UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MEASUREMENT OF TOP QUARK PAIR PRODUCTION CROSS SECTION IN PROTON ANTIPROTON COLLISIONS AT $\sqrt{s} = 1.96$ TeV IN THE $\tau + jets$ FINAL STATE USING $1 fb^{-1}$ OF DATA

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

BY

SOHRAB HOSSAIN

Norman, Oklahoma

2009

MEASUREMENT OF TOP QUARK PAIR PRODUCTION CROSS SECTION IN PROTON ANTIPROTON COLLISIONS AT $\sqrt{s} = 1.96$ TeV IN THE $\tau + jets$ FINAL STATE USING $1 fb^{-1}$ OF DATA

A DISSERTATION APPROVED FOR THE
HOMER L. DODGE DEPARTMENT OF PHYSICS AND ASTRONOMY

BY

Dr. Phillip Gutierrez

Dr. Brad K. Abbott

Dr. Chung Kao

Dr. Gregory A. Parker

Dr. James Hawthorne

Acknowledgments

I would first like to thank my supervisor Phillip Gutierrez for his guidance, support and his unwavering confidence in my abilities. I am really privileged to have him as my advisor and my dissertation chair.

I am grateful to Mikhail Arov, Dhiman Chakraborty, Amnon Harel, Haryo Sumowidagdo, Ike Hall, and Supriya Jain both professionally and personally. I am extremely appreciative to the members of my dissertation committee, Brad Abbott, Gregory Parker and James Hawthorne for all their support throughout my PhD program. I would also like to thank members of DØ collaboration for their suggestions and for their invaluable inputs that helped me immensely towards the completion of this dissertation.

I am grateful to my father Emdad Hossain, my mother Rubia Begum for all their love and care. I would like to especially thank my best friend and my wife Nafeesa Yunus for the sacrifices she made for me all through my PhD years. Last but not least I would like to thank our son Ian, whose wonderful smiles make every single day worth living to the fullest.

Contents

List Of Tables

List Of Figures

Abstract

This dissertation presents a new measurement of $p\bar{p} \rightarrow t\bar{t}X$ production at $\sqrt{s} =$ 1.96 TeV using 974.2 pb^{-1} of data collected with the DØ detector between 2002 and 2006. We focus on the final state where the W boson from one of the top quarks decays into a τ lepton and its associated neutrino, while the other W boson decays into a quark-antiquark pair. We aim to select those events in which the τ lepton subsequently decays hadronically, meaning to one or three charged hadrons, zero or more neutral hadrons and a tau neutrino (the charge conjugate processes are implied in all of the above). The observable signature thus consists of a narrow calorimeter shower with associated track(s) characteristic of a hadronic tau decay, four or more jets, