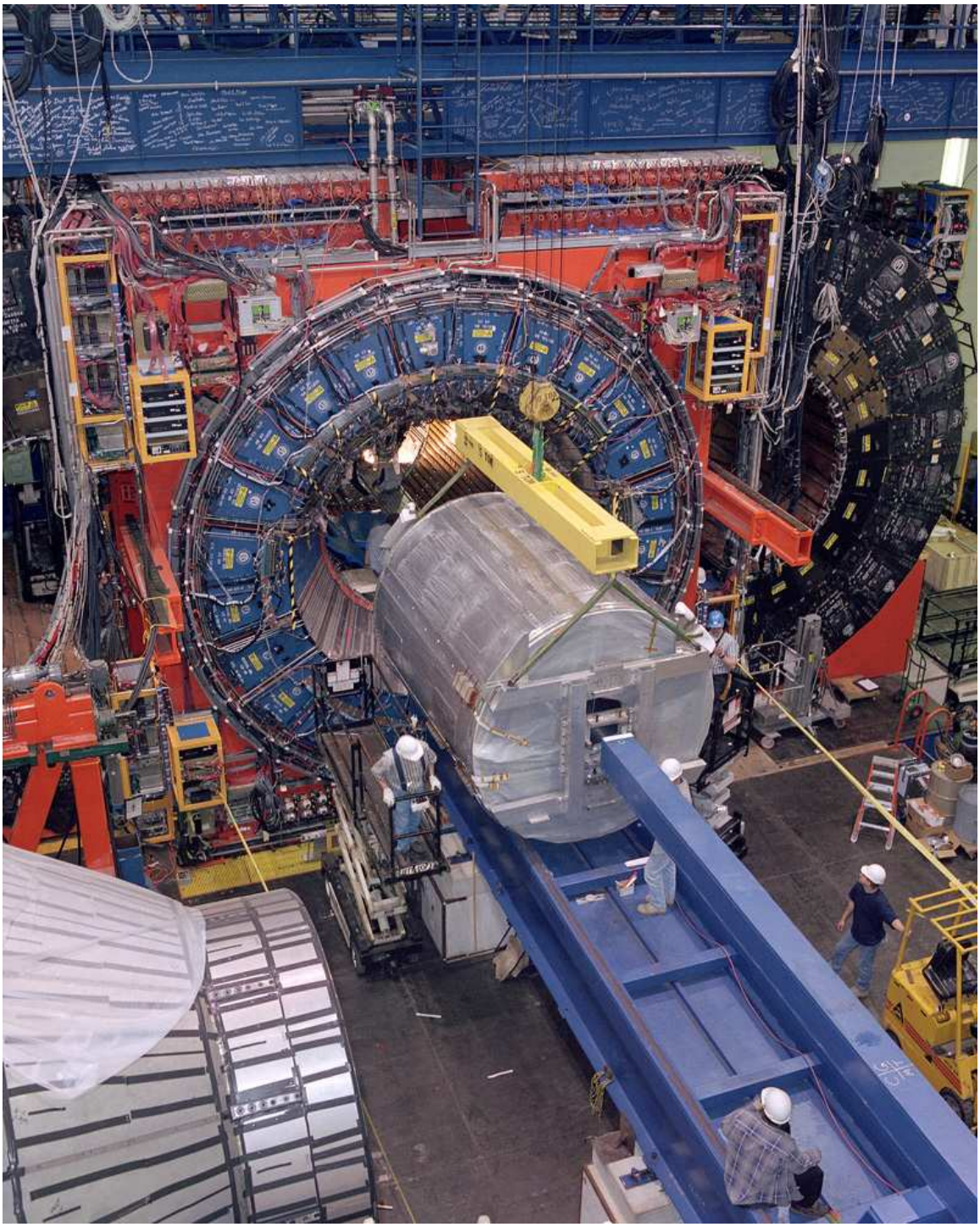


Figure 2.6: Installing the COT, from [82].

2.4.2 Calorimetry

As the tracking system serves to determine the momenta of the charged particles, the calorimeter system determines the energies of the particles.

The calorimeter uses plastic scintillators to collect energy from the charged particles. A higher energy particle passing through the scintillator excites the atoms in it, causing light to be emitted. The light is collected by the PMTs. At the same time, the particle loses its energy and creates showers of less energetic particles, which in turn, scatter off to create more and more showers of less energetic particles. Each charged particle passing a scintillator produces light. The measurement of the full shower profile allows a more precise determination of the particle's energy; thus, most calorimeters have multiple layers. However, CDF reduces the large amount of the scintillator needed by alternating layers of scintillator with layers of a dense absorber. This creates a sampling calorimeter, which measures a fraction of the total energy and estimate the true energy based on calibrations with test beams.

Different particles interact differently with the scintillators. Electrons rapidly lose their energy primarily through the radiation of photons (bremsstrahlung) and form electromagnetic showers. Photons also radiate their energy much faster than the massive particles. Charged hadrons are not deflected by the electric fields produced by the scintillators (losing far less energy to ionization). However, when they pass sufficiently close to the nuclei, their loss energy mediated by the strong force result in showers of less energetic particles. For muons, which are massive version of electrons, the electric field is not large enough to make muons ionizing their energy. On the other hand, they do not interact by the strong force. Hence they generally pass straight through the calorimeters depositing very little of their energy. Figure 2.7 shows what particles may leave what signatures in the sequential layers of CDF.

This motivates the construction of two distinct calorimeters; an electromagnetic calorimeter (EM) and a hadronic calorimeter (HA). Muons on the other hand must

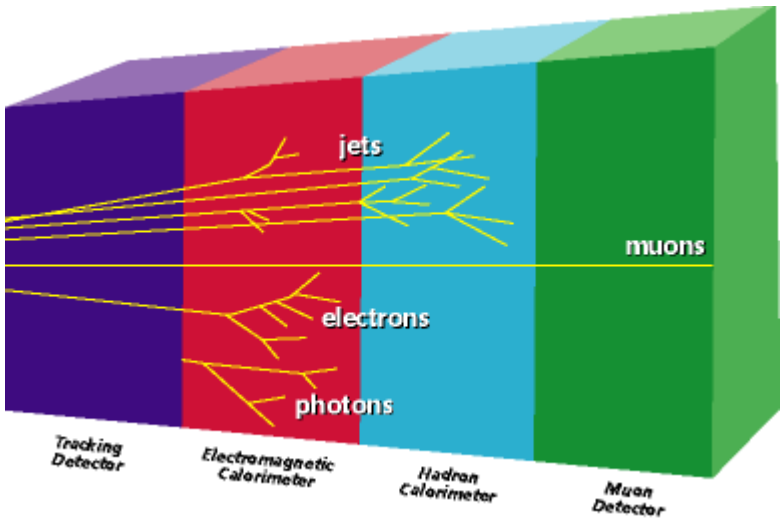


Figure 2.7: The detection of different particles at CDF, from [82].

be measured separately by muon chambers.

Lead and steel are used as the absorbing material for the EM and HA calorimeters respectively. The reason behind using a much denser material for the EM calorimeter is to ensure that the electrons are all absorbed in the EM layer of the calorimetry system. On the other hand, and because of the slower rate of radiation of hadrons, the HA calorimeter must be much larger than the EM calorimeter. The energy resolution of the EM calorimeter is better than that for the HA. The total energy of the initial particle is related to the number of layers it interacts with in the calorimetry system before losing all of its energy, as well as to the amount of light captured by the PMTs. This measured energy must be corrected for different aspects. This will be explained in detail in section 3.3 .

Both the EM and the HA calorimeters are composed of subsystems to fully envelop the rest of the detector. The Central Electromagnetic calorimeter (CEM) [89] and the Central Hadronic calorimeter (CHA) [90] are located in the central region of the detector ($\eta < 1.1$), immediately outside the solenoid, with CEM being closer to the beam. The CEM has 31 layers of lead and scintillator. It is organized into towers in

a projective geometry in η and ϕ. The CHA has 32 alternating layers of steel and scintillator with similar geometry of the CEM.

In the CEM, and at the point where the shower has the greatest number of particles (called shower maximum), a special proportional wire chamber called the Central Electromagnetic Shower Maximum (CES) is embedded in the CEM. This CES with finer position resolution allows for better precision than the surrounding layers. On the other hand, the CHA calorimeter can not be extended as far forward as the CEM one. The gap that is left due to this limitation is filled by the Wall Hadronic calorimeter (WHA), which has 15 layers of steel and scintillator. It extends the hadronic calorimetry coverage in η to the endwall region ($0.8 < \eta < 1.2$).

The plug calorimeters, so named because they fit into the end of the detector like a plug, have analogous components to the central calorimetry. The Plug Electromagnetic calorimeter (PEM) followed by the Plug Hadronic calorimeter (PHA) cover the region from $1.1 < \eta < 3.6$. Both PEM (PHA) have 23 layers of alternating scintillator and lead (steel) respectively [91]. The Plug Electromagnetic Shower Maximum calorimeter (PES) serves the same as the CES.

2.4.3 Muon System

Figure 2.7 shows how muons pass through the calorimeters without leaving much calorimeter information. Consequently, CDF has additional drift chambers and scintillators located at the outer most part of the detector to identify where and when muons passed through. These chambers are mounted behind thick steel slab which serves to decrease the non-muonic particles that may make it through the calorimetry, as depicted in Figure 2.3.

The muon system composed of wire drift chambers (CMU, CMP, CMX and MBU) and scintillators (CSP, CSX and BSU) as explained in the next paragraph. Muons that pass through the wires will be reconstructed as a stub. Because muons stubs

can also be generated by cosmic ray muons, timing information helps identify muons that come from collisions. Scintillators paired with the drift wires provide timing information that is used to reduce the cosmic background. The reconstructed tracks in the tracking system (COT and silicon) are then extrapolated to these stubs to form muon candidates. A track which is linked to the CMX is called a CMX muon, while a track linked to both the CMU and CMP is called a CMUP muon. Both of these muons are used in our measurement. The muon identification algorithm is discussed further in section 3.2 .

Being located at the outermost part of the CDF detector, the muons detectors are constrained by the size and geometry of the collision hall (B0), sometimes resulting in unusual configurations. The Central Muon Detector (CMU), directly outside the CHA calorimeter, sits in the very central region ($\eta < 0.6$). However, this region lacks of steel shielding. Hence, the Central Muon Upgrade (CMP) is constructed with large piece of steel (the return yoke for the solenoid magnet) to reduce the background [92]. This is why CMP forms a box rather than a cylinder. The Central Muon Extension (CMX) extends the chamber coverage to cover the rest of the COT fiducial region, ($0.6 < \eta < 1.0$), using several pieces (arches, keystone and miniskirt) which fit in (on the sides, on top and beneath) the detector respectively. Additional forward coverage , ($1.0 < \eta < 1.5$), is produced by the Barrel Muon Detector (BMU).

The Central Muon Scintillator Upgrade (CSP) and the Central Muon Extension scintillators (CSX) are used for the timing information and are paired with the CMP and CMX respectively. Therefore, they provide the same coverage as the chambers. Similarly, the Barrel Muon Scintillator (BSU) are attached to the BMU. The coverage of the drift chambers CMU, CMP and CMX which are used in our measurement is shown in Figure 2.8

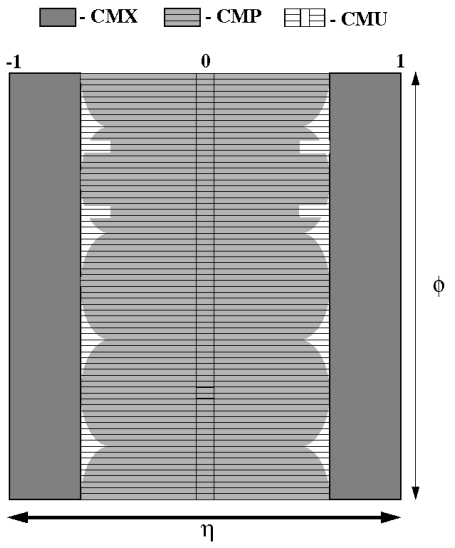


Figure 2.8: Muon coverage in CDF, from [84].

2.4.4 Čerenkov Luminosity Counters

CDF has luminosity counters that sit near the beamline, in the very forward region ($3.7< \eta <4.7$). The Čerenkov Luminosity Counters (CLC), filled with gaseous Isobutane, are designed to detect the Čerenkov radiation that results from charged particles traversing it. The light is collected and sent to PMTs to be read out. The CLC has 3 layers with each layer having 16 counters on each side [93].

This detection rate is converted to a measure of the instantaneous luminosity via:

$$\mathcal{L} = \frac{fnN_pN_{\bar{p}}}{4\pi\sigma_x\sigma_y}\ \mathrm{cm}^{-2}\mathrm{s}^{-1} \tag{2.4}$$

where n is the number of the bunches, f is the frequency of bunch revolution, N_p and $N_{\bar{p}}$ are the total number of protons and anti-protons per bunch, and $4\pi\sigma_x\sigma_y$ represents the transverse area of the beam, assuming a Gaussian distribution of particles with widths σ_x and σ_y in the x and y directions. During the store, the instantaneous luminosity falls off exponentially as expected. On the same token, the integral of the instantaneous luminosity over a given time period (the integrated luminosity)

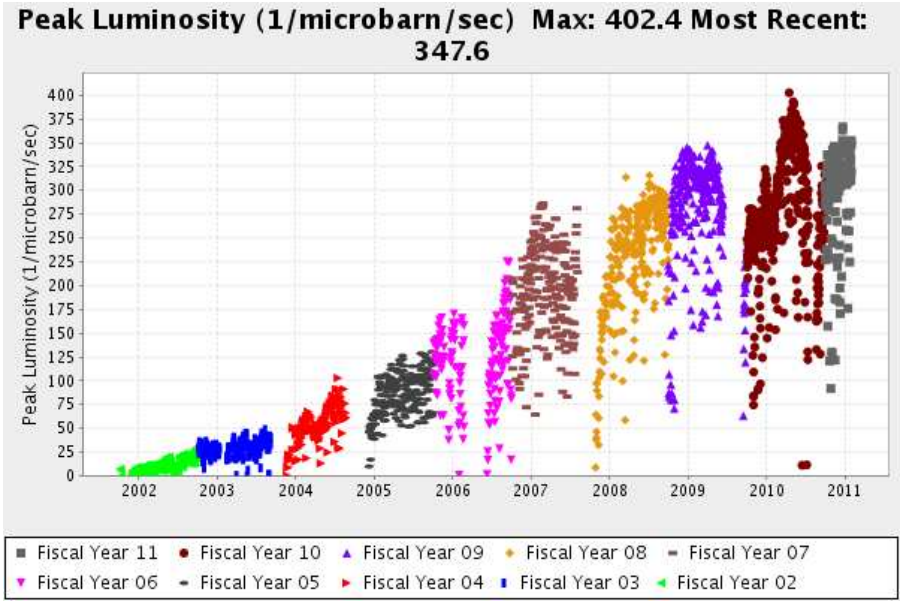


Figure 2.9: The peak instantaneous luminosity at the Tevatron during RunII, from [94].

increases by the time.

The performance of the Tevatron has improved dramatically since the beginning of RunII. Figure 2.9 shows the initial luminosity the stores start with during RunII; the current peak is around $\sim 400{\cdot}10^{30}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The integrated luminosity observed at CDF is shown in Figure 2.10. Currently, more than 10 fb^{-1} of luminosity has been delivered at CDF with more than 8 fb^{-1} stored to tape. By comparison, the total integrated luminosity for RunI was 126 pb^{-1}. The current estimation on the uncertainty of the CDF luminosity measured by CLC is around 5.8% which is derived from the CLC accuracy and the uncertainty on the theoretical cross section for inelastic $p\bar{p}$ collisions. The systematic will be addressed in section 4.2.2 .

2.5 Data Acqusition System

In Section 2.2.3, it was stated that the bunch crossings occur at a rate of 2.5 MHz and several events result from each bunch crossing. With an event size of $\sim 1/4$ MB,

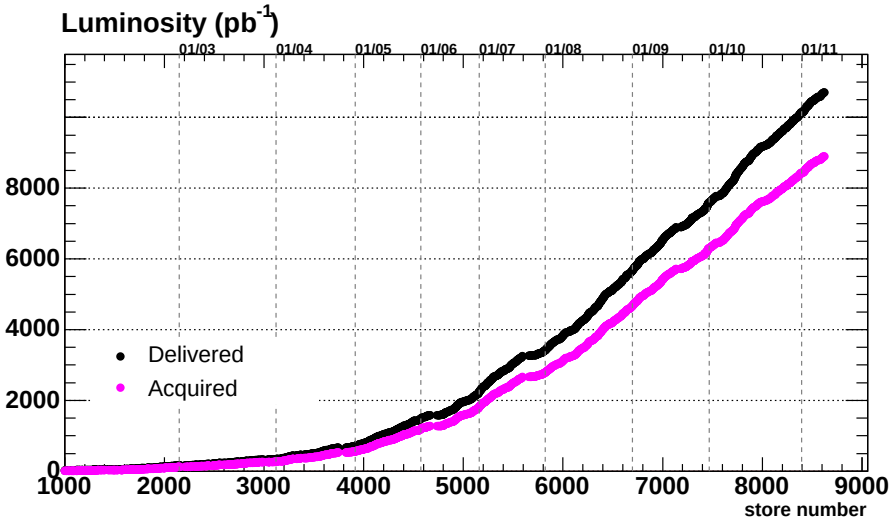


Figure 2.10: The integrated luminosity at CDF during RunII, from [95].

saving the full output would be impractical. On the other hand, a small fraction of the events contain interesting physics. The actual CDF output to tape is about 20 MB/sec [6]. This means that CDF rejects $\sim 99.995\%$ of events!

Therefore, the trigger and acquisition system should be designed to keep -with high efficiency- events that are useful for physics analysis and reject the unwanted ones. It should also be dead-timeless; meaning the system should make a decision before the next crossing occurs [96]. Three tiered-decision levels, called trigger levels, are used to pick the desired events. Each level makes a decision using more detailed information, and thus with more time allowed for the decision than the previous level. An outline of the data flow through the trigger system is shown in Figure 2.11. Each level is described briefly below.

The Level 1 Trigger system (L1) is hardware-based, which reduces the event rate to approximately 25 kHz. The XFT, eXtremely Fast Tracker, reads out only some information from the COT, calorimetry system and the muon chambers and then performs very basic event reconstruction [97]. These reconstructed events are then passed to the Extrapolation Unit (XTRP) to make a decision to keep or reject the

[6] Data in CDF can only be written to a tape at about 200 Hz.

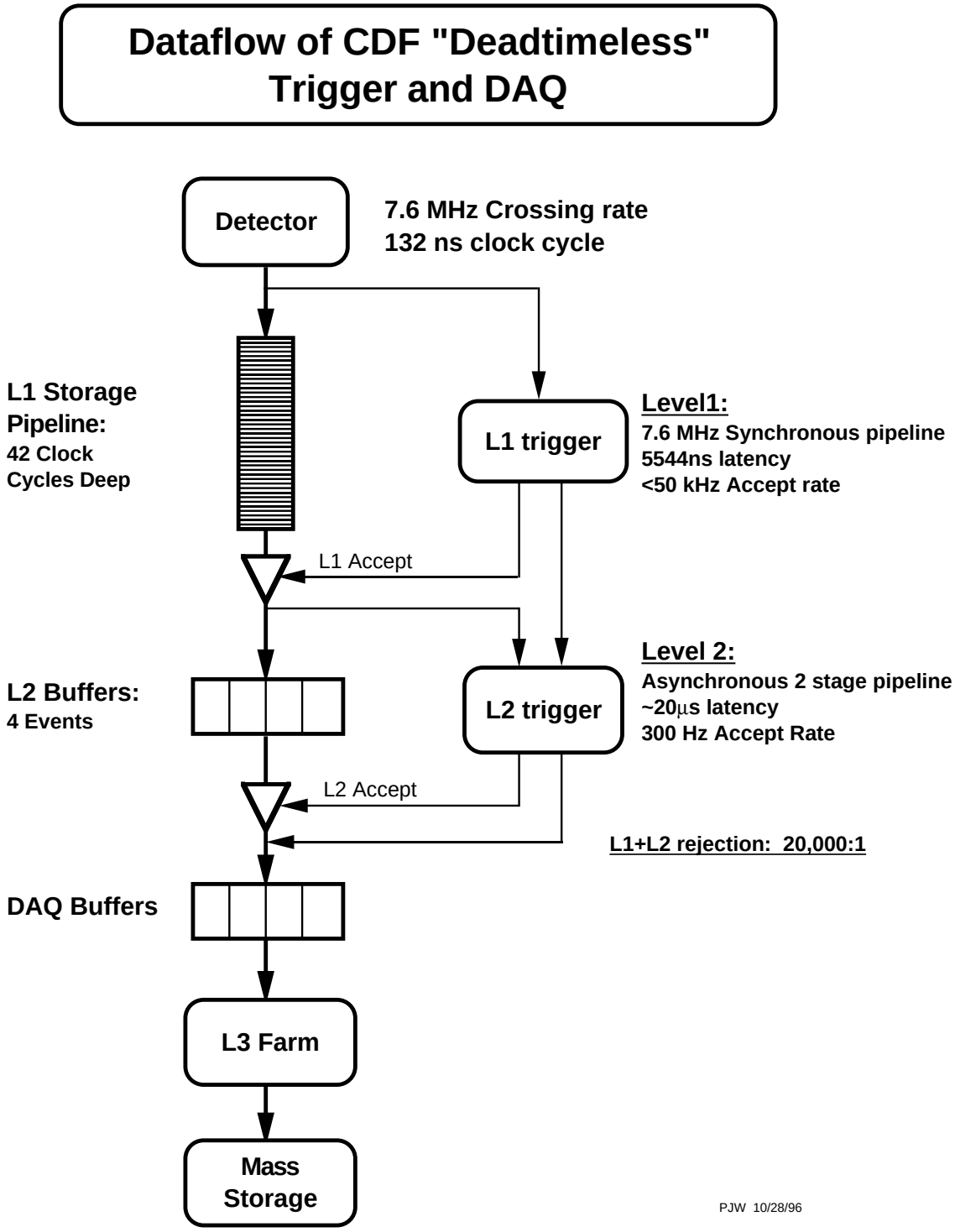


Figure 2.11: Data flow through CDF trigger system, from [84]. The diagram depicts 132 ns bunch crossing times (L1 42 clock cycles deep). However, the current crossing time is 396 ns (corresponds to L1 14 clock cycles deep).

event. These accepted events are then buffered for the Level 2 Trigger system (L2), while the rejected events are simply discarded.

The L2 system is hardware and software-based and reduces the event rate further to $\sim$ 300-400 Hz. One of the hardware components, Silicon Vertex Trigger (SVT), [98], reads Silicon information from the SVX and looks for displaced secondary vertices which are important for our analysis (see section 3.4). L2 reads also the clustering information from the CES. This information is sent to a single computer that analyzes the data and decides whether to accept the event. Like the L1 system, events that are accepted by L2 are buffered for the Level 3 Trigger system (L3), while those that are not are discarded.

The highest level of trigger at CDF, L3, is solely a software-based system, and is implemented as a large computer farm . The events that pass L2 are sent first to the Event Builder (EVB), which is a small farm of CPUs. The L1 and L2 triggers only use a small subset of the event data while making a decision, in order to reduce the time needed, while the rest of the event data is stored in several buffers. The EVB reads out the disparate information and assembles it into a data format readable by the L3 farm. L3 fully reconstructs the event and analyzes its topology before making the final decision to accept or reject the event. The accepted events are then sent to the Consumer Server Logger (CSL), where they stored temporarily before being written to tape for permanent storage at a rate of about 200 Hz. Details about the data set used in this analysis is addressed in section 3.6 .

Within the context of the CDF three-level trigger system, each level has several different trigger paths. Each trigger path has specific acceptance requirements that differ than the other paths. These requirements mimic the selection cuts used in the different analyses. An event then can be accepted by more than one path in the same trigger level and hence passed on to the next trigger level. The $t\bar{t}$+jet triggers used are explained in section 3.2 .

Chapter 3

Event Selection & Background Modeling

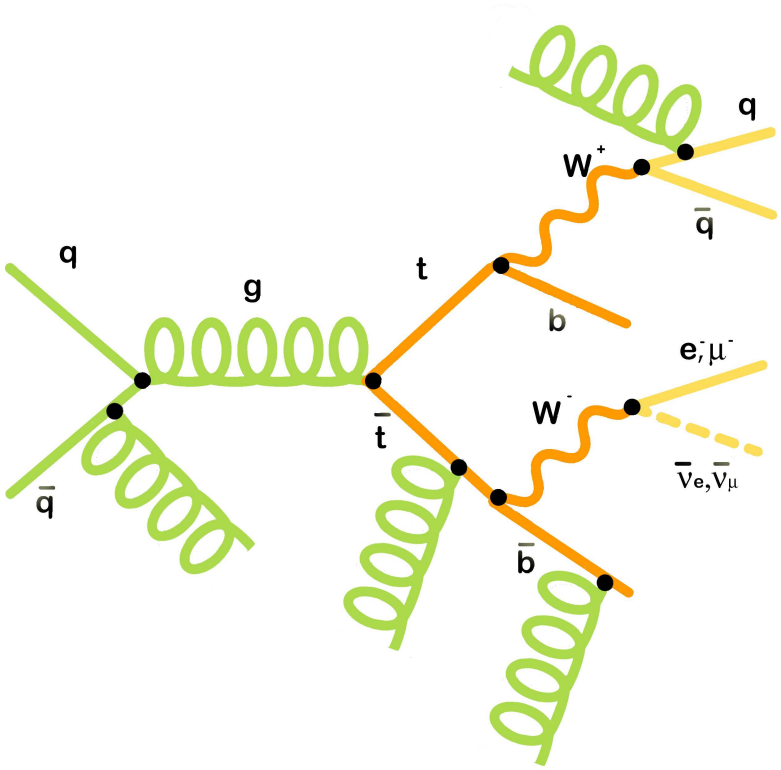


Figure 3.1: A cartoon depicts the topology of $t\bar{t}$+jets event where extra hard jets are produced by the ISR, FSR and radiation off from the decay products of the top quark and the hadronic W boson.

3.1 Introduction

In order to measure the cross section for $t\bar{t}$+jet as described in Equation (1.53), one needs to quantify the signal and the background events and to model the efficiency for detecting the signal. This means that we need to associate: the tracks in the tracking detectors, the clusters in the calorimetry and the stubs in the muon chamber (as discussed in detail in chapter 2) with the physical objects: electrons, muons, jets and neutrinos. This chapter details how we: trigger leptons and select them afterword, cluster jets and tag those that come from b quarks, calculate the $\not\!\!E_{\mathrm{T}}$ and the total transverse energy (H_{T}), cluster the extra hard jet, and finally simulate the different backgrounds of our sample and estimate their contributions. The main goal we keep in mind while we are performing all these tasks is to maximize the number of signal events ($N_{t\bar{t}+\mathrm{jet}}$) and to minimize the background ones (N_{bkg}).

3.2 Lepton Identification

The identification of electrons, muons, jets etc at CDF begins during data-taking with the online CDF trigger system, which was described in section 2.5. Each trigger path applies only basic selection requirements. The events that pass these basic selections will be processed offline after being written to tape. This allows more comprehensive and tighter requirements to be applied, which in turn allows for a reconstruction of the objects with a higher quality. A complete information can be found in [99].

3.2.1 CEM, CMUP & CMX Triggers

We use three lepton path triggers for $t\bar{t}$+jet events: one for electrons detected in the CEM detector and two for muons detected in the CMUP (CMU & CMP) and CMX detectors. The trigger path dataset for electron is `ELECTRON-CENTRAL-18` which is a composite of three CDF trigger level (`L1-CEM8-PT8`, `L2-CEM16-PT8` and

Trigger Level	Property	Requirement
Level 1	E_{T}	≥ 8 GeV
	P_{T}	≥ 8 GeV
	$E_{\mathrm{HAD}}/E_{\mathrm{EM}}$	≤ 0.125
Level 2	E_{T}	≥ 16 GeV
	P_{T}	≥ 8 GeV
	$E_{\mathrm{HAD}}/E_{\mathrm{EM}}$	≤ 0.125
Level 3	E_{T}	≥ 18 GeV
	P_{T}	≥ 8 GeV
	$E_{\mathrm{HAD}}/E_{\mathrm{EM}}$	≤ 0.125
	L_{shr}	≤ 0.4

Table 3.1: `ELECTRON-CENTRAL-18` CEM Trigger.

Trigger Level	Property	CMUP Requirement	CMX Requirement
Level 1	P_{T}	(CMU) ≥ 6 GeV	≥ 6 GeV
	COTP_{T}	≥ 4 GeV	≥ 8 GeV
	# hits	≥ 1 CMP stub	≥ 4 COT hits
Level 2	COTP_{T}	≥ 8 GeV	no trigger
Level 3	COTP_{T}	≥ 18 GeV	≥ 18 GeV

Table 3.2: `MUON-CMUP18` and `MUON-CMX18` Triggers.

`L3-ELECTRON-CENTRAL-18`). The trigger path for CMUP muons is `MUON-CMUP18`, with three corresponding triggers (`L1-CMUP6-PT4`, `L2-CMUP6-PT8` and `L3-MUON-CMUP-18`) while the CMX muon dataset is `MUON-CMX18` with its triggers (`L1-CMX6-PT8` and `L3-MUON-CMX-18`). Each level applies stricter cuts than the previous one. The CEM trigger requirements are given in Table 3.1 while the CMUP and CMX triggers requirements are summarized in Table 3.2.

An electron is identified as an isolated track in the COT matched to an EM calorimeter cluster. At level 1 (Table 3.1), its calorimeter transverse energy (E_{T}) is required to be approximately the same [1] as its COT track momentum P_{T}. The ratio of energy deposited in the hadronic calorimeter to that in the electromagnetic, $E_{\mathrm{HAD}}/E_{\mathrm{EM}}$, should be small to reject showers from hadrons that might imitate

[1] The ratio of E_{T} to P_{T} for a true electron depositing all of its energy in the EM calorimeter is 1 . It can become larger than 1 if the electron emits photon while traveling through COT. This reduces the electron's momentum, and because the photon tends to be collinear with the electron, the calorimeter energy doesn't decrease.

electromagnetic showers. The calorimeter energy restriction is tightened more at level 2 while the track requirement and the ratio remain the same. Finally, at level 3, the calorimeter energy has to be at least 18 GeV and both the track requirement and the ratio are also enforced. In order to remove more background from hadronic showers, a lateral sharing profile requirement for the EM shower (L_{shr}) is applied. This variable, which quantifies the distribution of EM energy in adjacent EM calorimeter cells, must be consistent with the expected profile for electrons derived from test beam data.

Likewise, the muon is identified as an isolated track in the COT matched to a stub in the muon chambers. At level 1 (Table 3.2), a $\geq$ 6 GeV stub in the CMU chamber should be extrapolated to a $\geq$ 4 GeV track in the COT. In order to label this stub as a CMUP muon, there must be also at least one track stub in the CMP. Tracks of $\geq$ 18 GeV should be extrapolated to both CMU and CMP at level 3. The distance between the tracks and the stubs should be $\leq$ 10 cm and 20 cm for CMU and CMP respectively. In order to reduce fake muons that result from particles that penetrate all the way through to the muon chambers, an ionizing energy threshold is applied at level 2.

The forward CMX trigger operates very similarly. However, it's efficiency is not as great as the CMUP because the coverage is not as extensive. A $\geq$ 6 GeV stub in the CMX chamber should be extrapolated to a $\geq$ 8 GeV track in the COT at level 1. While no additional requirements are applied at level 2, tracks of $\geq$ 18 GeV should be extrapolated to CMX chamber (within $\leq$ 10 cm) at level 3.

3.2.2 Offline Selection

The offline lepton selection imposes tighter cuts than those used in the triggers in order to purify the sample by removing fake leptons. Tables 3.3 and 3.4 respectively describe the standard cuts used for electron and muon selection at CDF. Leptons that pass these cuts are referred to as isolated or (tight) leptons; otherwise, they are

Property	Requirement
$E_{\rm T}$	≥ 20 GeV
$\lvert\eta\rvert$	≤ 1.0
$P_{\rm T}$	≥ 10 GeV
$E_{\rm HAD}/E_{\rm EM}$	$\leq 0.055+(0.00045)E$
L_{shr}	≤ 0.2
$E_{\rm T}/P_{\rm T}$	≤ 2 GeV (if $E_{\rm T} < 100$ GeV)
Isolation	< 0.1
COT Quality	True
Photon Conversion	False
$Q \cdot \Delta x$ (cm)	$-3.0 < Q \cdot \Delta x < 1.5$ and $\lvert\Delta z\rvert < 3.0$
$\chi^2_{\rm strip}$	≥ 10

Table 3.3: Offline selections for tight electron.

non-isolated or (loose).

Table 3.3 shows the tighter requirements for each of the electron variavles $E_{\rm T}$, $P_{\rm T}$, $E_{\rm HAD}/E_{\rm EM}$ and L_{shr} required at the the offline level.

A true electron deposits almost all of its energy in a single calorimeter cell i.e. its energy is well collimated. Isolation is defined as:

$$\text{Isolation} = \frac{E_{\rm T}^{\Delta {\rm R}<0.4} - E_{\rm T}^{\rm Electron}}{E_{\rm T}^{\rm Electron}} \tag{3.1}$$

where $E_{\rm T}^{\Delta {\rm R}<0.4}$ is the the energy deposited in EM calorimeter cells surrounding the cell containing the electron within a cone of radius ($\Delta R < 0.4$) (see Equation (2.3)). The isolation variable should be less than 0.1 in order to discriminate between the true electrons and some jets that fake electrons. The COT track quality cut demands that the electron's track must have hits in at least 3 axial and 2 stereo superlayers (see section 2.4.1). Such an electron is accepted for physics analysis. Photons passing through matter can convert into electron-positron pairs. An electron resulting from photon conversion is rejected by searching for its corresponding positron track; an opposite charged track separated from the electron by less than 2mm in the $r-\phi$ plane at the point of closest approach. If such tracks are found, the electron is rejected

because it is more likely to come from photons and not from a hard collision of interest. The COT tracks should be extrapolated to the CES calorimeter (see section 2.4.2). The difference between the extrapolation position and the actual hits in the CES in the $(r-\phi)$ plane (Δx) is required to be small as shown in the Table, where Q is the charge of the track. On the other hand, a similar difference in the $(r-z)$ plane $(|\Delta z|)$ should be less than 3.0 cm. Lastly, $\chi^2_{\rm strip}$ is comparing the shape of the shower profile measured in the CES to the expected shower profile of the electron.

Table 3.4 summarizes the standard cuts that are applied to all muons. However, additional cuts are required for the different muon chambers. An important cut is the minimum ionization energy that the muon should leave in the calorimetry system, as discussed in the muon trigger. This will remove fake muons that punch through the calorimeters, like pions (in the EM calorimeter), and energetic hadrons (in the HAD calorimeters). As for electrons, Isolation and COT quality cuts are enforced for muons [2]. A Δx cut for the stubs in the different subdetectors is also applied. Cosmic ray vetos and impact parameter cuts (d_0) are applied to reject muons from cosmic rays. Cosmic rays passing through the detector would leave two tracks in the muon chamber separated in ϕ by 180° (back-to-back) and separated in time as well. Using a specialized algorithm, such cosmic muons are vetoed (see section 2.4.3) . On the other hand, the distance between the position of the primary vertex and the z axis, d_0, must be small. If there is no match with the Silicon detector, then the impact parameter cut is made looser. Besides removing cosmic ray muons, the d_0 cut also rejects the decays-in-flight of pions $(\pi \rightarrow \mu\nu)$ and kaons $(K \rightarrow \mu\nu)$ that do not originate from the primary vertex.

[2] However, the muon's Isolation variable is defined differently: $E_{\rm T}^{\Delta R<0.4}/P_{\rm T}^{\mu}$.

Property	Requirement
E_{T}	≥ 20 GeV
$\lvert\eta\rvert$	≤ 1.0
E_{EM}	$<$ Maximum$(2.0\,, 2.0 + 0.0115(P - 100.0))$
E_{HAD}	$<$ Maximum$(6.0\,, 6.0 + 0.0280(P - 100.0))$
Isolation	< 0.1
COT Quality	True
Δx (cm)	< 3.0 (CMU), < 5.0 (CMP), < 6.0 (CMX)
Cosmic Rays Veto	True
d_0 (cm)	< 0.02 (with Silicon) and < 0.2 (with no Silicon)

Table 3.4: Offline selections for tight muon.

3.3 Jet Reconstruction & Corrections

The concept of hadronization and forming jets was introduced earlier in section 1.3.2. When the colored partons receive large amount of energy in a collision, they hadronize into color neutral hadrons. Although the state of the event before hadronization is not a physical observable, it is useful when discussing the phenomenology of hadron physics and is referred to as the parton level. This is used as the basis for the QCD MC event generators as will be addressed in section 3.7 . These partons, while interacting with each other, radiate quarks and gluons. The latter, in turn, can also radiate more quarks and gluons and this chain of radiation is called a parton shower. After all the partons in the event hadronize, the colorless content of the event (such as pions and kaons) is referred to as the hadron or particle level. The particles at this level are observables, and they are the states which interact with the calorimetry system. After the particles interact with the calorimetry system, the resulting description of the event is referred to as the detector or calorimeter level [101]. In other words, what we observe/measure in the calorimeter is the jet that is collimated in the direction of the parton. However, the MC generators can predict all the three levels. The three levels of the event are depicted in Figure 3.2.

In principle, the sum of the total energy at the calorimeter level within a cone around the direction of the parton level should be strongly correlated with the parton's

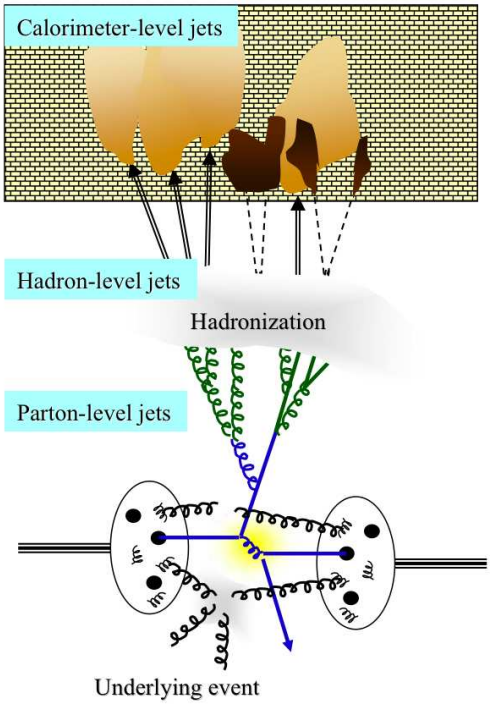


Figure 3.2: The three levels of a jet at CDF event. The time flow direction in the cartoon is upward, from [100].

energy. The task left then is to reconstruct/cluster the calorimeter jet in order to measure its energy.

Jets at the calorimeter level are first reconstructed offline using a cone-based clustering algorithm called `JETCLU` [102]. Briefly, an E_{T} ordered list of seed towers with $E_{\mathrm{T}} > 1.0$ GeV is created. Preclusters are formed from the adjacent towers, provided that the towers are within a cone of $\Delta R \leq 0.4$ of the seed [3]. The centroid of the jet is calculated by summing the $\eta - \phi$ positions of the towers weighted by their E_{T}. The preclusters in the cone centered on this new centroid are then summed (the new centroid is iterated until the jet is stable, i.e. the set of towers contributing to the jet is stable and a single tower does not contribute to more than one jet). Overlapping jets are merged if they share $\geq 75\%$ of the smaller jet's energy. Otherwise, the towers are assigned to the closest cluster.

[3] A smaller cone size would let too much information of the parton level to be lost. On the other hand, a much larger size would not discriminate the multiple jets expected in $t\bar{t}$ events.

However, the jet energy measured is raw jet energy and has to be corrected for various detector and physics effects before the experimental measurement can be compared with theoretical and MC predictions. The jets can be corrected to the hadron level and/or the parton level. The precision to which this correction can be achieved determines the precision of many measurements. For example, a 1% uncertainty on the overall jet energy correction (referred to as Jet Energy Scale – JES) results in a 1 GeV uncertainty on the top quark mass [103]. The JES systematic uncertainties for our measurement will be addressed in section 4.2.2 .

The parton transverse momentum $P_{\mathrm{T}}^{\mathrm{parton}}$ is related to the calorimeter transverse momentum $P_{\mathrm{T}}^{\mathrm{calorimeter}}$ and the particle transverse momentum $P_{\mathrm{T}}^{\mathrm{particle}}$ through a series of corrections [104]:

$$P_{\mathrm{T}}^{\mathrm{parton}} = (P_{\mathrm{T}}^{\mathrm{calorimeter}} \times C_{\eta} - C_{\mathrm{PU}}) \times C_{\mathrm{Abs}} - C_{\mathrm{UE}} + C_{\mathrm{OOC}} = P_{\mathrm{T}}^{\mathrm{particle}} - C_{\mathrm{UE}} + C_{\mathrm{OOC}} \quad (3.2)$$

This equation manifests a total of five levels of corrections, numbered 1 through 7 (levels 2 and 3 correspond to corrections which are no longer used at CDF).

The corrections are:

- Level 1: Relative Correction (C_{η}).

  The CDF calorimeter response to jets is not flat in η. The non-uniformity in η arises from the separation of calorimeter components (cracks) at $\eta = 0$ where the two halves of the central calorimeter join and at $|\eta| \approx 1.1$, where the plug and central calorimeters join. The response varies across η also due to the different amounts of material that both the plug and central calorimeters have. This makes the central region ($0.2 < |\eta| < 0.6$) better understood and instrumented than the plug region. Using a sample of pure dijet events, the relative correction is derived by balancing the E_{T} of a jet in the central region against a random (probe) jet in the other different regions. Since the relative jet energy response

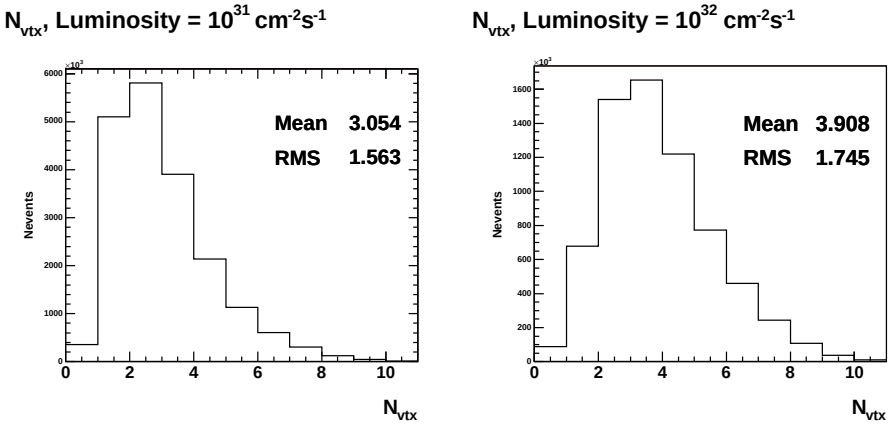


Figure 3.3: The dependence of $\langle N_{vtx} \rangle$ on the instantaneous luminosity $\mathcal{L}$.

is different between data and MC, corrections are derived separately for data and MC.

- Level 4: Pileup Correction ($C_{\rm PU}$).

  At high instantaneous luminosity, more than one $p\bar{p}$ interaction can occur in the same bunch crossing at the Tevatron due to the large $p\bar{p}$ cross section. The average number of $p\bar{p}$ interactions per bunch crossing (pileups) depends linearly on the instantaneous luminosity. The number of the reconstructed primary z-vertices, N_{vtx}, is the best estimate of the pileups. This is shown [4] in Figure 3.3 and Figure 3.4 . Each extra interaction is found to add ~ 0.3 GeV of energy on average [100]. This extra energy should be subtracted from the jet energy.

- Level 5: Absolute Correction ($C_{\rm Abs}$).

  By applying the absolute correction, the calorimeter jet can be fairly compared with the particle jet. As particles pass through the CDF calorimeters, some of their energy is not collected due to calorimeter cracks and the overall nonlinear detector response. The absolute correction is derived entirely using MC by comparing particle level and calorimeter level jets. After this correction, the

[4] I used 1.74 fb^{-1} of data written to Jet20, Jet50, Jet70 and Jet100 trigger paths and `JETCLU` of $\Delta R \leq 0.7$ to generate these histograms.

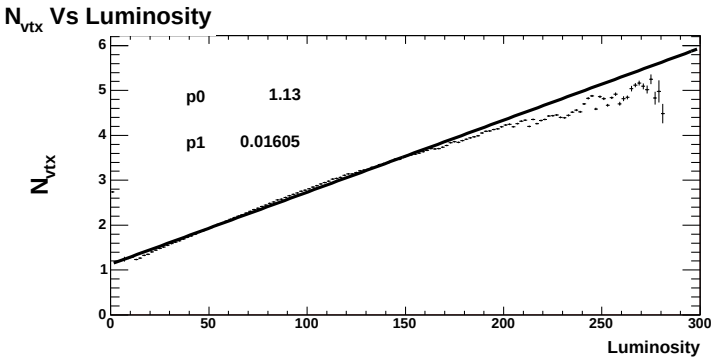


Figure 3.4: $\langle N_{vtx} \rangle$ verses the instantaneous luminosity $\mathcal{L}$.

energy scale of a jet is independent of the CDF detector. All of the jets used in our measurement are corrected up to level 5 i.e. they have the absolute correction applied.

- Levels 6 & 7: Underlying Event Correction ($C_{\rm UE}$) & Out-Of-Cone Correction ($C_{\rm OOC}$).

  Levels 6 and 7 corrections are not applied in correcting the jets in our measurement, but are only used for the systematic uncertainty calculations. In order to correct the calorimeter jet further back to the parton level, both of these corrections should be involved.

  Most of the time when p and $\bar{p}$ beams collide, they pass through each other and undergoes a soft collision. The UE is everything in a collision except the production of the hard scattering (the UE is also referred to as the beam remnants). It is an unavoidable background to many collider observables and can not be separated out on an event by event basis. By studying distributions which are sensitive to the UE, the MC can be tuned to fit the effects observed in data. The correction thus derived then removes the UE energy associated with the beam remnant. [105].

  On the other hand, the OOC correction accounts for the energy lost due to the

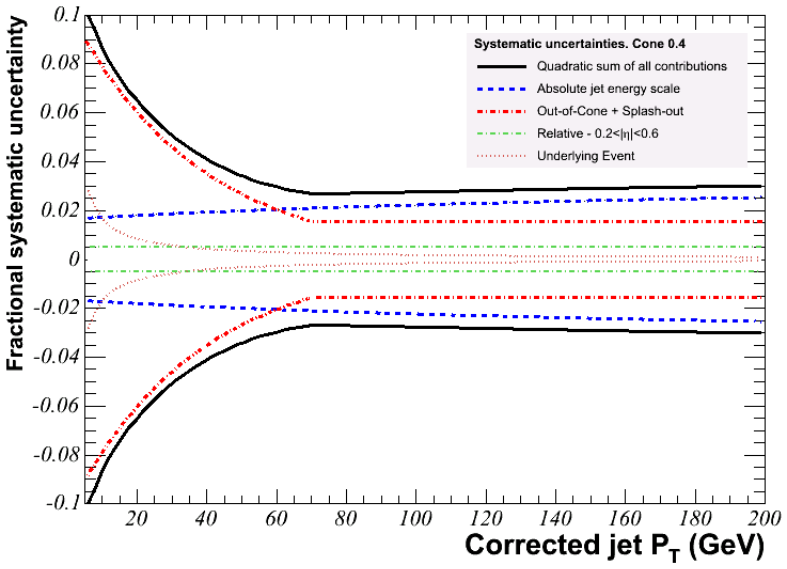


Figure 3.5: The JES total systematic uncertainties, from [100].

finite size of the jet cone (0.4 in our case). Part of the parton jet energy can lie outside the cone, either because of fragmentation effects or FSR. For example, this OOC correction adds $\sim$ 2.2 GeV of energy to a 20 GeV particle jet [100]. On average, the UE correction tends to partially remove the OOC correction.

The combined systematic uncertainties on the JES summing all these corrections are shown in Figure 3.5. For a high P_{T} jet, the uncertainty shifts the jet energy by $\sim$ 3% while at low P_{T}, the uncertainty can shift the jet energy by $\sim$ 10%. Converting the JES uncertainty into a systematic uncertainty on our own measurement, the resulting systematic uncertainty is by far the dominant source of uncertainty in our result. This will be further described in section 4.2.2 .

3.4 b-jet Tagging

In section 1.4.2, I discussed the importance of identifying the jets that come from b quarks. This tagging process, takes advantage of the relatively long lifetime of the B hadrons [5]. This will give the B hadrons a chance to travel, transversely away from the primary vertex, a few millimeters ($\sim$ 5mm) before they decay. The tracks that result from the displaced decay (referred to as the secondary vertex) play the key role in tagging the b-jets. With the secondary vertex tagging algorithm (`SecVtx`) [106], using the excellent position resolution the Silicon detector provides (as discussed in section 2.4.1), CDF can efficiently tag b-jets.

The `SecVtx` algorithm uses a large number of selection requirements that can be found in Ref. [106]. Very briefly, the algorithm starts by using the precise knowledge of the primary vertex itself. It looks then for a large value for the impact parameter, d_0, (defined in Table 3.4) and for its significance (d_0 devided by its uncertainty). The distance in the transverse plane that the secondary vertex is displaced from the primary vertex (L_{xy}) plays a very key role in the `SecVtx` algorithm. It determines whether the vertex is sufficiently far from the primary vertex. The algorithm looks for L_{xy} significance as well. However, L_{xy} is signed relative to the jet direction. It is defined to be positive when the secondary vertex is displaced in the same direction as the jet (a signature of a b-jet). A negative L_{xy} (i.e. the vertex displaced on the opposite side) indicates a tagging of jets that do not contain true b quarks (i.e. the light jets *u, d, s*). This case is called a mistag. A cartoon depicting both b-tagging and mistagging is shown in Figure 3.6. If the predicted secondary vertex passes all the selection requirements, the jet is marked as a tight `SecVtx` jet. With less stringent selections, the `SecVtx` algorithm identifies a loose `SecVtx` jet as well [6]. We use the tight `SecVtx` tag jet in our measurement.

[5] Included in these hadrons are B mesons (B^+, B^- and B^0) and B baryons (Ξ and Σ).

[6] The ultra-tight `SecVtx` tagged jet is the third tune the `SecVtx` algorithm is designed to identify.

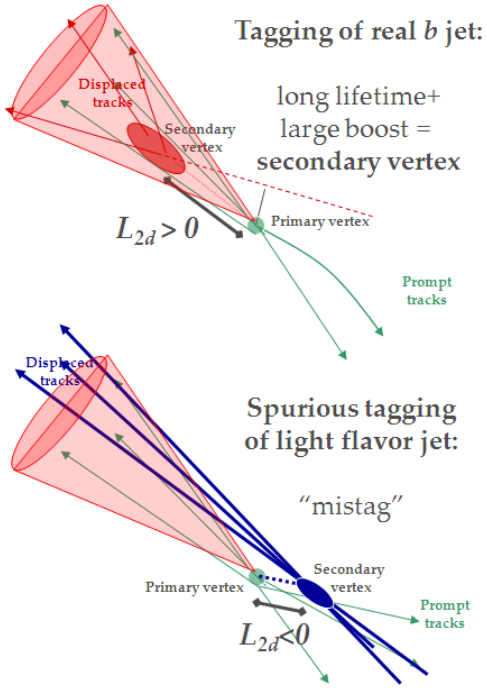


Figure 3.6: b-tagging and mistag. L_{2d} in the cartoon is L_{xy} defined in the text, from [107].

Due to the complexity of the `SecVtx` algorithm, the MC samples can not simulate the tagging process properly. As a matter of fact, MC events have more tagged b-jets than CDF data has. However, this overestimated efficiency can be compensated with the use of a scale factor. The latter is the ratio of the data over MC efficiencies. Applying this factor to the MC tagging rate gives a corrected estimate of the efficiency of b-tagging in the MC. The scale factor is calculated to be 0.95 ± 0.05 [107]. Unlike the b-tagging, MC samples tend to underestimate the mistag rate compared to the data. The mistag rate is badly modeled in MC as has a strong dependence on the kinematic distributions of the tagged events. The mistag rate can be parametrized by five jet variables (jet E_{T}, H_{T}, η, ϕ and number of good SVX tracks). Using this mistag matrix gives the mistag rate [107].

Figure 3.7 shows the CDF efficiency to tag b-jets in top quark MC samples as a function of the jet E_{T}. The bands represent the systematic error on the scale factor.

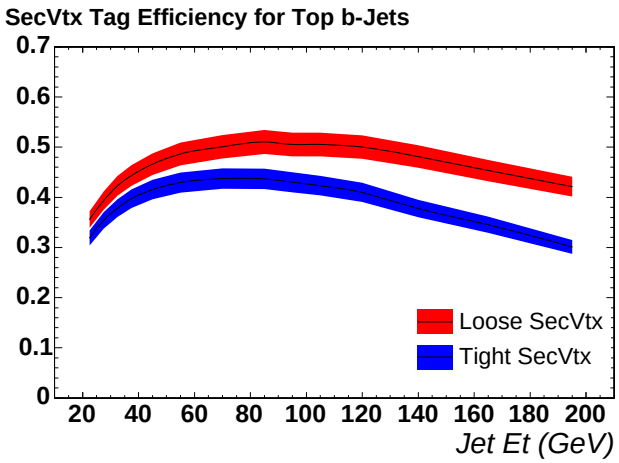


Figure 3.7: CDF `SecVtx` algorithm b-tagging efficiency, from [107].

While the efficiency for finding a tight b-jet is only about 40%, the efficiency for a mistag is less than 1%, as shown in Figure 3.8. This is a clear manifestation of how b-tagging increases our sample purity. One more feature that we can read from both plots is that while the loose b-jet requirement increases the acceptance, and hence the statistics of our sample, it does decrease, on the other hand, the sample's purity. The decrease in efficiency at high jet E_{T} is due to the declining yield of good Silicon tracks passing the quality cuts.

However, the `SecVtx` algorithm can not distinguish a b-jet from a charm jet. This is why `SecVtx` algorithm is sometimes called the heavy flavor tagging algorithm. The charm jet tagging efficiency is about 8%. Because it's scale factor is not measured, the same b jet scale factor is applied to jets matched to charm, with the uncertainty doubled to be conservative. The systematic uncertainties for both the b-tagging and the mistag will be addressed in section 4.2.2 .

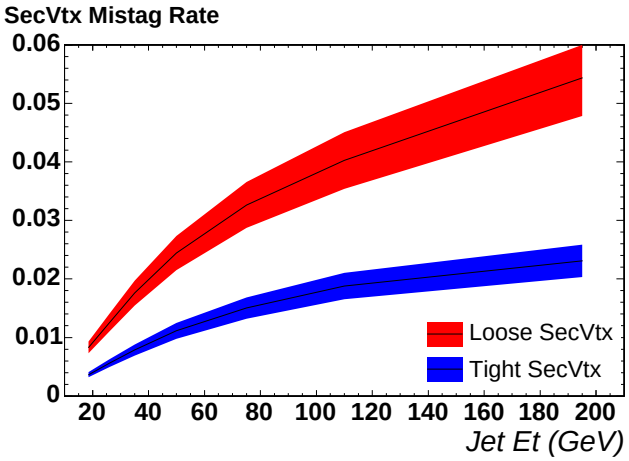


Figure 3.8: CDF `SecVtx` algorithm mistag efficiency, from [107].

3.5 Missing E_{T} & H_{T} Reconstructions

I mentioned in section 2.3 that the $\not\!\!E_{\mathrm{T}}$ is the distinctive signature of neutrinos in our $t\bar{t}$+jet sample. In general, the $\not\!\!E_{\mathrm{T}}$ could be caused by poor energy measurement from particles passing through the cracks in CDF and due to detector noise. Another source for the $\not\!\!E_{\mathrm{T}}$ is from cosmic ray events. However, a large portion of the $\not\!\!E_{\mathrm{T}}$ is due to neutrinos. In lepton plus jet channel, many of the backgrounds in our sample, which have no final state neutrinos, can be removed by applying a $\not\!\!E_{\mathrm{T}}$ cut.

The missing transverse energy vector ($\overrightarrow{\not\!\!E_{\mathrm{T}}}$) is defined as follows:

$$\overrightarrow{\not\!\!E_{\mathrm{T}}} = -\sum_{i} E_{\mathrm{T}}^{i}\hat{\mathbf{n}}_{i} \tag{3.3}$$

where $\hat{\mathbf{n}}_i$ is a unit vector perpendicular to the beam axis and pointing from the primary vertex to the i^{th} calorimeter tower. The sum in this equation is over all the towers in the calorimeter. The energy E_{T} is the raw energy measured in the towers. The $\not\!\!E_{\mathrm{T}}$ is simply the absolute magnitude of the $\overrightarrow{\not\!\!E_{\mathrm{T}}}$ after correcting the E_{T} up to level 5 and correcting the primary vertex position at the trigger level. This definition applies exactly for events with electrons. However, if muons are found in

the event (which leave minimal energy in the calorimeters), the $\not\!\!E_{\rm T}$ is corrected by first subtracting the muon $E_{\rm T}$'s, and adding instead, the muon $P_{\rm T}$'s to the vector sum in the equation. In our measurement, of $t\bar{t}$+jet events, we require $\not\!\!E_{\rm T} > 20$ GeV.

The total sum of the transverse energy of the event in the detector, $H_{\rm T}$, is simply the scalar sum of the transverse energy of all the physical objects the event contains. Therefore, $H_{\rm T}$ is the sum of the electron $E_{\rm T}$ or the muon $P_{\rm T}$, the $E_{\rm T}$ of all jets and the $\not\!\!E_{\rm T}$. For an event such as a $t\bar{t}$+jet event, the $H_{\rm T}$ is expected to have a larger value than the backgrounds. The standard $H_{\rm T}$ cut used by the Top Physics Group at CDF is 200 GeV. However, to further reduce the backgrounds, the $H_{\rm T}$ in our measurement is required to be at least 220 GeV.

3.6 Selection Cuts & Data Sets

The cartoon in Figure 3.9 depicts a $t\bar{t}$+jet event topology in the lepton plus jet channel. We thus would expect to see 5 jets, of which two are b tagged, one lepton and large $\not\!\!E_{\rm T}$ in the final state. Consequently, each event is required to have in particular the following cuts:

- One tight electron (muon) with $E_{\rm T}$ ($P_{\rm T}$) ≥ 20 GeV in the region $|\eta| \leq 2$ (CEM, CMUP and CMX) as described in section 3.2 .

- At least three tight L5 corrected jets with $E_{\rm T} \geq 20$ GeV and $|\eta| \leq 2$ (see section 3.3).

- At least one of the jets above must be a tight `SecVtx` tagged jet (see section 3.4).

- $\not\!\!E_{\rm T} \geq 20$ GeV & $H_{\rm T} \geq 220$ GeV (see section 3.5).

- Dilepton veto: in order to separate the $t\bar{t}$+jet lepton plus jet channel from the dilepton channel, any event with a second tight lepton is removed.

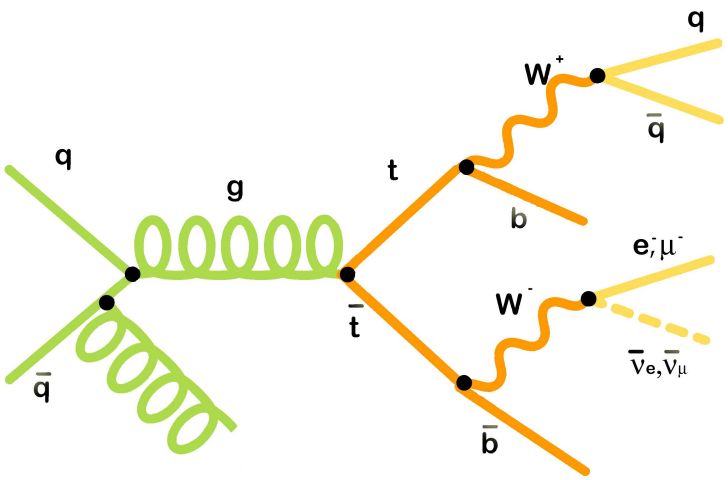


Figure 3.9: $t\bar{t}$+jet event's topology where the extra hard jet is produced by the ISR.

- Z boson veto: the events are removed if a tight lepton and a second object form an invariant mass consistent with a Z boson ($76 < M_{ll} < 106$ GeV). If the tight lepton is an electron, the second object must be an isolated EM object, a reclustered jet with 95% of energy deposited in the EM calorimeter, or an opposite-signed isolated track. If the tight lepton is a muon, the second object must be an isolated muon or an opposite-signed isolated track.

- Primary vertex reconstruction: The z-position of the primary vertex is used to cluster jets and to ensure that leptons and jets belong to the same interaction. To ensure the lepton originates from the primary vertex, events are rejected if the lepton z-intercept is not within 5 cm of the primary vertex z-position. This reduces the number of events in the sample where jets and the lepton are part of different interactions [108].

Data at CDF is divided into run periods to take into account the detector conditions that may vary over time. In our measurement, we have used data collected through the periods 0 to 22 (from February 4, 2002 to February 10, 2009), corresponding to a total of 4.1 ± 0.2 fb^{-1} of integrated luminosity [109]. All the data sets with the triggers used are listed in Table 3.5. These runs are contained in the Good

Data set ID	Trigger	Run Range
bhel0d	CEM	138425-186598
bhmu0d	CMUP&CMX	138425-186598
bhel0h	CEM	190697-203799
bhmu0h	CMUP&CMX	190697-203799
bhel0i	CEM	203819-228596
bhmu0i	CMUP&CMX	203819-228596
bhel0j	CEM	228664-241664
bhmu0j	CMUP&CMX	228664-241664
bhel0k	CEM	241665-261005
bhmu0k	CMUP&CMX	241665-261005
bhel0m	CEM	261119-272214
bhmu0m	CMUP&CMX	261119-272214

Table 3.5: Summary of the data sets used in our analysis.

3jets	4jets	≥5jets
1115	882	329

Table 3.6: The observed $t\bar{t}$ event counts as a function of jet multiplicity.

Run List v28, as outlined in detail in [109].

After the selection cuts been applied over the data collected, the number of events that pass through, as a function of jet multiplicity, is shown in Table 3.6 . The signal region requires ≥ 3 jets, leading to a total of 2326 events. This represents the number ($N_{t\bar{t}+\text{jet}}$) in Equation (1.53). Next I will estimate the other terms in Equation (1.53).

3.7 MC Models

3.7.1 Event Generator

In the MC world, the role of the collider machine, namely to produce events, is taken over by event generators. As the name indicates, the output of an event generator should be in the form of events, with the same average behavior and the same fluctuations as in real data. An event generator can be used for multiple purposes. It tells us what kind of events one may expect to find, and at what rate. It can be used

as a tool for devising the analysis strategies that should be used on data, so that signal-to-background conditions are optimized. It also helps in estimating detector acceptance corrections that have to be applied to raw data, in order to extract the true physical signal, as we will address in section 3.8.2 [110].

For these purposes, an event generator should contain a simulation of several physics aspects. If we try to follow the evolution of an event in some semblance of time order, one may arrange these aspects as follows: the partonic distribution for the two incoming particles, the ISR shower (spacelike) developed from the partons from each incoming particle, the hard process involving the two incoming partons, the FSR shower (timelike) from the outgoing partons from the hard process. In addition, further semi-hard scatterings may occur between the other partons in the two incoming particles, and a beam remnant is left behind when the shower initiator is taken out of the beam particle. At relatively long time scale, the hadronization of the outgoing partons takes place due to QCD confinement mechanism, and finally the unstable hadrons decay into relatively stable particles [110].

This means that the components/routines of an event generator are: PDFs (for the incoming partonic distributions), Parton Shower Algorithms (for ISR, FSR), Matrix Element Calculator (for the hard process) and Color String Framework (for hadronization). `PYTHIA` [111] and `HERWIG` [112] are the most convenient generators currently used. We used `PYTHIA 6.216` to simulate our signal and then cross-checked the results using `HERWIG 6.510` (for systematic studies). For the background processes in our analysis, `ALPGEN` [113] and `MadGraph/MadEvent` [114] models are used. However, both `ALPGEN` and `MadGraph/MadEvent` do not contain parton showering or color string routines. Thus, both are used to calculate the matrix elements which then are passed to `PYTHIA` or `HERWIG` for parton showering and decay. In our analysis, we used `ALPGEN 2.10` and `MadGraph/MadEvent 4` and passed the information to `PYTHIA`.

The response of the detectors (how the final particles produced by the event gen-

erator traverse the detector, bend in the magnetic fields, shower in the calorimeters, or escape through cracks) is simulated using the CDF detector simulation programs, which use the `GEANT` and `GFLASH` routines [115, 116] . Ideally, the output of this simulation has exactly the same format as the real data recorded by CDF, and can therefore be put through the same event reconstruction and physics analysis chain, except that here we know what the results should be, and so can cross-check the reconstruction chain [110].

Each process that makes up our own signal and the different backgrounds are modeled using the MC samples presented here, and then are passed through the same selection cuts. In section 3.8, I will present the signature of these backgrounds to understand how they are able to mimic our signal and then show in detail how to estimate their contributions to the overall observed events.

3.7.2 The Extra Hard Jet

In order to model (define) the theoretical extra hard jet, the `PlusJetFinder` [7] tool [125], was built to operate on `PYTHIA`. `PlusJetFinder` identifies the objects at the parton/particle levels in `PYTHIA` that arise from ISR, FSR and from the internal decay of the top system. It only counts how often a `PYTHIA` event produces an ISR jet (as depicted in Figure 3.9). It doesn't contain jets produced by UE or jets from the decays products of the top quark or the W boson (Figure 3.10 and Figure 3.11) as these objects are not accounted for in the NLO calculation in [79] (the top quark is stable). However, it includes the small number of objects that are radiated off by the top quark before it decays i.e. the FSR (Figure 3.12).

To identify the extra jet, we used the `SpartyJet` [8] framework [117] with $\Delta R \leq 0.4$ to cluster the objects that are fed by the `PlusJetFinder`. If the jet passes our P_{T}

[7] This tool was written by Steve Mrenna, one of the authers of `PYTHIA` .

[8] The framework bears the name of MSU mascot `Sparty`; a Spartan warrior statue on campus. It was created by Joey Huston's group and currently used at both of Fermilab and CERN.

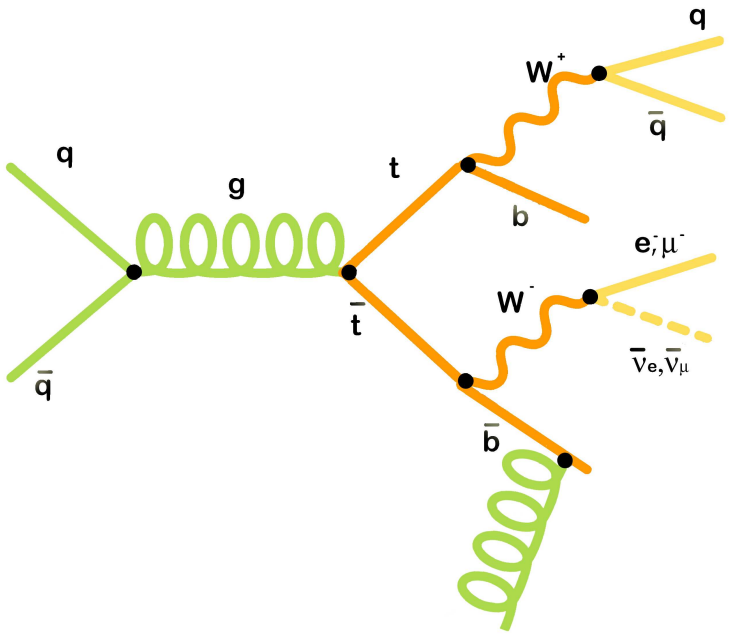


Figure 3.10: $t\bar{t}$+jet event's topology where the extra hard jet is produced by the radiation off of the decay prodcuts of the top quark.

threshold cut (20 GeV to match [79]), then it is counted as an extra hard jet in the $t\bar{t}$ event. In the lepton plus jet channel, $t\bar{t}$ events may manifest with 3, 4, 5 or even 6 jets in the final state, depending on the kinematic and acceptance cuts applied, and the presence of the additional ISR/FSR and decay products. Hence, we separated our sample into two: the $t\bar{t}$+jet sample allows the extra jet to be produced from ISR or FSR but from no other sources and the $t\bar{t}$+0jet sample allows a possible extra jet to be produced from any other source; either from the internal decay of the top system whether or not it passes the threshold cut or from ISR/FSR not passing the threshold cut. It also contains the $t\bar{t}$ pair events. This separation has allowed us to relate the $t\bar{t}$+jet sample to the NLO calculation of [79]. This will let us also know the probability for the extra jet to be radiated from each of the two cases. When the cross section values of both of the samples are combined, the inclusive cross section of $t\bar{t}$ events can be determined and hence compared to the results listed in Table 1.3 .

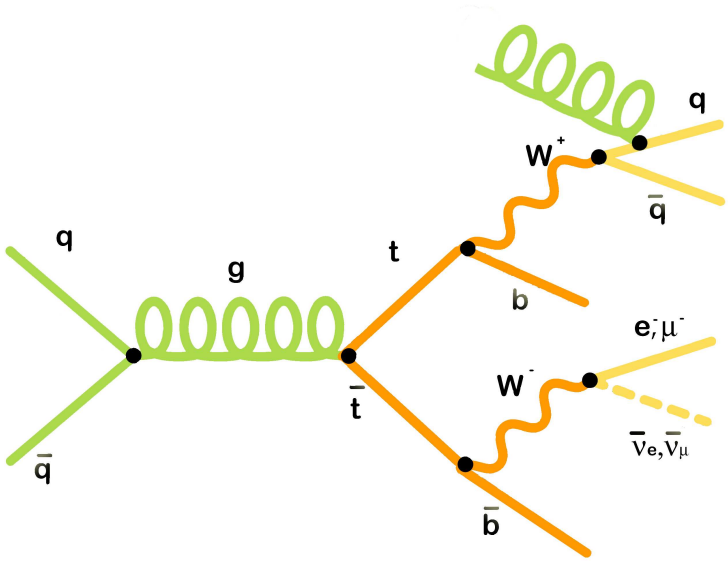


Figure 3.11: $t\bar{t}$+jet event's topology where the extra hard jet is produced by the radiation off of the decay prodcuts of the hadronic W boson.

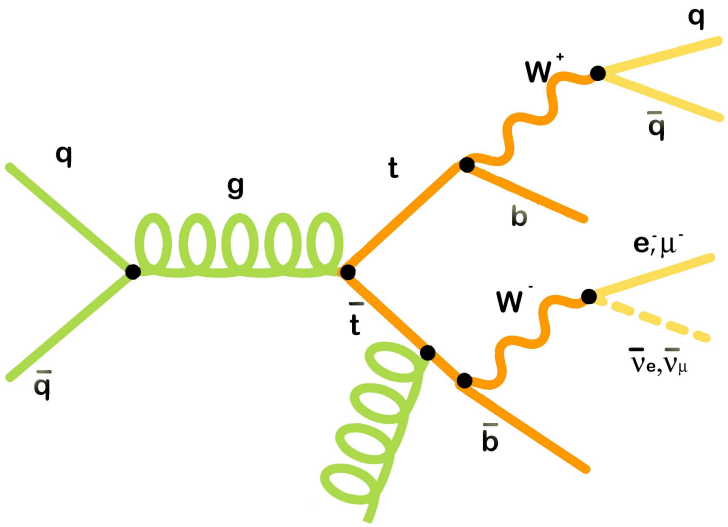


Figure 3.12: $t\bar{t}$+jet event's topology where the extra hard jet is produced by the FSR.

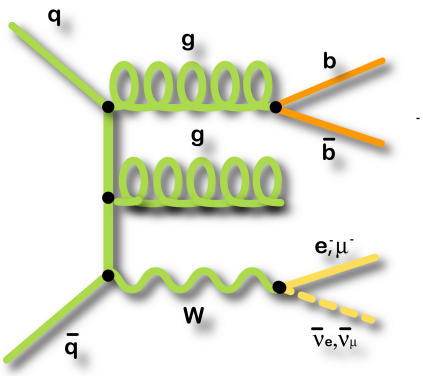


Figure 3.13: $W+b\bar{b}$ event's topology, from [28].

3.8 Backgrounds

3.8.1 The Potential Backgrounds

The following processes mimic the topology of a $t\bar{t}$+jet event as depicted in Figure 3.9 and thus are considered as backgrounds:

- W + heavy flavor: The largest background to our signal comes from events where a W boson is produced in association with heavy flavor jets ($b\bar{b}$, $c\bar{c}$ or c). As depicted in Figure 3.13, the $W+b\bar{b}$ events contain a real W and two b jets as well as a real lepton and neutrino. The `SecVtx` requirement fails to help in this case because the two b-jets, produced from a radiated gluon, are real b-jets. This is in particular the largest background in our analysis. Figure 3.14 shows the other two processes where additional jets come from higher order QCD radiation. However, the background created by the later two processes and by the Z + jet production is less than that from $W+b\bar{b}$ events.

- W + light flavor (Mistag): A substantial contribution also comes from W + jets events with lighter quarks in the final states; called mistags as addressed in section 3.4 . Such events enter our sample when u, d, s or gluon jet is mistagged as a b jet.

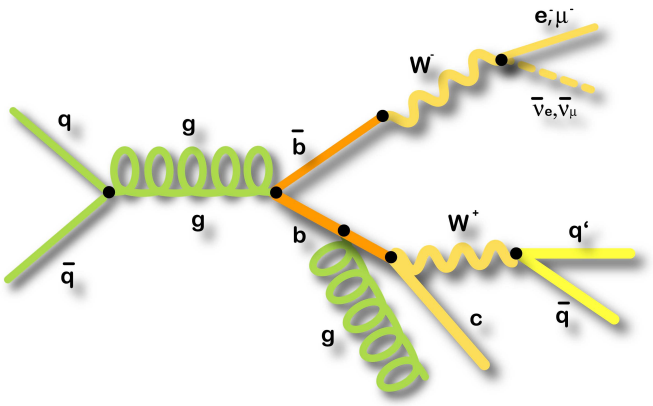


Figure 3.14: $W+c\bar{c}$ event's topology, from [28].

- Electroweak: Diboson electroweak processes such as WW, WZ and ZZ can produce a real lepton and neutrino, as well as a number of jets. Therefore, they form a possible background. $Z \rightarrow ll$ +jets is another background when one of the leptons is missed or misidentified as a jet. However, their contribution is small.

- Single top: As shown in Figure 1.6 , single top events contain a real W boson and a b jet and might enter our sample. Similar to the electroweak backgrounds, single top contributions (s- and t- channels) are small.

- Non-W (QCD): Many pure QCD multijet events, that contain no W bosons, can mimic our signal and fake every feature of its topology. The very large cross section of such processes compared to the electroweak ones makes it a significant background to our sample. QCD background mainly result from gluon multi-jet events: a dijet production (where a jet escapes through a crack the calorimeter, and another fakes an electron, leaving a track in the COT associated to an EM energy deposit) plus additional jets.

Table 3.7 lists all the MC samples used to simulate these backgrounds, as well as the signal. The ttop25 `PYTHIA` sample uses a top mass of 172.5 GeV. This new

Process	Data set ID	Type
$t\bar{t}$	ttop25	PYTHIA
$t\bar{t}$	dtop0s	HERWIG
$t\bar{t}$	dtop1s/dtop2s	PYTHIA IFSR
QCD	bhelX	Fake Electrons for $\not\!\!E_{\rm T}$ fit
QCD	gjt1X	Jet Electrons for Kinematics
$Wb\bar{b}$	btopXw, dtopXw, btopxx	ALPGEN
$Wc\bar{c}$	ctopXw, etopXw, btopxx	ALPGEN
Wc	stopXw	ALPGEN
Wlf	ptopXw, utopXw, utopxx	ALPGEN
$WW/WZ/ZZ$	itopww, itopwz, itopzz	PYTHIA
Single Top (s-channel)	stop00	MadGraph/MadEvent
Single Top (t-channel)	stop0m	MadGraph/MadEvent
$Zb\bar{b}$	ztopbX, btopzX	ALPGEN
$Zc\bar{c}$	ztopcX, btopzX	ALPGEN
Zlf	ztoppX, btopzX	ALPGEN

Table 3.7: Summary of the background and signal models used in our analysis.

sample improved our signal modeling compared to the ttop75 (175 GeV) sample that CDF used in the past. It is worthwhile to mention at this point that because of the extreme complexity of modeling the QCD background, a pure data-based approach is used instead. Specifically, we use an anti-electron data set to model the QCD as described in [118, 119].

The modeling is based on the idea that for a QCD event to pass the selection cuts, it has to have low $\not\!\!E_{\rm T}$ because the strong interaction does not produce neutrinos. The electron has also to pass all the offline selection cuts for electron as listed in Table 3.3 , but to fail the Isolation cut. The rational behind this is that low Isolation (i.e. > 0.1) means more tracks thus fake lepton contained in a jet. This is why these fake electrons are misnamed anti-electrons.

3.8.2 Background Estimate (Method II)

Estimating the background content of processes in the SecVtx tagged lepton plus jet data samples has been implemented at CDF in a program, colloquially referred to as,

`Method II For You`. This method is documented in detail in [120, 121]. The method begins by assuming the composition of the sample; meaning we assume we know of all possible processes contributing to the data. From this assumption we calculate the normalization of processes sequentially, basing the normalization on each type of process on the previous one. Because the MC is not a perfect model of the content of our sample, `Method II` relies heavily on data to fill in the gaps, so to speak, for the parts of the background not well understood.

Following the assumption of `Method II` , the number of events observed in data ($N_{\mathrm{t\bar{t}+jet}}$) should equal the sum of the predicted signal itself and all the backgrounds:

$$N_{\mathrm{t\bar{t}+jet}} = N^{\mathrm{signal}}_{\mathrm{t\bar{t}+jet}} + N_{\mathrm{t\bar{t}+0jet}} + N_{\mathrm{EW}} + N_{\mathrm{Z+jets}} + N_{\mathrm{single\ top}} + N_{\mathrm{QCD}} + N_{\mathrm{W+jets}} \quad (3.4)$$

The first step of `Method II` is to calculate the MC-based estimates of the processes. This includes the electroweak (EW in the equation), Z+jet, single top, $t\bar{t}$+0jet, as well as the predicted signal $t\bar{t}$+jet cross section itself. This is done for both pre-tagged (after applying all the selection cuts except for the `SecVtx` one) and tagged (with the `SecVtx` cut) samples. The QCD fraction is the next step in `Method II`. Because it is by far the least understood and the most poorly predicted, `Method II` fits the $\not\!\!E_{\mathrm{T}}$ spectrum of the QCD template to data (as mentioned in the previous section) to estimate the fraction of our sample that is entirely due to QCD processes. Finally, what is not QCD, EW, Z+jet, or single top is considered W+jet. In the following we will go step by step through the procedure. The results presented here are currently preliminary and documented in detail here [122, 123, 124].

Process	Cross Section [pb]
$t\bar{t}$+0jet	5.5±1.0
$t\bar{t}$+jet	1.5±1.0
Single Top (t-channel)	1.98±0.08
Single Top (s-channel)	0.88±0.05
WW	12.4±0.25
WZ	3.96±0.06
ZZ	1.58±0.02
$Z \rightarrow jet$	787.4±50

Table 3.8: Theoretical cross sections for MC based backgrounds.

MC Based Background Estimate

The calculated event numbers for each process (EW, Z+jet, single top and the predicted signal) in our sample is given by:

$$N_{\mathrm{p\bar{p}} \rightarrow \mathrm{X}} = \sigma_{\mathrm{p\bar{p}} \rightarrow \mathrm{X}} \cdot A \int dt \cdot \mathcal{L} \qquad \text{(pre-tagged sample)} \tag{3.5}$$

$$N_{\mathrm{p\bar{p}} \rightarrow \mathrm{X}} = \sigma_{\mathrm{p\bar{p}} \rightarrow \mathrm{X}} \cdot A \cdot \epsilon \int dt \cdot \mathcal{L} \qquad \text{(tagged sample)} \tag{3.6}$$

where $\sigma_{\mathrm{p\bar{p}} \rightarrow \mathrm{X}}$ is the theoretical cross section, $\int dt \cdot \mathcal{L}$ is the total luminosity, A is the pre-tagged selection acceptance (or efficiency) derived from `PYTHIA` and ϵ is the tagged selection efficiency. The theoretical cross sections used in `Method II` along with their uncertainties are shown in Table 3.8 . The acceptance A is simply the number of events that pass our selection cuts divided by the total number of events generated in MC:

$$A = \frac{N_{\mathrm{selected}}}{N_{\mathrm{generated}}} \tag{3.7}$$

However, this raw acceptance need to be corrected for the lepton scale factor and the trigger efficiency as described in [99].

The tagging selection efficiency is more complicated. Because MC simulations do not model tagging correctly, a tagging scale factor and mistag matrix must be integrated into the tagging efficiency calculation as described in section 3.4. The

Trigger	3 Jets	4 Jets	≥5 Jets
CEM	1.9%	1.9%	0.4%
CMUP	1.2%	1.2%	0.2%
CMX	0.5%	0.5%	0.1%
LOOSE	1.4%	1.4%	0.3%
TOTAL	5.0%	5.0%	1.0%

Table 3.9: The predicted pre-tagged acceptance (A) for $t\bar{t}$+0jet.

Trigger	3 Jets	4 Jets	≥5 Jets
CEM	1.1%	2.0%	2.1%
CMUP	0.7%	1.3%	1.3%
CMX	0.3%	0.5%	0.6%
LOOSE	0.8%	1.5%	1.6%
TOTAL	2.9%	5.3%	5.6%

Table 3.10: The predicted pre-tagged acceptance (A) for $t\bar{t}$+jet.

expected tag yield is calculated by weighting MC events by the probability to be tagged. The probability that an event is tagged is:

$$P_{\text{event}}^{\text{tag}} = 1 - \prod_{i}^{\text{jets}} (1 - p_i^{\text{tag}}) \tag{3.8}$$

For jets matched to heavy flavor, $p_{\text{jet}}^{\text{tag}}$ is the tagging scale factor if tagged and zero if not tagged. If the jet is matched to light flavor, $p_{\text{jet}}^{\text{tag}}$ is the mistag probability. The tagging efficiency is then:

$$\epsilon = \frac{\sum_j^{\text{events}} P_j^{\text{tag}}}{N_{\text{t}\bar{\text{t}}+\text{jet}}^{\text{pre-tagged}}} \tag{3.9}$$

Tables 3.9 through 3.12 show the corrected acceptance and tagging efficiency for the predicted events ($t\bar{t}$+jet) and for $t\bar{t}$+0jet events as well.

QCD Based Background Estimate

To determine the QCD fraction in both the pre-tagged and tagged samples (F_{QCD}), we fit the $\not\!\!E_{\text{T}}$ distribution of a QCD template and a MC signal template to data. The pre-tagged fraction is essential to the rest of `Method II` since it provides the

Trigger	3 Jets	4 Jets	≥5 Jets
CEM	57%	61%	62%
CMUP	57%	61%	63%
CMX	57%	62%	64%
LOOSE	56%	62%	62%
TOTAL	57%	61%	62%

Table 3.11: The predicted tagging efficiency (ϵ) for $t\bar{t}$+0jet.

Trigger	3 Jets	4 Jets	≥5 Jets
CEM	53%	58%	61%
CMUP	53%	58%	61%
CMX	54%	59%	62%
LOOSE	54%	57%	62%
TOTAL	53%	58%	61%

Table 3.12: The predicted tagging efficiency (ϵ) for $t\bar{t}$+jet.

starting point for the heavy flavor fraction and mistag estimate. To perform the fit in the pre-tagged sample, we take the $\not\!\!E_{\rm T}$ distribution of W+jets MC events as the template for signal and for the QCD template we use the anti-electron sample as described previously in [118, 119]. Both templates are fitted to the $\not\!\!E_{\rm T}$ distribution of isolated pretag data events using a binned likelihood fitter to calculate the QCD fraction. Once the fraction is calculated the normalization is:

$$N_{\rm QCD}^{\text{pre-tagged}} = F_{\rm QCD} \cdot N_{\rm t\bar{t}+jet}^{\text{pre-tagged}} \tag{3.10}$$

For the tagged QCD fraction we again use anti-electron sample but this time weighted by the tagging rate (the scale factor and the mistag rate). After calculating the fraction from the fit, the normalization is:

$$N_{\rm QCD}^{\text{tagged}} = F_{\rm QCD} \cdot N_{\rm t\bar{t}+jet}^{\text{tagged}} \tag{3.11}$$

Notice that $N_{\rm t\bar{t}+jet}^{\text{tagged}}$ is 2326, as given in Table 3.6. Tables 3.13 and 3.14 show the QCD

Trigger	3 Jets	4 Jets	≥5 Jets
CEM	25.5%	17.8%	22.7%
CMUP	3.3%	1.2%	4.6%
CMX	3.2%	0.0%	1.0%
LOOSE	7.2%	3.5%	5.0%

Table 3.13: Fitted pre-tagged QCD fraction, F_{QCD} .

Trigger	3 Jets	4 Jets	≥5 Jets
CEM	11.8%	5.5%	6.4%
CMUP	3.0%	0.8%	1.5%
CMX	7.0%	0.0%	0.0%
LOOSE	3.6%	1.6%	0.0%

Table 3.14: Fitted tagged QCD fraction, F_{QCD} .

fraction results for the pre-tagged and tagged samples respectively [9].

W+jet Background Estimate

In the pre-tagged sample, the number of W+jets events is calculated by subtracting the MC based processes and the QCD backgrounds from data:

$$N_{\text{W+jet}}^{\text{pre-tagged}} = N_{\text{t}\bar{\text{t}}\text{+jet}}^{\text{pre-tagged}} - N_{\text{QCD}} - N_{\text{EW}} - N_{\text{Z+jet}} - N_{\text{single top}} \tag{3.12}$$

$$N_{\text{W+jet}}^{\text{pre-tagged}} = N_{\text{t}\bar{\text{t}}\text{+jet}}^{\text{pre-tagged}} \cdot (1 - F_{\text{QCD}}^{\text{pre-tagged}}) - N_{\text{EW}} - N_{\text{Z+jet}} - N_{\text{single top}} \tag{3.13}$$

For the tagged estimate, the W+jet sample is broken into two categories: heavy and light flavor. Each of these processes produces a tagged jet very differently and therefore requires different treatment in calculating the normalization.

The contribution of the heavy flavor background to our signal region is calculated by:

$$N_{\text{W+HF}}^{\text{tagged}} = N_{\text{W+jet}}^{\text{pre-tagged}} \cdot f_{\text{HF}} \cdot K \cdot \epsilon \tag{3.14}$$

This equation reads that the estimated number of W with heavy flavor is a fraction

[9] The fractions should be lower as a function of jet multiplicity. The higher values in the ≥5 jets bins are due to statistical fluctuations.

$f_{\rm HF}$	1 Jet	2 Jets	3 Jets	4 Jets	$\geq$5 Jets
$f_{\rm HF}^{\rm 1b}$	0.76%	1.60%	2.54%	3.40%	3.80%
$f_{\rm HF}^{\rm 2b}$	0.00%	0.90%	1.82%	2.99%	4.16%
$f_{\rm HF}^{\rm 1c}$	5.66%	9.22%	11.72%	13.01%	12.79%
$f_{\rm HF}^{\rm 2c}$	0.00%	1.53%	3.31%	5.53%	7.69%

Table 3.15: The heavy flavor fraction ($f_{\rm HF}$) for W+jet sample.

ϵ	1 Jet	2 Jets	3 Jets	4 Jets	$\geq$5 Jets
$\epsilon_{\rm 1b}$	28.7%	30.7%	33.3%	33.7%	38.2%
$\epsilon_{\rm 2b}$	0.0%	52.8%	52.5%	53.0%	55.8%
$\epsilon_{\rm 1c}$	6.4%	7.6%	8.9%	9.8%	12.6%
$\epsilon_{\rm 2c}$	0.0%	12.7%	14.5%	16.0%	18.0%

Table 3.16: The tagging efficiency (ϵ) for W+jet sample.

of the W+jet sample. The fraction of the sample with jets matched to heavy flavor, $f_{\rm HF}$, is calculated by `ALPGEN` . Therefore, $f_{\rm HF}$ has to be corrected by another scale factor K. The calculation of the K factor is described here [123]. We used a K factor of $(1.5 \pm 0.3\,_{\rm stat+sys})$. ϵ is the tagging efficiency as described in (Equation 3.9). Both $f_{\rm HF}$ and ϵ are calculated for $Wb\bar{b}$, $Wc\bar{c}$ and Wc separately, which defines the rates for each of these processes. The uncorrected $f_{\rm HF}$ and ϵ for our sample are shown in Tables 3.15 and Table 3.16 , where 1 or 2 jets are matched to a b or c quark.

As discussed in section 3.4, the mistag matrix is parametrized by five jet variables. This matrix then gives the probability that a jet with given values of the tag parametrization variables will be negatively tagged. This technique is applied to estimate the number of events in our sample due to mistag. The predicted number of background events from W with light flavor jets is:

$$N_{\rm W+LF}^{\rm tagged} = (N_{\rm W+jet}^{\rm pre\text{-}tagged} - N_{\rm W+HF}^{\rm pre\text{-}tagged}) \cdot \epsilon_{\rm Mistag} \tag{3.15}$$

where the pre-tagged predicted amount of all the background events described so far is subtracted from the total pre-tagged observed events leaving an estimate for the W plus light flavor fraction. This estimate is multiplied by the mistag matrix $\epsilon_{\rm Mistag}$.

Process	1 Jet	2 Jets	3 Jets	4 Jets	≥5 Jets
Pre-tagged	7445	10947	6380	2724	782
Wbb	50.2±15.5	176.2±54.3	128.4±39.8	50.9±16.9	10.2±6.9
Wcc	24.4±7.7	76.9±24.3	65.8±20.8	27.3±9.2	6.0±4.0
Wc	32.6±10.3	75.2±23.7	41.6±13.2	13.1±4.4	2.4±1.6
Mistags	111.4±11.2	181.7±26.8	101.2±18.2	33.2±9.4	6.2±7.4
Non-W	41.6±12.5	116.4±34.9	71.7±21.5	25.5±20.4	9.3±7.5
WW	2.9±0.3	19.0±2.5	14.8±2.0	6.1±0.8	2.0±0.2
WZ	1.0±0.1	7.1±0.8	5.0±0.6	1.9±0.2	0.5±0.1
ZZ	0.1±0.0	0.9±0.1	1.2±0.2	0.5±0.1	0.2±0.0
Z+jets	3.8±0.4	16.3±1.9	16.7±2.1	6.6±0.8	1.8±0.2
Top -S	1.2±0.1	32.6±3.2	16.5±1.6	4.1±0.4	0.8±0.1
Top -T	0.4±0.0	32.9±2.9	18.7±1.6	4.9±0.4	0.9±0.1
$t\bar{t}$+0jet	8.6±1.7	179.3±35.0	534.4±104.2	555.1±108.1	105.7±20.6
$t\bar{t}$+jet	0.5±0.3	16.4±10.3	86.7±54.5	163.1±102.6	182.1±114.5
Predicted	278.6±37.2	930.9±117.3	1102.8±144.6	892.3±157.0	328.2±118.1
Observed	304	917	1115	882	329

Table 3.17: Summary of the background and signal models used in our analysis.

3.8.3 Full Background Prediction

Table 3.17 shows the full background estimate used in our measurement as a function of jet multiplicity. The bins with ≤ 4 jets are exclusive bins, while the ≥ 5 bin is obviously an inclusive sample. A stacked plot illustrating this composition is shown in Figure 3.15 . In the plot, the $t\bar{t}$+jet signal clearly emerges from the background in the ≥ 5 jets bin: $\sim 55\%$ of the events in this bin are attributed to $t\bar{t}$+jet, with $\sim 40\%$ due to $t\bar{t}$+0jet and the rest to backgrounds. It is ~ 1.5 times more likely for the extra hard jet to be radiated off of the ISR/FSR than to be radiated off from the decay products of the top quark and the W boson as discussed before in section 3.7.2 . The requirement of having at least three jets in the events to minimize the background can be justified by looking at the first two bins; where the background dominates the data sample.

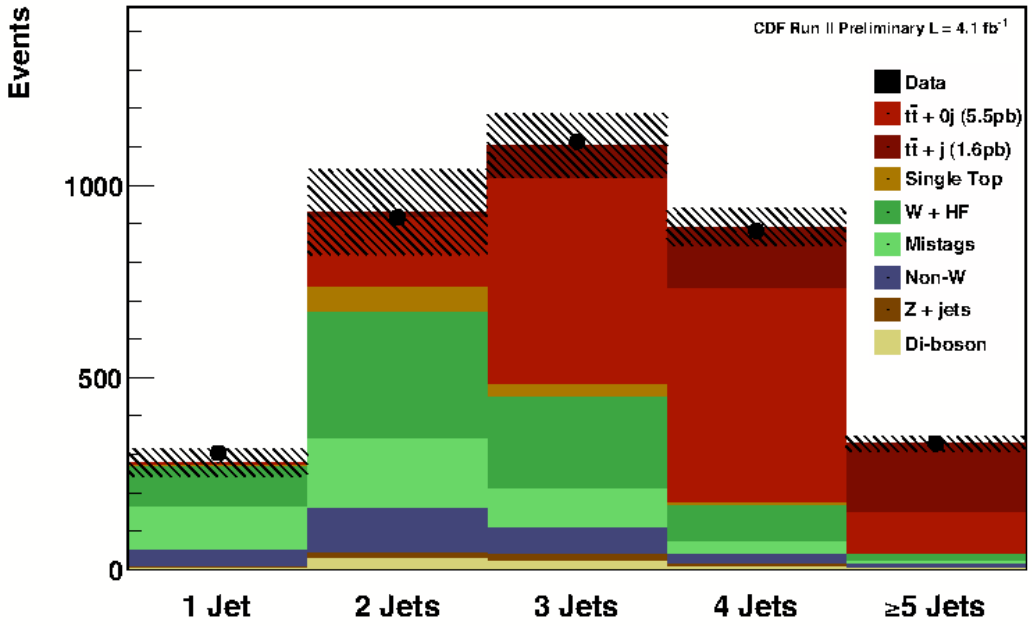


Figure 3.15: The predicted versus the observed composition of $t\bar{t}$ events as a function of jet multiplicity. Normalizations here are the same as in Table 3.17 .

3.9 Validation Plots

The kinematic plots normalized to the `Method II` predictions are shown here to validate that the MC samples do correctly model data and that `Method II` indeed performs adequately. Figures 3.16 through 3.19 show the H_{T}, $\not\!\!E_{\mathrm{T}}$, lepton P_{T} and lead jet [10] E_{T} distributions for the events that pass our selection cuts as described in section 3.6 . Figures 3.20 and 3.21 show the fifth lowest E_{T} jet and η distributions for events satisfying the cuts and having at least five jets.

[10] The jet with the highest E_{T} in an event is referred to as the lead jet in that event.

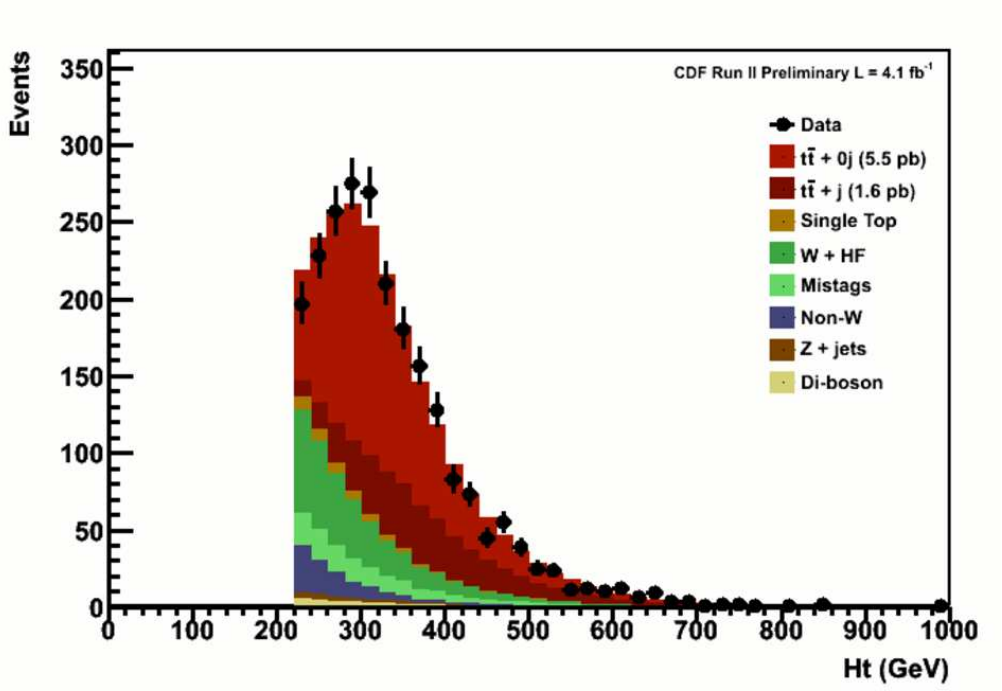


Figure 3.16: H_{T} distribution for the events that pass the selection cuts of section 3.6 .

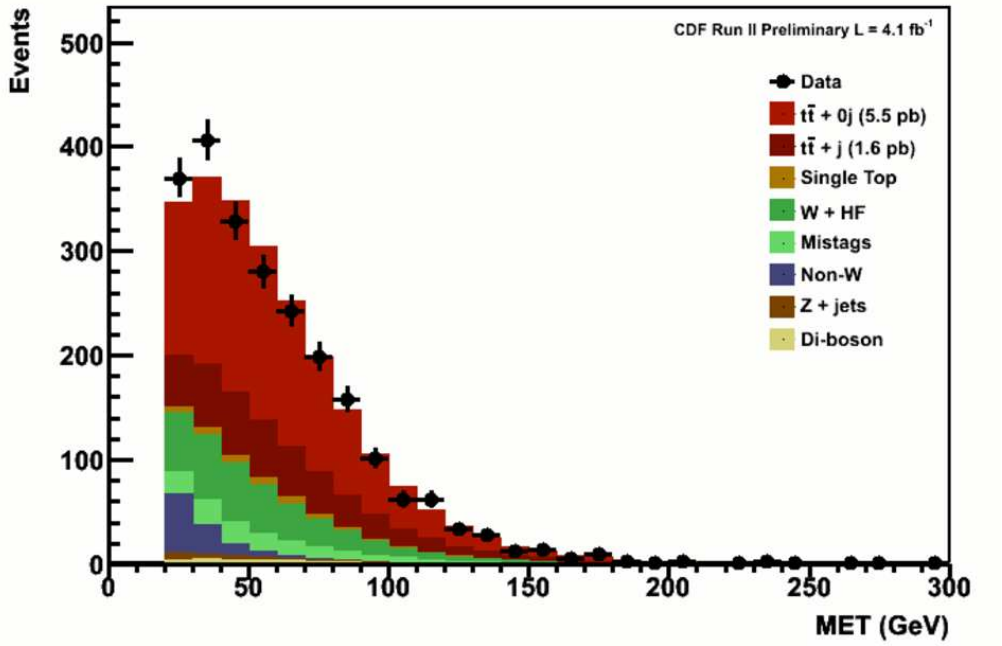


Figure 3.17: The missing E_{T} distribution for the events that pass the selection cuts of section 3.6 .

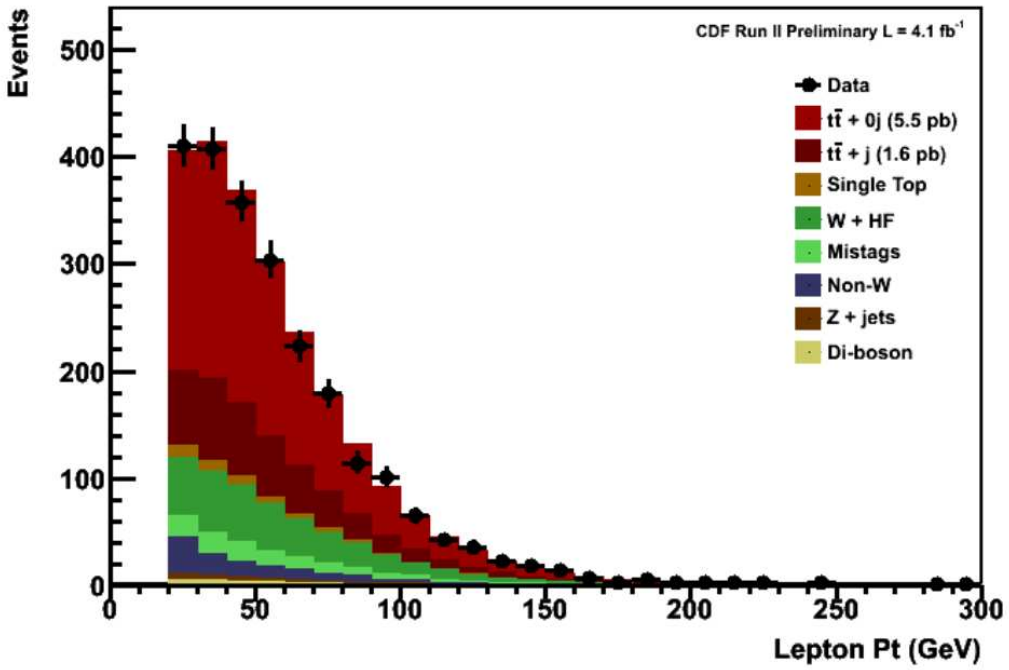


Figure 3.18: Lepton P_{T} distribution for the events that pass the selection cuts of section 3.6 .

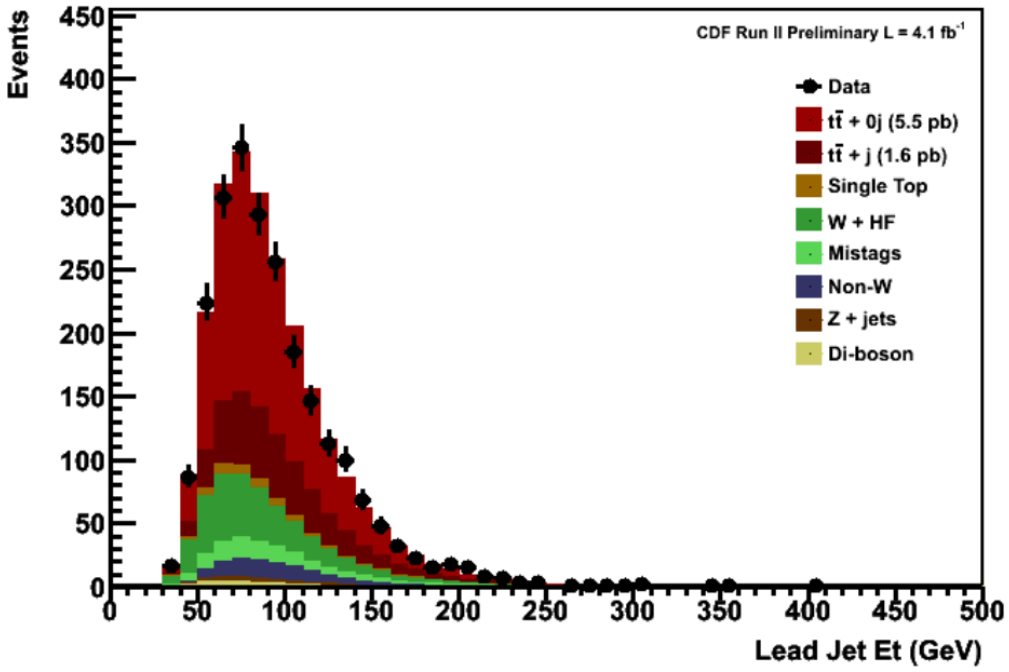


Figure 3.19: Lead jet E_{T} distribution for the events that pass the selection cuts of section 3.6 .

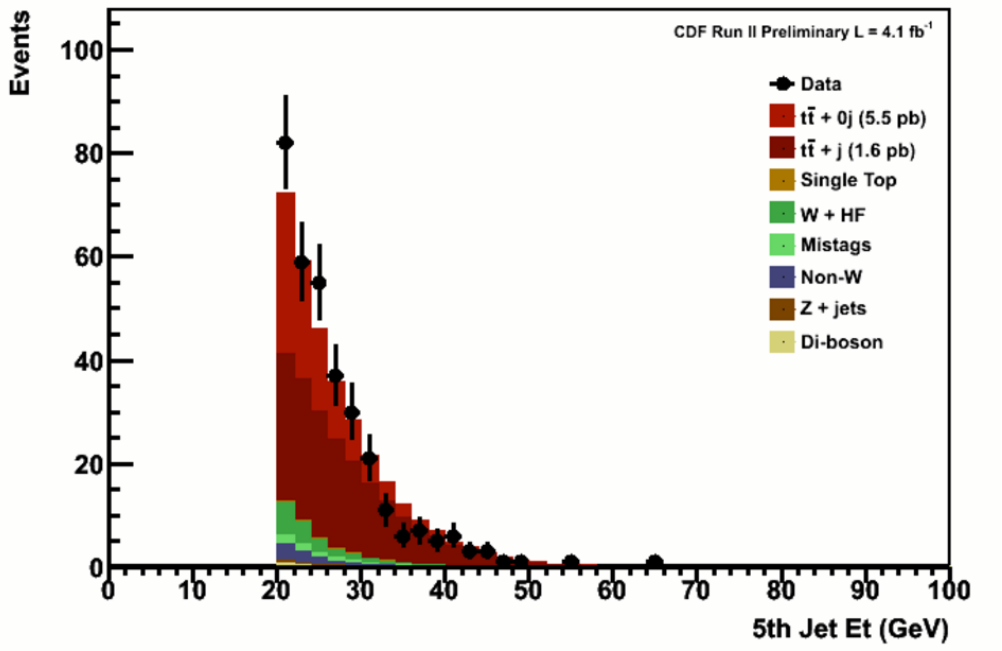


Figure 3.20: The fifth lowest jet E_{T} distribution for the events with at least five jets that pass the other selection cuts of section 3.6 .

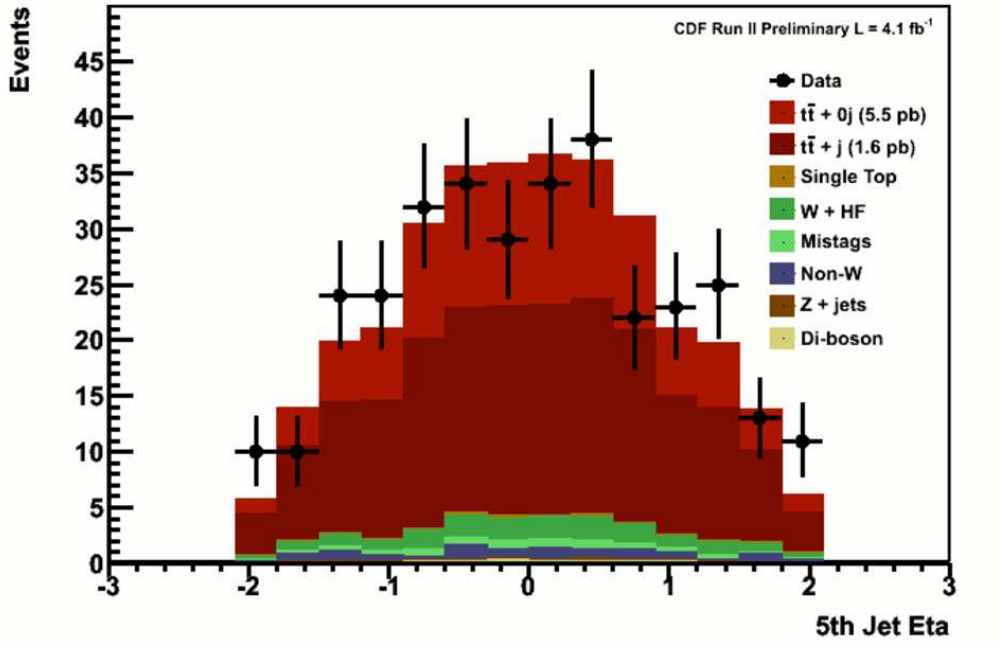


Figure 3.21: The fifth lowest jet η distribution for the events with at least five jets that pass the other selection cuts of section 3.6 .

Chapter 4

Results

There are two possible outcomes: if the result confirms the hypothesis, then you've made a measurement. If the result is contrary to the hypothesis, then you've made a discovery.

Enrico Fermi (1901-1954)

4.1 Introduction

With the background estimate in hand, and after validating the efficiency of `Method II`, we now perform the measurement of the $t\bar{t}$+jet cross section (section 4.2) and study the different kinematics of the extra hard jet (section 4.3). The book will conclude by a brief conclusion in section 4.4 .

4.2 The Counting Analysis

4.2.1 Calculating The Cross Section

We saw how `Method II` is dependent on the measured cross section itself, and hence extracting the latter value is not simple (i.e. we can't use Equation (1.53) directly).

Instead, we construct a Poisson likelihood where we take into account the background dependence. In order to extract simultaneously the measured values of the $t\bar{t}$+jet and $t\bar{t}$+0jet cross section, we construct a 2D likelihood from the data and the prediction for events with three, four, or five jets:

$$P_i = \frac{\lambda_i^{k_i} \cdot e^{-\lambda_i}}{k_i!} \tag{4.1}$$

where k is the number of events in data with (i) jets, and λ is the predicted number of events with (i) jets. More specifically:

$$\lambda = A_{0\mathrm{j}} \cdot \epsilon_{0\mathrm{j}} \cdot \mathcal{L} \cdot \sigma_{0\mathrm{j}}^{\mathrm{t\bar{t}}} + A_{+\mathrm{j}} \cdot \epsilon_{+\mathrm{j}} \cdot \mathcal{L} \cdot \sigma_{+\mathrm{j}}^{\mathrm{t\bar{t}}} + \text{Background } (\sigma_{0\mathrm{j}}^{\mathrm{t\bar{t}}}\ ,\ \sigma_{+\mathrm{j}}^{\mathrm{t\bar{t}}}) \tag{4.2}$$

where A is the acceptance, ϵ is the tagging efficiency, and $\mathcal{L}$ is the luminosity. The likelihood is then [126]:

$$L = -\,ln\ (P_3 \cdot P_4 \cdot P_5) \tag{4.3}$$

The likelihood is calculated for several values of the cross section and the resulting points are fit to a 2D second order polynomial. The minimum point on the fitted parabolic is taken as the measured cross section values while the uncertainty is extracted from the tangents to the ellipse at the point where the likelihood is one-half unit, 0.5, above the minimum. For the joint probability curves, we choose an ellipse formed at 1.15 units (68% CL) and 3.0 units (95% CL) above the minimum point. The 2D likelihood curve is shown in Figure 4.1 [1]. The measured values with statistical uncertainty are:

$$\sigma_{\mathrm{t\bar{t}+jet}} = 1.6 \pm 0.2_{\mathrm{stat}}\ \mathrm{pb} \tag{4.4}$$

$$\sigma_{\mathrm{t\bar{t}+0jet}} = 5.5 \pm 0.4_{\mathrm{stat}}\ \mathrm{pb} \tag{4.5}$$

[1] The peak value of this curve lies at the most probable value of the cross section given the data, and the uncertainty is defined as the region around this value that contains 68% and 95% of the total area of the probability density curve.

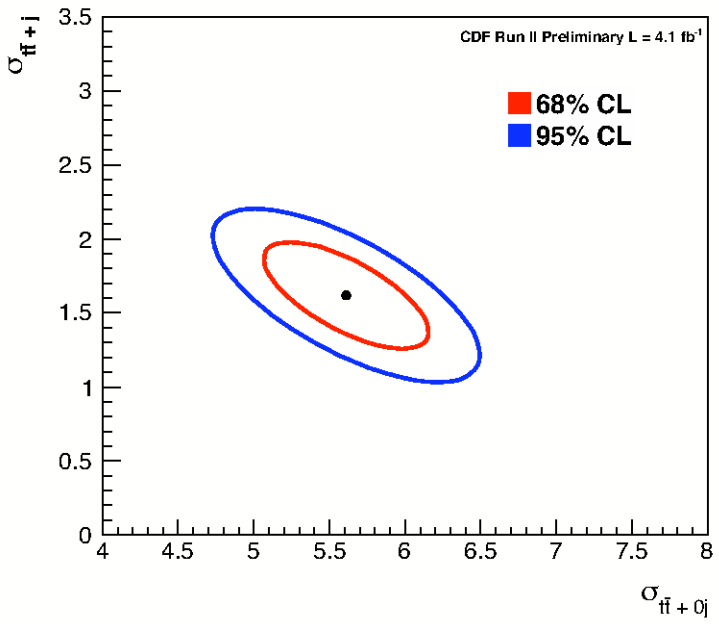


Figure 4.1: Likelihood curve for measured $t\bar{t}$+jet and $t\bar{t}$+0jet cross section.

4.2.2 Systematics

The systematic uncertainties in our measured results are calculated by varying a given parameter within its uncertainty and redoing the entire measurement. Each systematic is described below along with any additional relevant quantities. The individual systematic uncertainties are shown in Figures (Tables) 4.2 and 4.3 at the end of the section.

- JES : We study the effect of the JES (as described in section 3.3) on the measurement by varying it for our signal MC and background models and then re-performing the measurement. The effect of the JES on this measurement is mainly through the acceptance of signal and background. For the $t\bar{t}$+jet sample, the JES is by far the largest systematic uncertainty, on the order of the statistical error. It is not hard to understand why, given how drastically its

variation changes the acceptance for the $t\bar{t}$+jet events. The E_{T} spectrum for the fifth jet is quite soft, very close to the 20 GeV cut. Since the majority of the $t\bar{t}$+jet events sit in the fifth jet bin (as shown in Figure 3.15), any change in the JES shifts the number of the events away from or into the fifth jet bin. A lower H_{T} cut (such as the standard cut used by the CDF Top Physics Group) will decrease the dependence on the JES, but to avoid too great dependence on the QCD and W+jet modeling (mostly for the three jet bin), we raise the H_{T} cut to 220 GeV.

- Tagging : The scale factor applied to each b or c jet is 0.95±0.05 as mentioned in section 3.3 . The uncertainty associated with the scale leads to a systematic on the measurement. The effect on the measured value is calculated by fluctuating the scale factor within its uncertainty, applying it to each jet and repeating the entire measurement. The uncertainty for both the b and c scales propagated to the $t\bar{t}$+jet cross section is $\sim$ 5% (as shown in Figure 4.3), much less than the uncertainty of the JES.

- Mistag : The uncertainty on the probability for a jet to be mis-tagged, as calculated by the mistag matrix, is on the order of 20% (see section 3.3). This is taken as a systematic uncertainty on the measurement. The mistag rate on any jet is then fluctuated up and down by 20% and the entire measurement is repeated to quantify the effect.

- K factor: For modeling the W+jet sample, the heavy flavor fraction was corrected by the K factor as described in section 3.8.2 (page 111). We used a K factor of $(1.5 \pm 0.3\ _{\mathrm{stat+sys}})$ for this measurement.

- Luminosity : The current estimation on the uncertainty of the luminosity measured by CLC is around 5.8% as discussed in section 2.4.4 . The luminosity used

in the measurement is fluctuated within this uncertainty and the measurement redone. This uncertainty is the second largest systematic in our measurement.

- QCD : As addressed in detail in section 3.8.1, a pure data-based set is used to estimate the QCD fraction instead of a MC model. In section 3.8.2 we used the anti-electron data set for the QCD background estimate. To estimate the uncertainty on this background, fits are performed with a different data set; the jet electrons as listed in Table 3.7 . The fits are based on the idea that for a QCD event to pass the selection cuts, a jet has to fake an electron. We require the event to pass all the offline selection cuts for electron as listed in Table 3.3 , but to fail two of the five non-kinematic cuts : Δx, Δz, L_{shr}, $E_{\rm HAD}/E_{\rm EM}$ and $\chi^2_{\rm strip}$. The events that fail these cuts should be QCD-enriched, since the non-kinematic cuts serve primarily to filter out non-W events (i.e. to reject fake electrons), but still have most of the kinematic properties of real W events, including isolation. This is why these fake electrons are named jet electrons. The resulting difference in the fits is 30% which is taken as a systematic uncertainty in the measurement.

- ISR/FSR : An interesting aspect of this measurement is that the result of Equation (4.4) is a measure of the ISR (see section 1.3.2 page 10 and section 3.7.2). Because of this, we do not include this effect as a systematic, rather we cross check that our result is invariant to using a MC sample where ISR/FSR have been increased (`PYTHIA` IFSR in Table 3.7). Unfortunately, the FSR effect, which would be a true systematic is present in this cross check, but we assume the effect is quite small based on previous cross section measurements.

- MC Generators : Differences in MC models for parton showering (section 3.7.1) are studied simply by replacing our `PYTHIA` $t\bar{t}$ sample with the `HERWIG` sample as listed in Table 3.7 . `HERWIG` is separated into $t\bar{t}$+0jet and $t\bar{t}$+jet events exactly as `PYTHIA` is and the measurement was repeated.

SYSTEMATIC	$\Delta\ \sigma_{0j}$ pb	$\Delta\ \sigma_{0j}/\sigma_{0j}$ %
JET ENERGY SCALE	0.3	4.9
BOTTOM TAGGING	0.3	4.6
CHARM TAGGING	0.0	0.2
MIS-TAGS	0.0	0.2
HEAVY FLAVOR CORRECTION	0.4	6.7
LUMINOSITY	0.3	5.6
QCD FRACTION	0.0	0.2
ISR/FSR	0.1	2.1
MC GENERATOR	0.2	3.5
TRIGGER EFFICIENCY	0.0	0.6
PDF	0.1	1.0
TOTAL	0.7	11.8

CDF Run II Preliminary Lumi = 4.1 fb^{-1}

Figure 4.2: Systematic uncertainites for the measured $t\bar{t}$+0jet cross section. The numbers in the ($\Delta\ \sigma$ pb) column are rounded but the percentile column represents the exact values.

- Lepton ID : Detector specific corrections are applied to the MC to more correctly model the relative trigger efficiencies between CEM, CMUP, and CMX events, as discussed in detail in section 3.2 . The corrections are data-derived from a pure sample of Z bosons and have a small uncertainty associated with them. There are two types of corrections, lepton ID and trigger efficiencies. Each are fluctuated within their uncertainty, separately, and the resulting errors are added in quadrature.

- PDF : Uncertainty in the PDFs (section 1.3.2 page 9) are evaluated by a re-weighting scheme. PDFs are re-weighted in our signal MC to simulate 46 different PDF parametrizations. The measurement is performed for each different parametrization. A prescription for evaluating the uncertainty derived from the result of this is documented in detail here [127].

SYSTEMATIC	$\Delta\ \sigma_{+j}$ pb	$\Delta\ \sigma_{+j}/\sigma_{+j}$ %
JET ENERGY SCALE	0.5	30.2
BOTTOM TAGGING	0.1	4.6
CHARM TAGGING	0.0	0.4
MIS-TAGS	0.0	0.6
HEAVY FLAVOR CORRECTION	0.1	3.4
LUMINOSITY	0.1	6.1
QCD FRACTION	0.0	0.4
ISR/FSR	0.1	3.3
MC GENERATOR	0.0	2.3
TRIGGER EFFICIENCY	0.0	0.6
PDF	0.0	1.0
TOTAL	0.5	31.6

CDF Run II Preliminary Lumi = 4.1 fb^{-1}

Figure 4.3: Systematic uncertainites for the measured $t\bar{t}$+jet cross section. The numbers in the ($\Delta\ \sigma$ pb) column are rounded but the percentile column represents the exact values.

4.2.3 The Measured $\sigma_{\mathbf{t\bar{t}+jet}}$ & $\sigma_{\mathbf{t\bar{t}+0jet}}$

The first measured cross section of $t\bar{t}$ events in association with an extra jet with $E_T \geq 20$ GeV and $|\eta| \leq 2$ is:

$$\sigma_{t\bar{t}+\text{jet}} = 1.6 \pm 0.2_{\text{stat}} \pm 0.5_{\text{syst}}\ \text{pb} \tag{4.6}$$

while the measured cross section for $t\bar{t}$ events without an extra jet is:

$$\sigma_{t\bar{t}+0\text{jet}} = 5.5 \pm 0.4_{\text{stat}} \pm 0.7_{\text{syst}}\ \text{pb} \tag{4.7}$$

which when combined together gives the inclusive $t\bar{t}$ cross section:

$$\sigma_{t\bar{t}} = 7.1 \pm 0.3_{\text{stat}}\ \text{pb} \tag{4.8}$$

The result of Equation 4.6 is in agreement with the NLO prediction [79, 80] $\sigma_{t\bar{t}+\text{jet}} = 1.791(1)^{+0.16}_{-0.31}$ as listed in Table 1.4 . On the other hand, the value of the measured inclusive cross section (Equation 4.8) is also in agreement with the predicted SM value as listed in Tables 1.2 and 1.3 .

4.3 Kinematic Analysis

So far we have studied the total production rate for the extra hard jet associated with the $t\bar{t}$ events. However, this result does not provide us with complete kinematic information about the extra jet. Therefore, and for the completeness of our analysis, we want to examine the kinematic profile of the extra jet to see if it agrees with expectations from the theoretical predictions. In order to do so, we have to develop an algorithm (a simple handle) that isolates correctly the extra jet from the other final state jets. In the following subsections, I'll introduce the analysis by studying the extra jet as defined in section 3.7.2 using the `PYTHIA` ttop25 MC sample. With the selection cuts we used in the cross section analysis (as listed in section 3.6), I will proceed by developing the isolation algorithm and verifying its efficiency, along with correcting the measured extra jet E_{T} distribution for the purity/efficiency. The question of whether the extra jet is the softest (least energetic) jet among the jets in the final state will be described afterword, followed by cross checking our results using a different MC model (`MCFM`).

4.3.1 The Truth Extra Jet

In order to be consistent with the counting analysis, we will apply the same selection cuts on the MC sample. We will first study the case where we have ≥ 2 `SecVtx` tagged jets in the events (We will, from now on, refer to these as double tag events) and compare to the case where we have only ≥ 1 `SecVtx` tagged jet (single tag

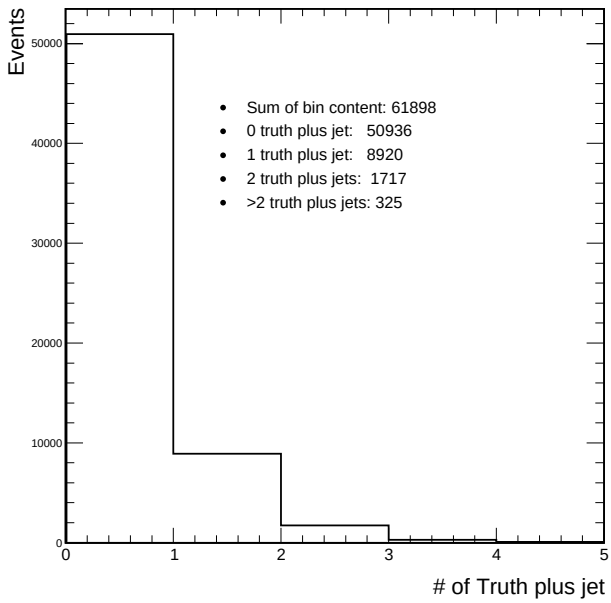


Figure 4.4: The truth plus jet multiplicity.

events). For the double tag events, Figure 4.4 shows the extra hard jet multiplicity in the final state as defined and clustered by the tools described in section 3.7.2. This is why we will term the hard extra jet the "truth plus" jet to distinguish it from the L5 corrected jets of the MC sample, the ttop25 sample [2]. The histogram demonstrates that having more than one truth jet is highly suppressed, as expected in QCD. The MC predictions indicate also a good agreement with the measured cross section of $t\bar{t}$+jet; the ratio of the $t\bar{t}$+jets events to the total inclusive $t\bar{t}$ sample ($\sim$ 11k/62k) is close to the ratio of the exclusive cross section (Equation 4.6) to the inclusive cross section (Equation 4.7) ($\sim$ 1.6 pb/7.1 pb). This reflects that the MC sample that we use provides a reasonable description of the data.

As we are interested in $t\bar{t}$ events with at least three jets after the $t\bar{t}$ pair decays, we investigated how often $t\bar{t}$ events with 4 ttop25 jets would have an extra jet which is a truth plus jet (Figure 4.5), and how often $t\bar{t}$ events with 5 ttop25 jets would have

[2] The jets in the ttop25 MC sample are L5 corrected while the truth plus jet is not corrected because, by definition, it is a theoretical jet.

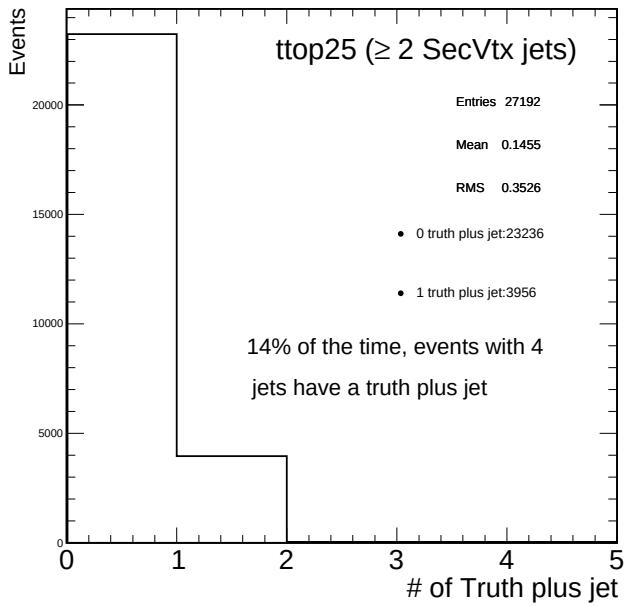


Figure 4.5: The number of truth plus jets in 4 ttop25 jet events.

an extra jet which is a truth plus jet (Figure 4.6). These plots verify that most of the time, events with only 4 jets would not have a truth plus jet, but more than half of the time events with 5 jets would have a truth plus jet. This agrees with the situation observed in single tag events as displayed in Figure 3.15 in the ≥ 5 jets bin. This, however, doesn't necessarily mean that the fifth ttop25 jet is actually the truth plus jet. This will be addressed in the next section.

4.3.2 The Isolation Algorithm

Double Tag Events

Based on the results from the previous plots, we use the events which have 5 ttop25 jets where the truth plus jet is found, i.e. the 4991 events of Figure 4.6 . These events have exactly five jets in the final state as depicted, for example, in Figure 3.9 . One jet is for sure the extra jet. The main difficulty is how to isolate or extract this extra jet from among the 5 jets.

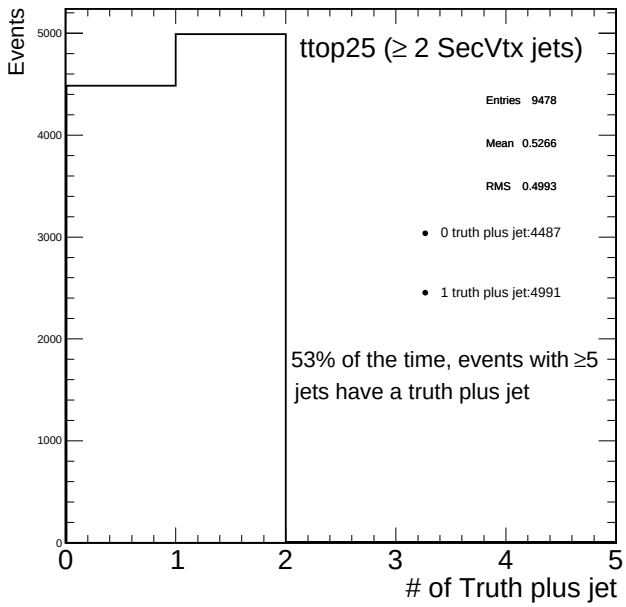


Figure 4.6: The number of truth plus jets in 5 ttop25 jet events.

Requiring two jets to be `SecVtx` jets would leave us with three jets, one of which should be the extra jet. Out of these three jets, two jets should reconstruct the hadronic W boson. There are three possible dijet combinations to reconstruct the mass of the hadronic W boson, among which we will choose the one that is closest to the correct mass. The other two combinations are considered as a background to the signal. The jet that remains should in turn represent the extra jet. We can further reconstruct the hadronic top quark by forming the 3-jet mass using the best dijet combination along with one of the `SecVtx` jets.

Figure 4.7 shows all the possible combinatoric mass distributions to reconstruct the hadronic W boson and the hadronic top quark masses [3], while Figure 4.8 shows only the best mass distribution. It is clear how the former distribution fills a much larger phase space than the best distribution. Figures 4.9 and 4.10 show separately the best reconstructed mass distributions of the hadronic W boson and the hadronic

[3] We have two possible combinatorics to reconstruct the top quark mass for each of the three possible combinatorics to reconstruct the W boson mass, total of six combinatorics.

top quark respectively. Figures 4.11 and 4.12 show ΔR separation (see Equation 2.3) between the dijet that best reconstructs the hadronic W boson [4] and the truth W boson, and ΔR separation between the trijet that best reconstructs the hadronic top quark [5] and the truth top quark respectively. Figure 4.13 indicates that by using this algorithm, and after calculating the ΔR separation between the candidate jet that is not part of the W boson and the truth plus jet, we are able to isolate the correct extra jet $\sim 52\%$ of the time. By requiring the dijet combination that best reconstructs the mass of the hadronic W boson to be within $\pm$ 10 GeV of the truth W boson mass, our efficiency of correctly isolating the extra jet increases to $\sim 57\%$ without a significant reduction in statistics. This is displayed in Figure 4.14 . With this $\pm$ 10 GeV constraint on the W boson mass, and by applying a $\pm$ 20 GeV constraint on the hadronic reconstructed top quark mass, the efficiency increases to $\sim 61\%$. This is presented in Figure 4.15 .

In data, we of course don't know which jet the extra jet is. However, we can still verify the ability of the algorithm to reconstruct the measured hadronic W boson and the measured hadronic top quark masses in the 5 jets bin events. Figure 4.16 shows that the mean values of the measured distribution are close to those of the ttop25 MC distribution of Figure 4.8 .

[4] i.e. the dijets of Figure 4.9 .
[5] i.e. the trijets of Figure 4.10 .

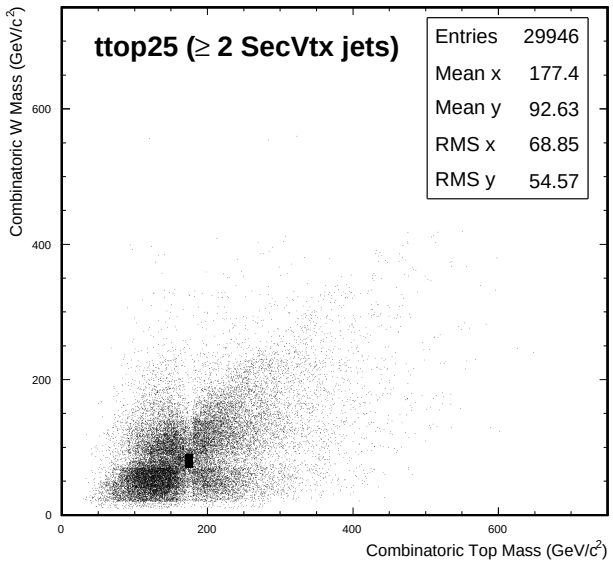


Figure 4.7: All the possible combinatoric W boson and top quark mass distributions using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

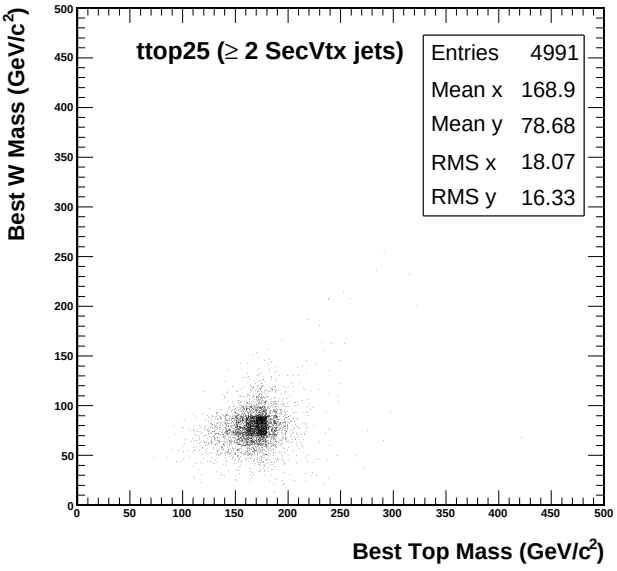


Figure 4.8: The best reconstructed W boson and top quark mass distributions using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

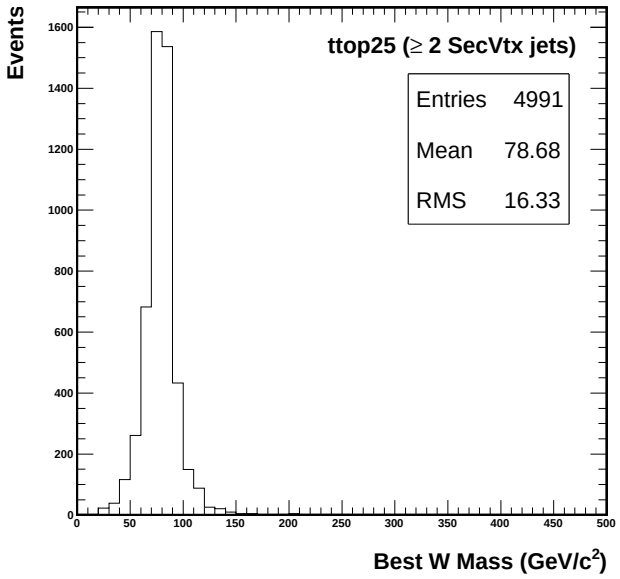


Figure 4.9: The best reconstructed W boson mass distribution using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

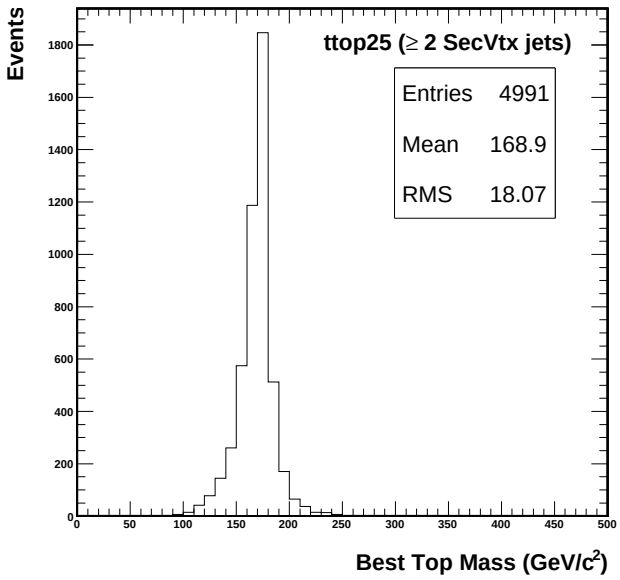


Figure 4.10: The best reconstructed top quark mass distribution using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

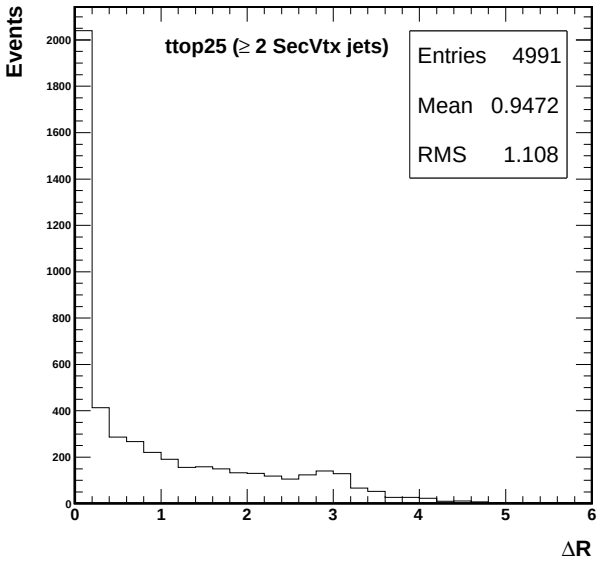


Figure 4.11: ΔR value between the dijet combination that best reconstructs the W boson and the truth W boson mass distributions using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

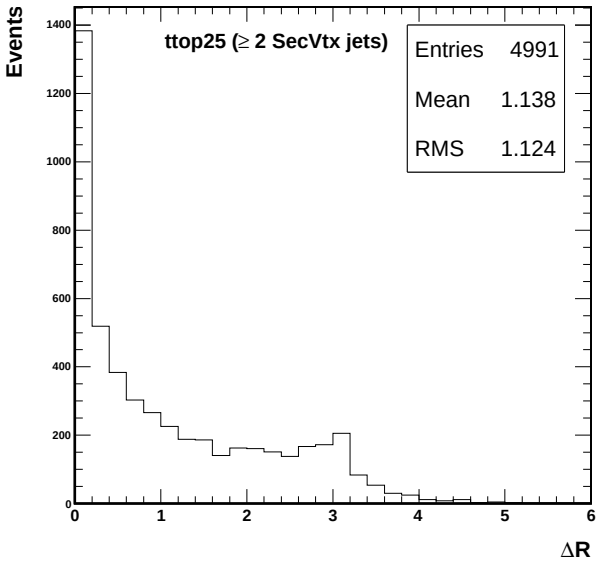


Figure 4.12: ΔR value between the trijet combination that best reconstructs the top quark and the truth top quark mass distributions using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

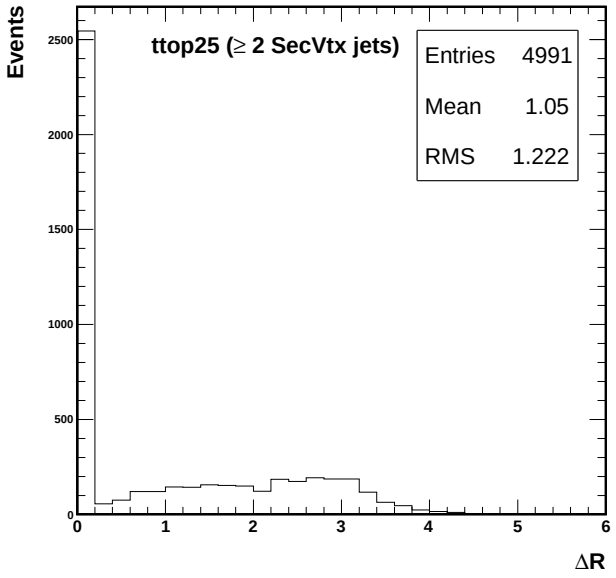


Figure 4.13: ΔR value between the extra ttop25 jet and the truth plus jet using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

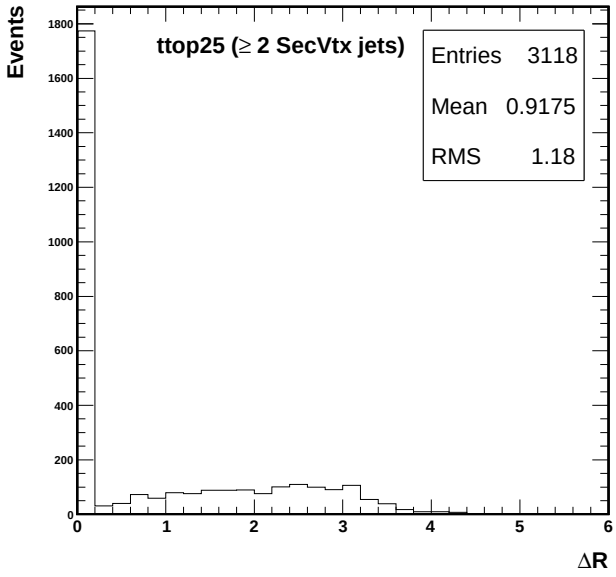


Figure 4.14: ΔR value between the extra ttop25 jet and the truth plus jet within a $\pm$ 10 GeV window of the truth W boson mass distribution using the ttop25 MC sample with ≥ 2 `SecVtx` tagged jets.

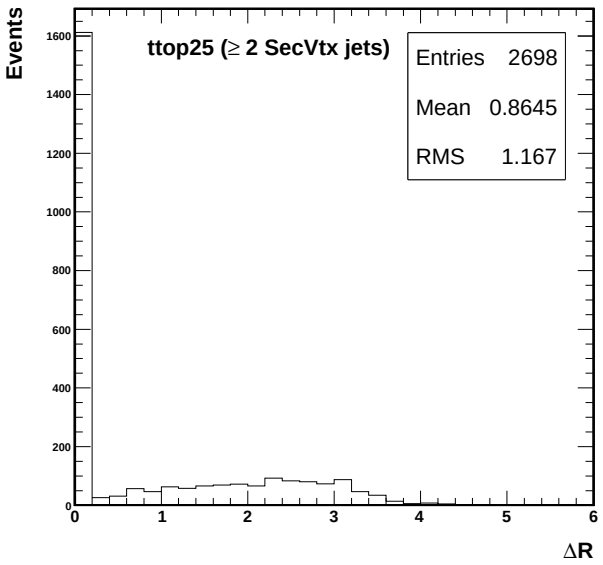


Figure 4.15: ΔR value between the extra ttop25 jet and the truth plus jet within a box of $\pm$ 10 and $\pm$ 20 GeV of the truth W boson and the truth top quark masses respectively, using the ttop25 MC sample with $\geq$ 2 `SecVtx` tagged jets.

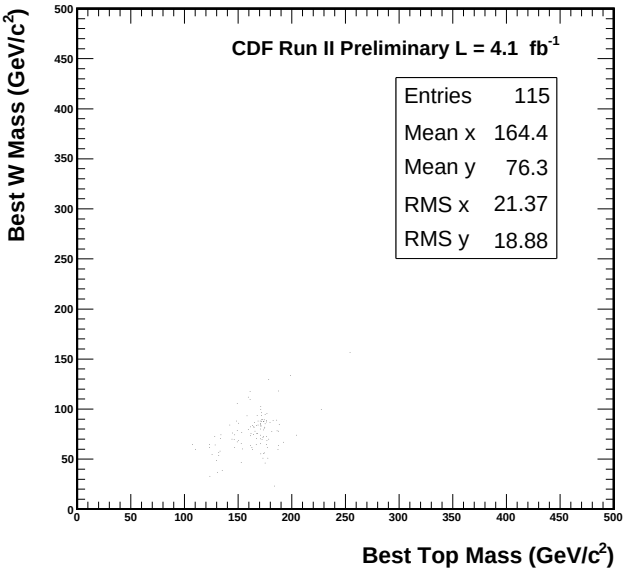


Figure 4.16: The best measured W boson and top quark mass distributions with $\geq$ 2 `SecVtx` tagged jets.

Single Tag Events

For the double tag case, in order to isolate the extra jet, we looked at three possible combinations to reconstruct the hadronic W boson. We want to study the impact of requiring only one `SecVtx` tagged jet in our sample since the majority of our sample has only one `SecVtx` tagged jet. The number of possible combinations will increase and this will have an impact on determining the extra jet. In single tag events, there are four non b-tagged jets, two of which should reconstruct the hadronic W boson. Therefore, we have six possible combinations for the reconstruction of the hadronic W boson; one combination is the signal and the other 5 are considered backgrounds. By choosing the combination that best reconstructs the W boson mass, two jets remain where, one of them should be the extra jet. We then calculate the mass of this best W boson combination with each of these two remaining jets and choose the trijet combination that best reconstructs the correct mass of the top quark. The jet selected in the best trijet combination should represent the other `SecVtx` jet that has to be present, by definition, in the final state, while the other jet should in turn represent the extra jet.

Figure 4.17 shows the dijet combination that best reconstructs the hadronic W boson mass and the trijet combination that best reconstructs the top quark mass. It is clear that the latter is much broader (more background) than that shown in Figure 4.8 for the double tag events. This is expected, because in the double tag case, one can always reconstruct the hadronic top quark because one of the `SecVtx` jets would definitely come from the hadronic top decay, while in the single tag events, the `SecVtx` jet could equally likely come from the leptonic top decay as well as from the hadronic decay, and the former can not be directly reconstructed. Figure 4.18 verifies that by using this algorithm and after calculating ΔR separation between the jet that is left over and the truth plus jet, we are able to isolate the correct extra jet $\sim 50\%$ of the time. With a box of $\pm$ 10 and $\pm$ 20 GeV around the truth W boson and the truth

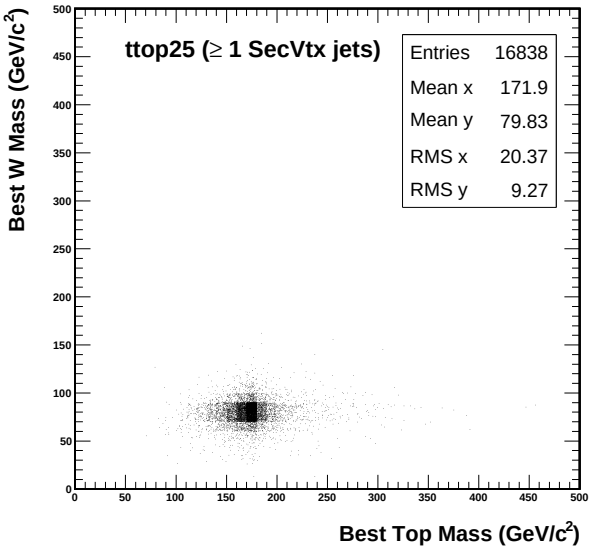


Figure 4.17: The best hadronic W boson and top quark mass distributions using the ttop25 MC sample with ≥ 1 `SecVtx` tagged jets.

top quark masses respectively, the efficiency (Figure 4.19) increases to $\sim 59\%$. This is slightly less than the analogous value for the double tag case as shown in Figure 4.15 .

4.3.3 The Corrected Measured Extra Jet

As we are able to isolate the correct extra jet more than 60% of the time, we can rely on our isolation algorithm to correct back our observed extra jet distribution to the truth plus jet distribution. Therefore, we use all the events in the 5 jets bin for data distribution, while for the ttop25 MC distribution, we use only the events in the 5 jets bin where we find the extra jet (not all the 5 jets bin MC events).

For the double tag events, we can assume that there are no non-top backgrounds (i.e. `Method II` backgrounds) and we have to care only about the combinatoric backgrounds as the ones in Figure 4.7. Figure 4.20 shows the E_{T} distribution of the extra jet as found in data, in the ttop25 MC distribution and in the truth distribution

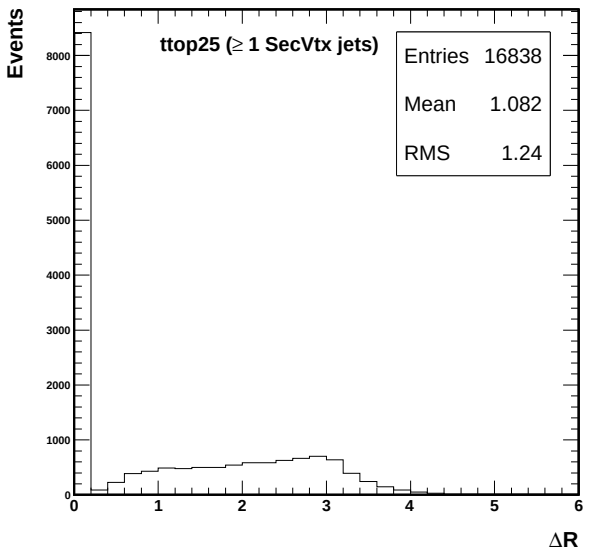


Figure 4.18: ΔR value between the extra ttop25 jet and the truth plus jet using the ttop25 MC sample with ≥ 1 `SecVtx` tagged jets.

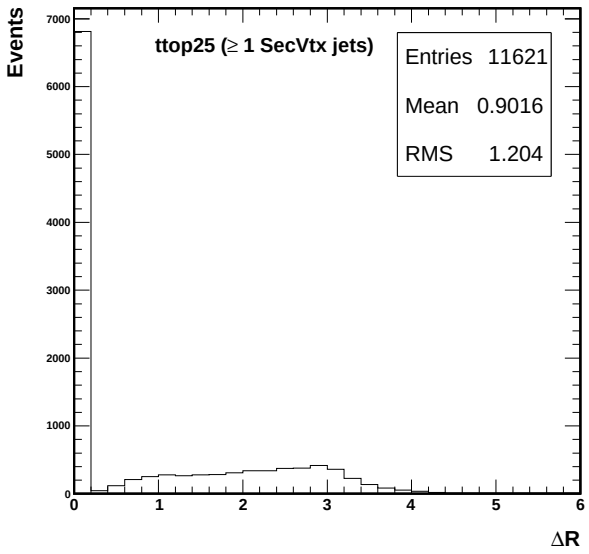


Figure 4.19: ΔR value between the extra ttop25 jet and the truth plus jet within a box of $\pm$ 10 and $\pm$ 20 GeV of the truth W boson and the truth Top quark mass distributions respectively, using the ttop25 MC sample with ≥ 1 `SecVtx` tagged jets.

(section 4.3.1) using the mass box constraint. By determining bin by bin the efficiency for keeping the correct extra jet, we correct the observed E_{T} distribution of the extra jet for the efficiency and compare it to the truth plus jet E_{T} distribution. This is shown in Figure 4.21 .

On the other hand, and for the single tag events, we first subtract the `Method II` backgrounds (as listed in Table 3.17) from the data, leaving only the signal and the combinatoric backgrounds. Then we correct the observed E_{T} distribution for the efficiency (within the mass box constraint) and compare it to the truth plus jet distribution. This is what Figure 4.22 displays. Figure 4.23 shows the extra jet E_{T} distribution as found in data (before being corrected), in the ttop25 MC distribution, in the truth distribution and in the `Method II` background distributions. It is obvious how the `Method II` background contributions are relatively small. This is why we ignore this kind of background for the more pure double tag event samples as shown in Figure 4.20 .

However, the corrected measured E_{T} distributions (Figure 4.21 and Figure 4.22) are slightly steeper (peak more at low E_{T} value) than the truth E_{T} distributions, especially for the double tag event samples. On the other hand, the theoretical predictions (the ttop25 MC distribution and the truth distribution) behave similarly as shown in Figure 4.20 and Figure 4.23 .

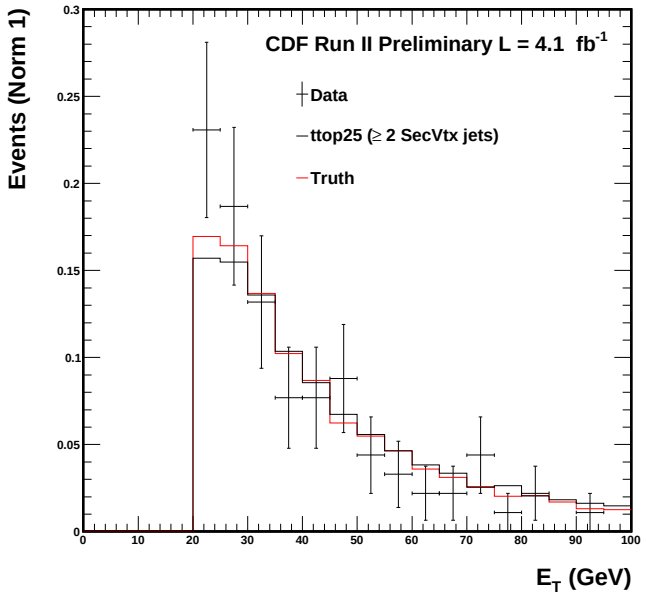


Figure 4.20: The E_{T} distribution of the extra jet as found in data, ttop25 and the truth samples for the double tag event samples.

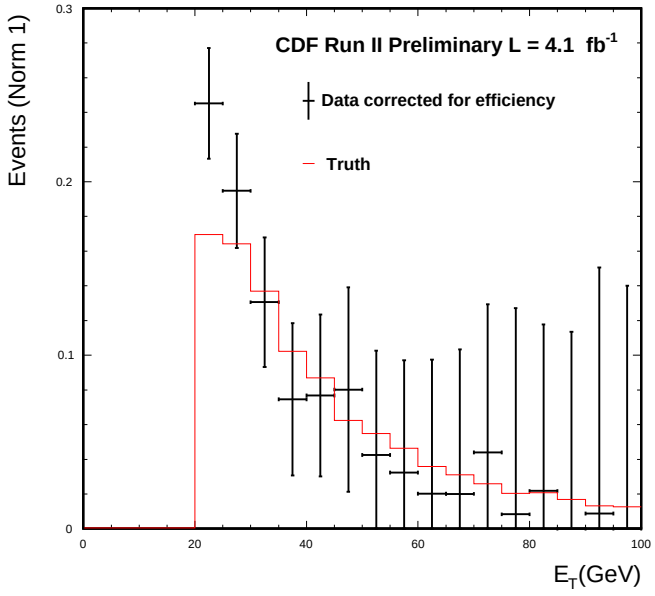


Figure 4.21: The corrected observed E_{T} distribution of the extra jet compared with the truth distribution for the double tag event samples.

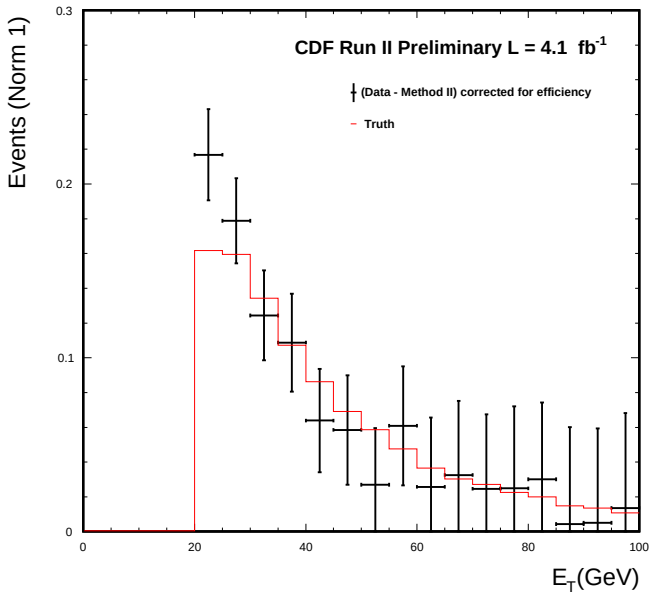


Figure 4.22: The corrected observed E_{T} distribution of the extra jet compared with the truth distribution for the single tag event samples.

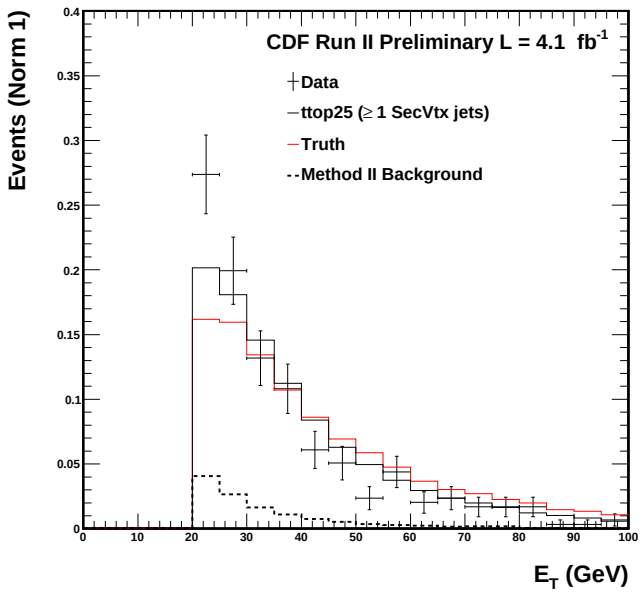


Figure 4.23: The E_{T} distribution of the extra jet as found in data, ttop25, truth and `Method II` background samples for the single tag event samples.

4.3.4 The Extra Jet Profile

Full Machinary ΔR Calculations

So far, we discussed how to correctly determine the extra jet and it is time now to determine which jet, among the five jets of the final state, the extra jet is. The two b jets tend to be the most energetic jets in the event and we have verified this using ttop25 MC events. The two light jets radiated from the hadronic W tend also to be more energetic than the extra jet, especially if the latter belongs to the $t\bar{t}$+jet sample (i.e. ISR) not to the $t\bar{t}$+0jet. This is why the extra jet is expected to be the fifth jet in the final state i.e. to be the softest (least energetic) jet.

To investigate the above statement, and for the double tag event samples, we calculated the ΔR separation between the truth plus jet and each of the five ttop25 MC jets. Figure 4.24 shows that the extra jet could be equally likely the fifth jet (24.0% of the time) and the fourth jet (24.1% of the time). The first (lead) and the second jets have a much small likelihood to be the extra jet. Figure 4.25 shows the results for the single tag event samples. It provides similar results that show it is equally likely for the fifth jet (23.0% of the time) and the fourth jet (23.9% of the time) to be the extra jet.

$\mathbf{E_T}$ and η Discriminants

Using the straightforward ΔR calculations on the ttop25 MC sample didn't solve the dilemma of whether the extra jet is the fifth or the fourth jet. Hence, we tried to tackle this issue using the following procedure: There are ten possible combinations for two out of the five jets in the final state to be tagged as `SecVtx` jets [6]. We do not have to look at all of these combinations because the above ΔR calculations have

[6] For double tag ttop25 events, we found that the pair of the lead jet (first jet) and the second lead jet is tagged as `SecVtx` jets $\sim$ 24% of the time, while the pair of the fifth jet and the fourth jet is tagged as `SecVtx` jets only $\sim$ 2% of the time. Out of the ten possible jet pairs, the probability for the jet pairs (first-second) or (first-third) or (second-third) to be `SecVtx` jets is $\sim$ 56%, leaving the 46% probability for the other seven jet pairs.

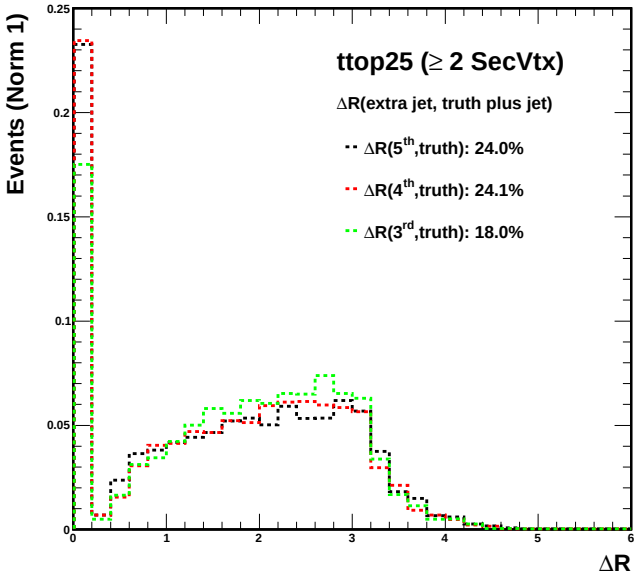


Figure 4.24: ΔR value between the ttop25 jets and the truth plus jet for the double tag event samples.

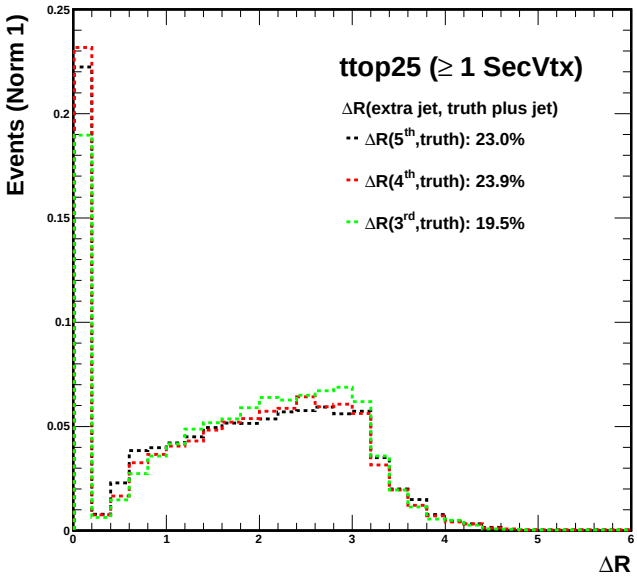


Figure 4.25: ΔR value between the ttop25 jets and the truth plus jet for the single tag event samples.

shown that the likelihood for the first, the second and the third jets to be the extra jet are small and that it is equally likely for the fifth and the fourth jets. Hence, we looked at only the following three combinations:

- If the first and the second jets are tagged as `SecVtx` jets, the third, fourth and fifth jets could be the extra jet candidate (Figures 4.26 and 4.27):
  - When the third and the fourth jets best reconstructed the W boson, we considered the fifth jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.
  - When the third and the fifth jets best reconstructed the W boson, we considered the fourth jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.
  - When the fourth and the fifth jets best reconstructed the W boson, we considered the third jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.
- If the first and the third jets are tagged as `SecVtx` jets, the second, fourth and fifth jets could be the extra jet candidate (Figures 4.28 and 4.29):
  - When the second and the fourth jets best reconstructed the W boson, we considered the fifth jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.
  - When the second and the fifth jets best reconstructed the W boson, we considered the fourth jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.
  - When the fourth and the fifth jets best reconstructed the W boson, we considered the second jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.

- If the second and the third jets are tagged as `SecVtx` jets, the first, fourth and fifth jets could be the extra jet candidate (Figures 4.30 and 4.31):

    - When the first and the fourth jets best reconstructed the W boson, we considered the fifth jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.

    - When the first and the fifth jets best reconstructed the W boson, we considered the fourth jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.

    - When the fourth and the fifth jets best reconstructed the W boson, we considered the first jet to be the extra jet and hence compared its E_T and η distributions with that of the truth plus jet.

We have seen that the first, second and third jets have the least likelihood to be the extra jet. However, both the E_T and η discriminants are unable to favor the fifth jet or the fourth jet over the other. The results show that they are still equally likely to be the extra jet. By definition, the fifth jet E_T distribution should drop off more rapidly than the fourth jet. This is why the η distribution for the former jet is broader than that for the latter jet.

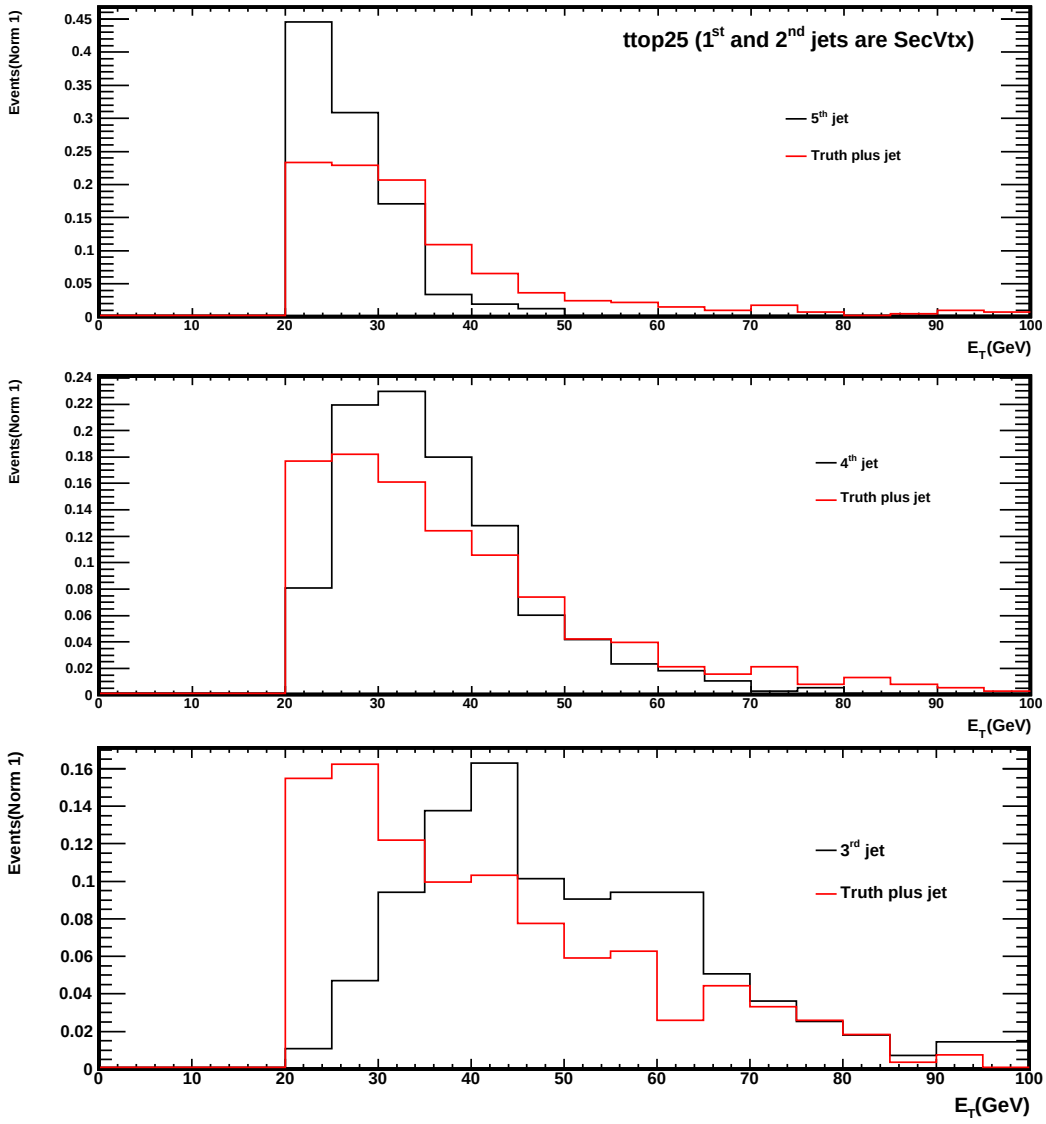


Figure 4.26: E_T distributions for the fifth, fourth and third jets when the first and the second jets are `SecVtx` tagged.

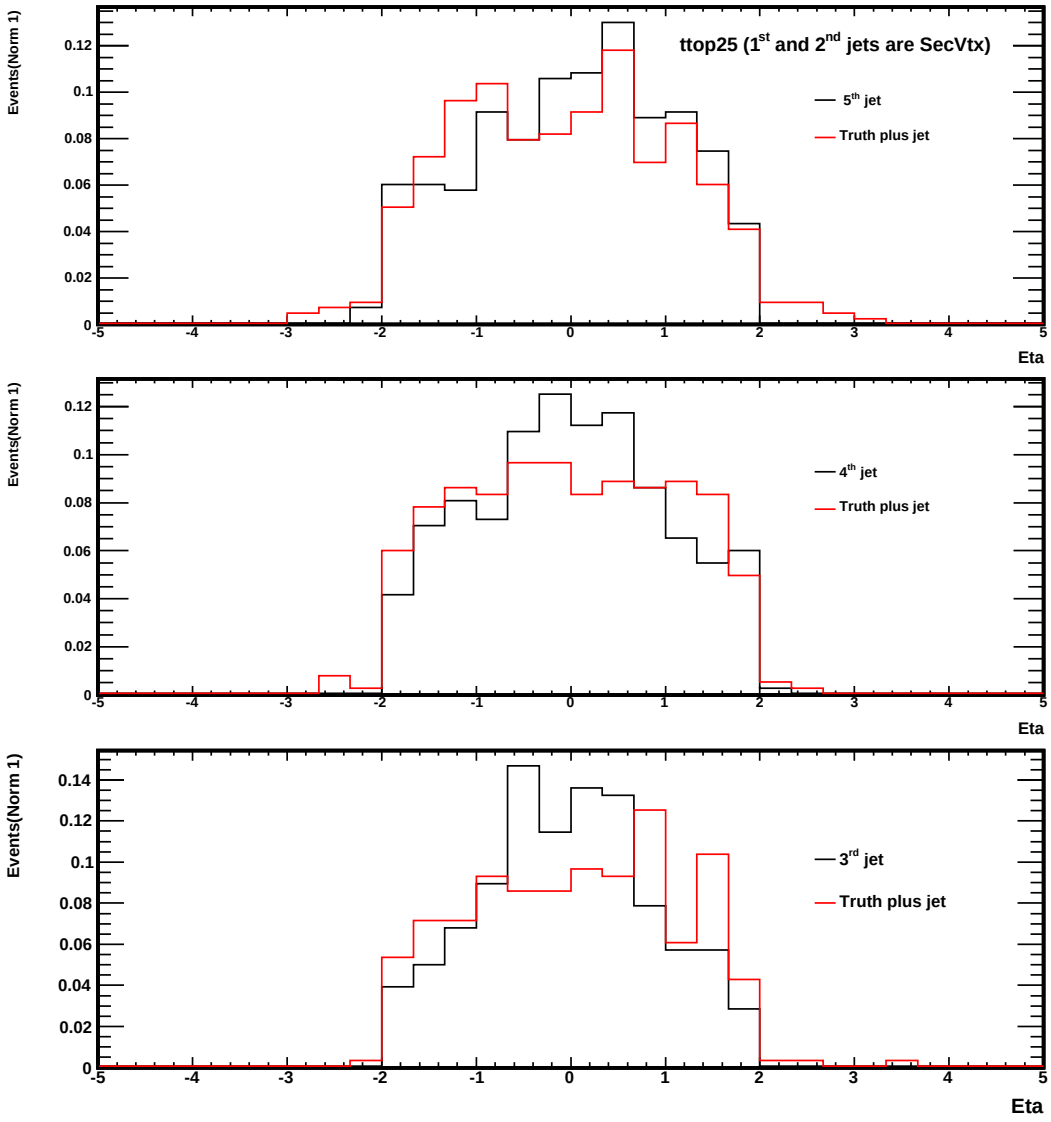


Figure 4.27: η distributions for the fifth, fourth and third jets when the first and the second jets are `SecVtx` tagged.

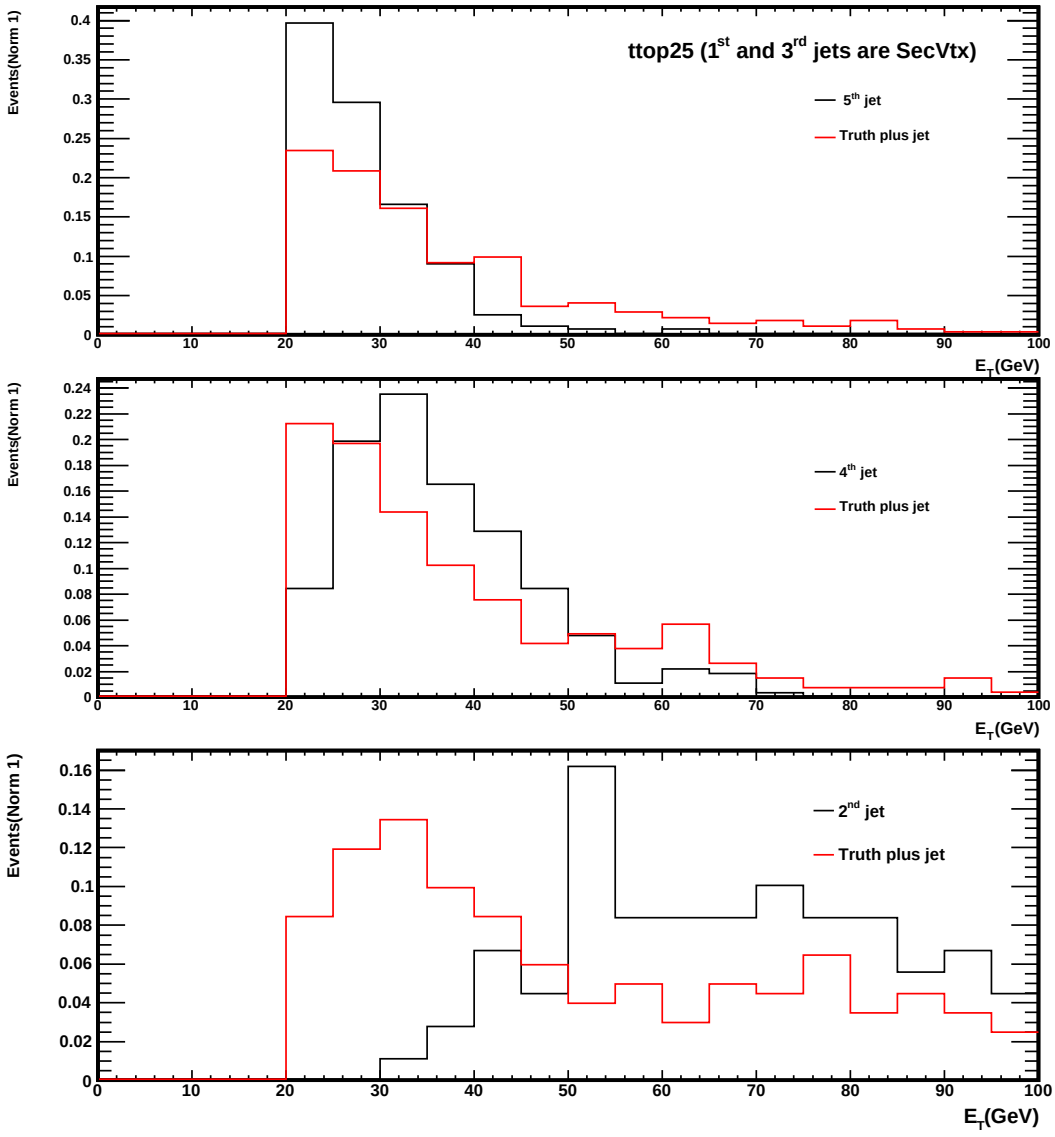


Figure 4.28: E_T distributions for the fifth, fourth and second jets when the first and the third jets are `SecVtx` tagged.

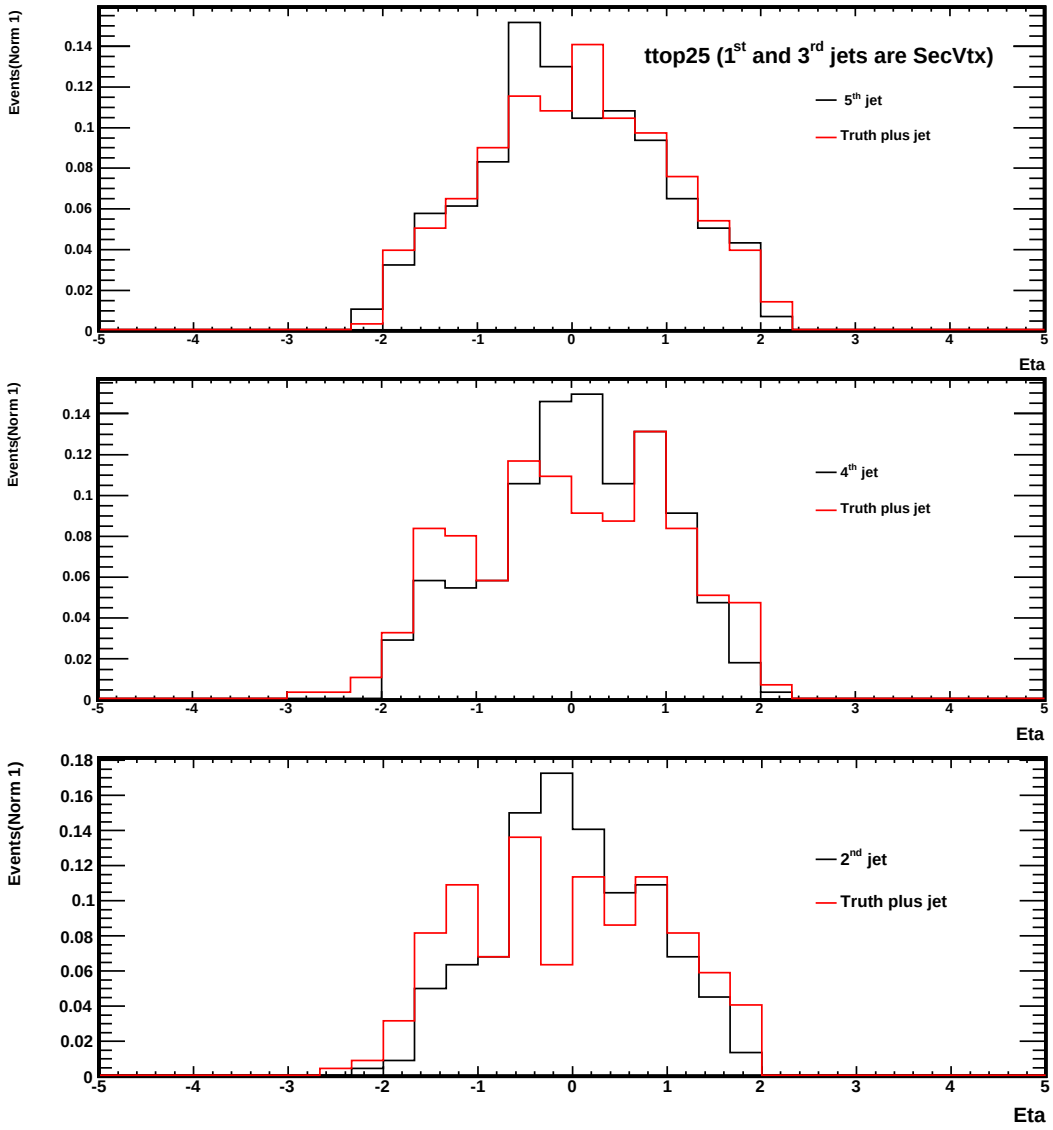


Figure 4.29: η distributions for the fifth, fourth and second jets when the first and the third jets are `SecVtx` tagged.

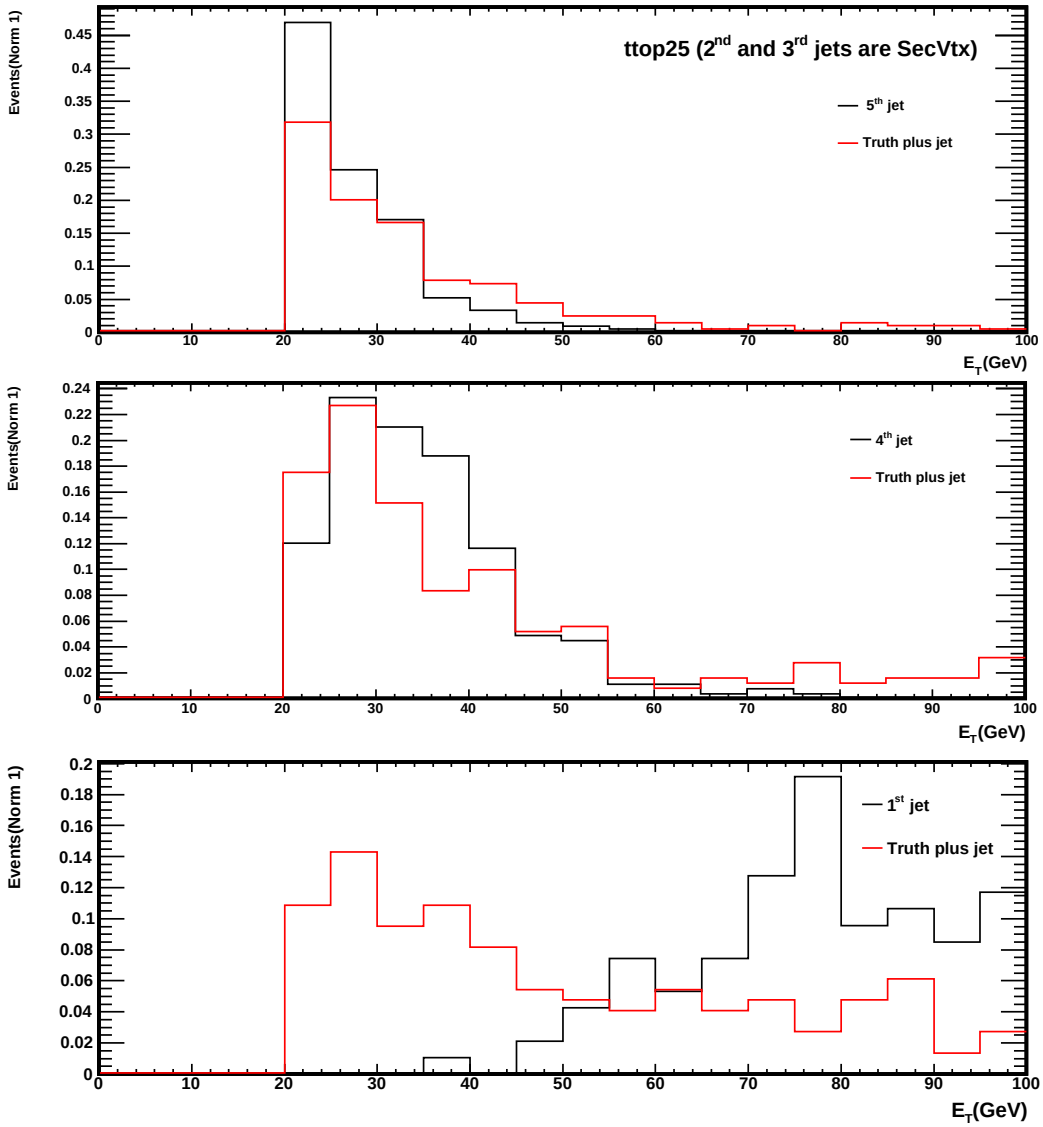


Figure 4.30: E_T distributions for the fifth, fourth and first jets when the second and the third jets are `SecVtx` tagged.

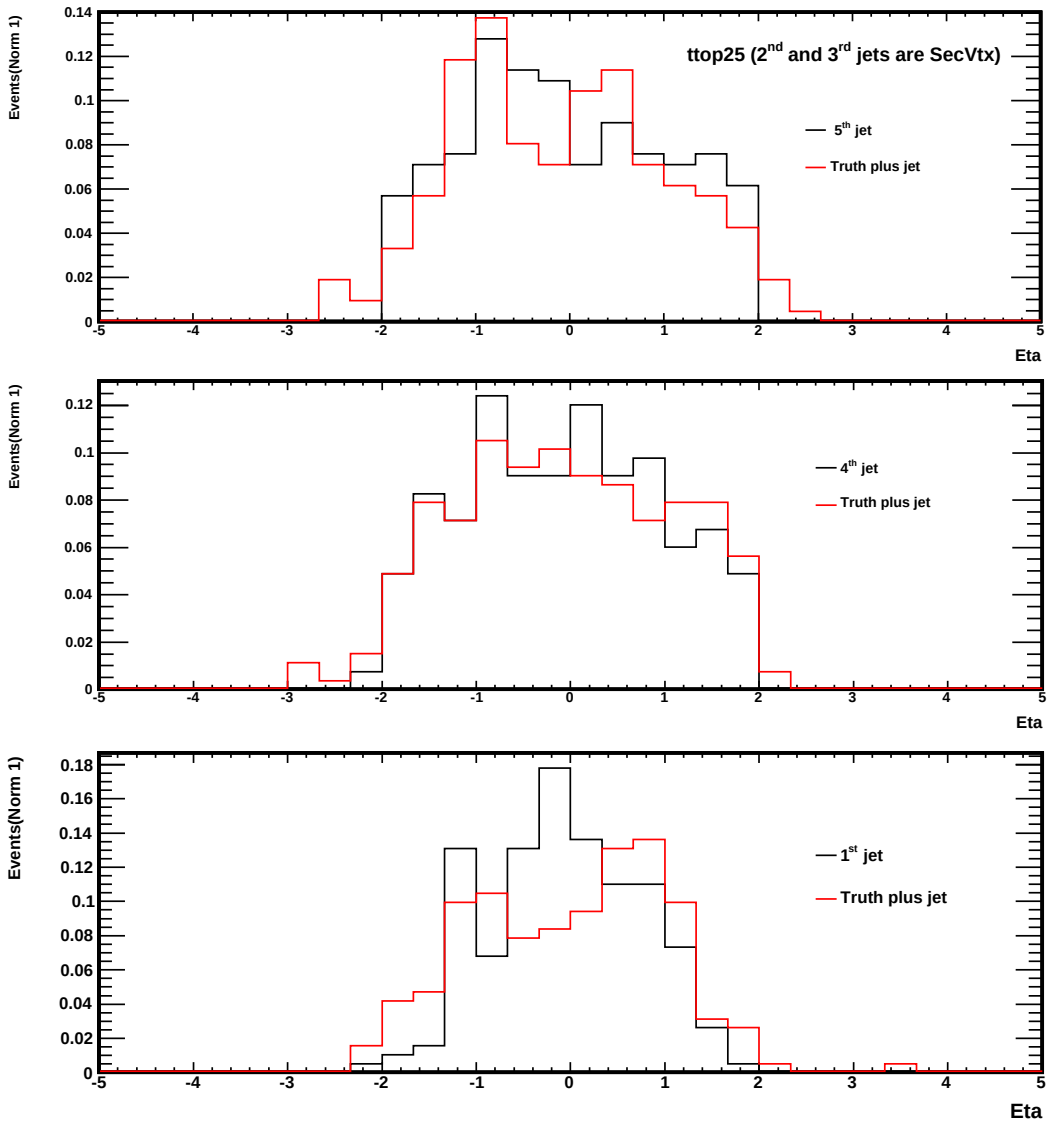


Figure 4.31: η distributions for the fifth, fourth and first jets when the second and the third jets are `SecVtx` tagged.

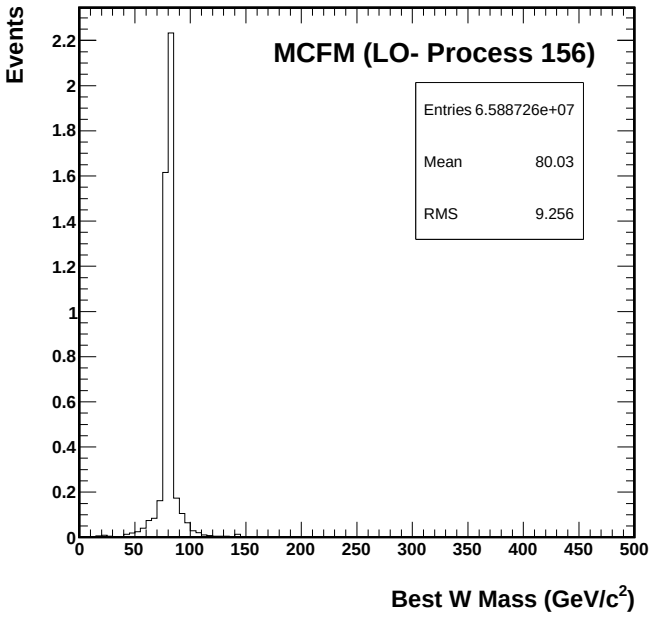


Figure 4.32: The best W boson mass distribution for LO $t\bar{t}$+jet `MCFM` sample.

4.3.5 The `MCFM` Extra Jet

`MCFM` (Monte Carlo for FeMtobarn) is a parton level MC program which gives LO and NLO predictions for a range of processes at hadron colliders [128]. Unfortunately, our $t\bar{t}$+jet process is not included in `MCFM` yet at the NLO level. However, we can still infer some knowledge from LO `MCFM` (process 156) in which an extra gluon is radiated as ISR. We can use `MCFM` to cross check our analysis as well. In this regard, and using our isolation algorithm, Figures 4.32 and 4.33 show the best reconstructed mass and one of the possible combinatoric mass distributions of the W boson respectively. Finally Figure 4.34 shows how often the extra ISR gluon could be the lowest P_T jet among the five jets in the final state. The distribution (the bold black) falls off as $\sim (1/P_T^2)$ as expected in QCD [49].

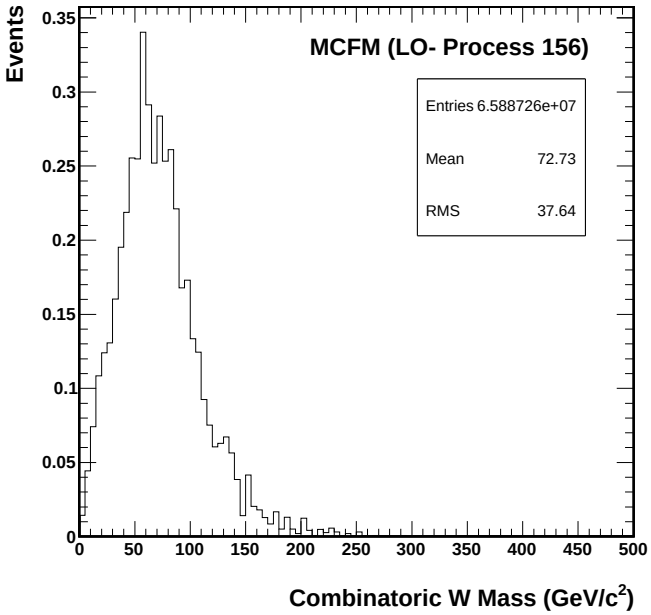


Figure 4.33: A combinatoric W boson mass distribution for LO $t\bar{t}$+jet `MCFM` sample.

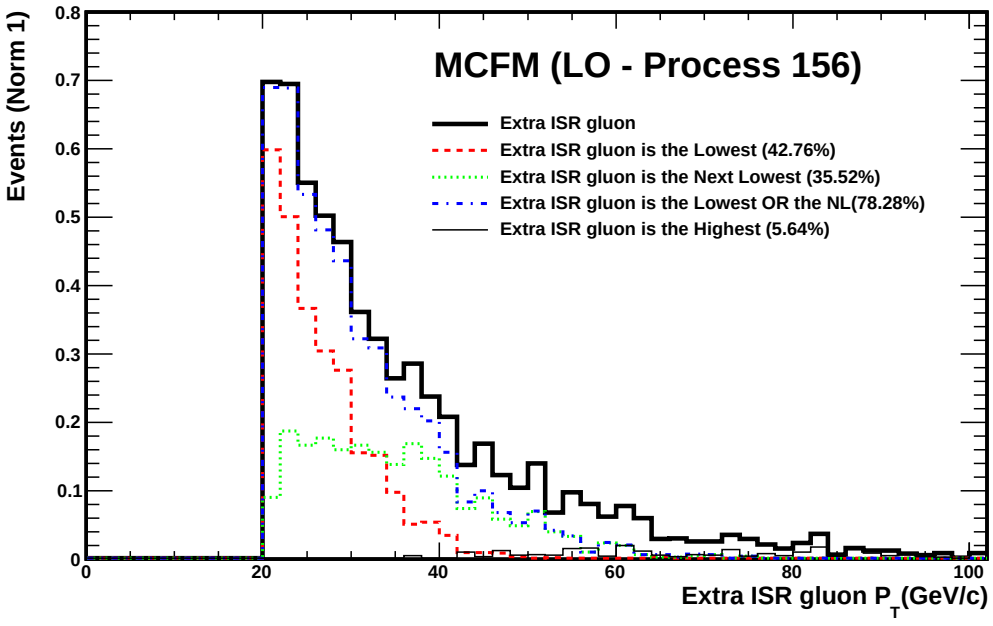


Figure 4.34: The likelihood for the extra ISR gluon to be the lowest P_T jet for LO $t\bar{t}$+jet `MCFM` sample.

4.4 Conclusions

The first measurement of the cross section of the top quark pair in association with an extra hard jet ($t\bar{t}$+jet) has been performed with 4.1 fb^{-1} of data collected at CDF. The measurement is an important test of perturbative QCD, as NLO effects play an important role in the calculation of the theoretical cross section. In addition, it is also important as a preview of the LHC, for which almost half of the top quark events will be produced with extra jets. Therefore, this process will be a substantial background for many new physics signals. The measurement is performed using `SecVtx` tagged events in the lepton plus jet channel. A data-driven approach is used to predict the background content, and a 2D likelihood is formed to simultaneously measure the $t\bar{t}$+jet and $t\bar{t}$ without extra jet ($t\bar{t}$+0jet) cross sections. The measured result is $\sigma_{t\bar{t}+\text{jet}}$= 1.6$\pm$0.2$_{\text{stat}}\pm$0.5$_{\text{syst}}$ pb which is in agreement with the recent NLO SM prediction $\sigma_{t\bar{t}+\text{jet}} = 1.791(1)^{+0.16}_{-0.31}$pb . The measured cross section for $t\bar{t}$+0jet is $\sigma_{t\bar{t}+0\text{jet}}$= 5.5$\pm$0.4$_{\text{stat}}\pm$0.7$_{\text{syst}}$ pb which when combined with $\sigma_{t\bar{t}+\text{jet}}$ gives the inclusive $t\bar{t}$ cross section $\sigma_{t\bar{t}}$= 7.1$\pm$0.3$_{\text{stat}}$ pb, in agreement with the SM prediction.

In order to elucidate the kinematic profile of the extra jet, an isolation algorithm has been developed. The algorithm has extracted correctly the extra jet out from the final state jets more than 60% of the time. This allowed for correcting the measured distributions of the extra jet for purity/efficiency in order to compare them with the MC distributions. The differences in the kinematic of the extra jet using different `SecVtx` requirements and different MC models (`PYTHIA` & `MCFM`) have been studied. The agreement between data and the simulations is reasonable. The calculations have shown that the likelihood for the first, the second and the third highest E_T jet in the final state of $t\bar{t}$+jet sample to be the extra jet is quite small. On the other hand, the fifth and the fourth highest E_T jet in the final state of $t\bar{t}$+jet sample are found to be equally likely the extra jet.

Bibliography

[1] Kronfeld, A., *et al.*, arXiv:1002.5032v4 [hep-ph].

[2] Nakamura, K., *et al.*, J. Phys. **G37**, 075021 (2010).

[3] *http://www.cpepweb.org/cpep_sm_large.html*

[4] Mann, R., An Introduction to Particle Physics and the Standard Model. (Taylor & Francis Group, 2010).

[5] Barger, V. & Phillips, R., Collider Physics . (Perseus Press , 1997).

[6] Martin, A., *et al.*, Eur. Phys. J. C**70**, 51 (2010).

[7] Pumplin, J., *et al.*, JHEP. **07**, 012 (2002).

[8] Fermi, E. , Z. Phys. **88**, 161 (1934).

[9] Glashow, S. L., Nucl. Phys. **22**, 579 (1961).

[10] Weinberg, S., Phys. Rev. Lett. **19**, 1264 (1967).

[11] Salam A., in Elementary Particle Theory, ed. N. Savrtholm, Almquist and Forlag, Stockholm (1968).

[12] 't Hooft, G., Nucl. Phys. **B33**, 173 (1971); 't Hooft, G., Nucl. Phys. **B35**, 167 (1971); 't Hooft, G. and Veltman, M., Nucl. Phys. **B44**, 189 (1972); 't Hooft, G. and Veltman, M., Nucl. Phys. **B50**, 318 (1972).

[13] Arnison, G., *et al.*, Phys. Rev. B **122**, 103 (1983); Bagnaia, P., *et al.*, Phys. Rev. B **122**, 476 (1983); Arnison, G., *et al.*, Phys. Rev. B **126**, 398 (1983); Bagnaia, P., *et al.*, Phys. Rev. B **126**, 130 (1983);

[14] Kraus, C., *et al.*, Eur. Phys. J. **C40**, 447 (2005)

[15] Higgs, P., Phys. Rev. Lett. **12**, 132 (1964); Phys. Rev. Lett. **13**, 508 (1964); Phys. Rev. Lett. **145**, 1156 (1965);

[16] Cabbibo, N., Phys. Rev. Lett **10**, 531 (1963); Kobayashi, M. and Maskawa, T., Prog. Theor. Phys. **49**, 652 (1973).

[17] Herb, S., *et al.*, Phys. Rev. Lett. **39**, 252 (1977)

[18] Abe, F., *et al.*, Phys. Rev. Lett. **74**, 2626 (1995); Abachi, S., *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[19] Deliot, F. and Glenzinski, D. , arXiv:1010.1202v1 [hep-ex].

[20] Wicke, D., arXiv:1005.2460v1 [hep-ex].

[21] Incandela, J., *et al.*, arXiv:0904.2499v3 [hep-ex].

[22] Pleier, M., arXiv:0810.5226v2 [hep-ex].

[23] Mangano, M. and Trippe, T., Phys. Lett. **B592**, 482 (2004).

[24] Kidonakis, E., *et al.*, Phys. Rev. D **64**, 114001 (2001).

[25] Aaltonen, T., *et al.*, Phys. Rev. Lett. **103**, 092002 (2009).

[26] Abazov, V., *et al.*, Phys. Rev. Lett. **103**, 092001 (2009).

[27] The TEVEWWG Working Group , arXiv:0908.2171v1 [hep-ex].

[28] *http://www-cdf.fnal.gov/internal/physics/top/feynman.shtml.*

[29] The TEVNPH Working Group , arXiv:1007.4587v1 [hep-ex].

[30] *http://tevnphwg.fnal.gov/results/SM_smHiggs_Summer_10/*

[31] Fiedler, F., arXiv:1003.0521v1 [hep-ex].

[32] The TEVEWWG Working Group , arXiv:1007.3178v1 [hep-ex].

[33] *http://tevewwg.fnal.gov/top/*

[34] Cembranos, J. , *et al.*, Europhys. Lett. **82**, 21001 (2008).

[35] The CDF Collaboratio, CDF note 10173 (2010).

[36] Abazov, V., *et al.*, Phys. Rev. Lett. **103**, 132001 (2009).

[37] Aaltonen, T., *et al.*, Phys. Rev. Lett. **102**, 042001 (2009).

[38] Abazov, V., *et al.*, arXiv:1009.5686v1 [hep-ex].

[39] Aaltonen, T., *et al.*, Phys. Rev. B **674**, 160 (2009).

[40] Abazov, V., *et al.*, Phys. Rev. Lett. **100**, 062004 (2008).

[41] Stelzer, T. and Willenbrock, S., arXiv:9512292v3 [hep-ph].

[42] Wagner, W., arXiv:0507207v2 [hep-ph].

[43] D. Mietlicki, Ph.D. thesis, University of Michigan (2010).

[44] The CDF Collaboratio, CDF/FNAL/TOP/PUBLIC/10211 (2010).

[45] *http://www-cdf.fnal.gov/physics/new/top/2010/tprop/LJ_spincorr_5invfm/*

[46] The CDF Collaboratio, CDF note 9824 (2009).

[47] Cacciari, M., *et al.*, arXiv:0303085v1 [hep-ph].

[48] The CDF Collaboratio, CDF note 9432 (2008).

[49] Ellis, R., *et al.*, QCD and Collider Physics.
(Cambridge University Press , 2003).

[50] Aaltonen, T., *et al.*, Phys. Rev. D **78**, 111101 (2008).

[51] The CDF Collaboratio, CDF note 8811 (2007).

[52] Aaltonen, T., *et al.*, Phys. Rev. D **79**, 031101 (2009).

[53] Chang, D., *et al.*, Phys. Rev. D **59**, 091503 (1999).

[54] Z. Ünalan, Ph.D. thesis, Michigan State University (2007).

[55] The CDF Collaboration, CDF/PUB/TOP/PUBLIC/9939 (2010).

[56] Abazov, V., *et al.*, Phys. Rev. Lett. **98**, 041801 (2007).

[57] Antunano, Q., *et al.*, Phys. Rev. D **77**, 014003 (2008).

[58] G. Strycker, Ph.D. thesis, University of Michigan (2010).

[59] Aaltonen, T., *et al.*, arXiv:1101.0034v1 [hep-ex].

[60] The D∅ Collaboratio, D∅ Note 6062-CONF.

[61] Takeuchi, Y., *et al.*, CDF/ANAL/TOP/CDFR/10398 v2.2 (2011).

[62] Ahrnes, V., *et al.*, arXiv:1003.5827v3 [hep-ph].

[63] Aliev, M., *et al.*, arXiv:1007.1327v1 [hep-ph].

[64] Langenfeld, U., *et al.*, Phys. Rev. D **80**, 054009 (2009).

[65] Kidonakis, N. and Vogt, R. , Phys. Rev. D **78**, 074005 (2008).

[66] *http://www-cdf.fnal.gov/physics/new/top/2009/xsection/ttbar_combined_46invfb/*

[67] Abazov, V., *et al.*, Phys. Rev. D **80**, 071102 (2009).

[68] CDF Collaboration, CDF note 9913 (2009).

[69] *http://www-cdf.fnal.gov/physics/new/top/2009/xsection/ttbar_combined_46invfb/*

[70] *http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_public.html*

[71] Aaltonen, T., *et al.*, Phys. Rev. Lett. **102**, 222003 (2009).

[72] *http://www-cdf.fnal.gov/physics/new/top/ 2008/tprop/dXs_27fb/webpage/dXs27fb_Public.htm*

[73] Abazov, V., *et al.*, Phys. Lett. B **693**, 515 (2010).

[74] *http://www-d0.fnal.gov/Run2Physics/WWW/results/final/TOP/T10A/*

[75] The ATLAS Collaboration, arXiv:1012.1792v2 [hep-ex].

[76] The CMS Collaboration, arXiv:1010.5994v1 [hep-ex].

[77] The CDF Collaboration, CDF note 10137 (2010), The D∅ Collaboration, D∅ note 6037-CONF (2010).

[78] Körner, J. , *et al.*, Phys. Rev. D **77**, 094011 (2008).

[79] Dittmaier, S., *et al.*, Phys. Rev. Lett. **98**, 262002 (2007).

[80] Dittmaier, S., *et al.*, Eur. Phys. J. C **59 3**, 625 (2009).

[81] Campbell, J., *et al.*, Rept. Prog. Phys. **70**, 89 (2007).

[82] *http://www-cdf.fnal.gov/events/CDFPictures.html*

[83] *http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html*

[84] *http://www-cdf.fnal.gov/upgrades/tdr/tdr.html* along with FERMILAB-Pub-96/390-E

[85] Antos, J., *et al.*, Nucl. Instr. Meth. **A 383**, 13 (1996).

[86] Affolder, T., *et al.*, Nucl. Instr. Meth. **A 526**, 249 (2004).

[87] Acosta, D., *et al.*, Nucl. Instr. Meth. **A 518**, 605 (2004).

[88] Minemura, H., *et al.*, Nucl. Instr. Meth. **A 238**, 18 (1985).

[89] Balka, L., *et al.*, Nucl. Instr. Meth. **A 267**, 272 (1988).

[90] Bertolucci, S., *et al.*, Nucl. Instr. Meth. **A 267**, 301 (1988).

[91] Oishi, R., *et al.*, Nucl. Instr. Meth. **A 453**, 227 (2000).

[92] Dorigo, T., *et al.*, Nucl. Instr. Meth. **A 461**, 560 (2001).

[93] Acosta, D., *et al.*, Nucl. Instr. Meth. **A 461**, 540 (2001); Nucl. Instr. Meth. **A 494**, 57 (2002).

[94] *http://www-bd.fnal.gov/pplot/today/DataSummaryTables.html*

[95] *http://www-cdfonline.fnal.gov/ops/opsshelp/stores/*

[96] Amidei, D., *et al.*, Nucl. Instr. Meth. **A 265**, 326 (1988).

[97] Abulencia, A., *et al.*, Nucl. Instr. Meth. **A 581**, 482 (2007).

[98] Adelman, J., *et al.*, Nucl. Instr. Meth. **A 572**, 361 (2007).

[99] *http://ncdf70.fnal.gov:8001/PerfIDia/PerfIDia.html.*

[100] *http://www-cdf.fnal.gov/internal/physics/top/jets/corrections.html*

[101] Group, R., Ph.D. thesis, University of Florida (2006).

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

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

[104] Bhatti, A., *et al.*, Nucl. Instr. Meth. **A 566**, 375 (2006).

[105] Field, R., CDF/PUB/JET/PUBLIC/7898 (2005).

[106] Acosta, D., *et al.*, Phys. Rev. D **71**, 052003 (2005).

[107] *http:www-cdf.fnal.gov/internal/physics/top/RunIIBtag/bTag.html*

[108] T. Schwarz, Ph.D. thesis, University of Michigan (2006).

[109] *http:www-cdf.fnal.gov/internal/physics/top/luminosity.shtml*

[110] Sjöstrand, T., *et al.*, JHEP **05**, 026 (2006).

[111] Sjöstrand, T., *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

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

[113] Mangano, L., *et al.*, JHEP. **0307**, 001 (2003).

[114] Alwall, J., *et al.*, JHEP. **0709**, 028 (2007).

[115] Sulkimo, J., *et al.*, Nucl. Instr. Meth. **A 506**, 250 (2003).

[116] *http://www-cdf.fnal.gov/internal/mcProduction/oksana/Overview.html.*

[117] *http://projects.hepforge.org/spartyjet/.*

[118] Budd, S., *et al.*, CDF/ANAL/TOP/CDFR/8489 (2006).

[119] Messina, A. and Sarkar, S., CDF/PHYS/PUBLIC/6636 (2003).

[120] Adelman, J., *et al.*, CDF/ANAL/TOP/CDFR/9185 (2008).

[121] Franklin, M., *et al.*, CDF/ANAL/TOP/CDFR/8766 (2007).

[122] The CDF Collaboration, CDF Note 9850 (2009).

[123] Schwarz, T., *et al.*, CDF/PUB/ANAL/CDFR/9676 (2009).

[124] *http://www-cdf.fnal.gov/physics/new/top/2009/xsection/ttj_4.1invfb/.*

[125] Mrenna, S., Fermi National Accelerator, private correspondence.

[126] Lyons, L., Statistics for Nuclear and Particle Physicists. (Cambridge University Press, 1992).

[127] *http://www-cdf.fnal.gov/internal/physics/joint_physics/ instructions/PDFUncertainties/pdf.html.*

[128] MCFM v5.0, Users Guide , *http://mcfm.fnal.gov/.*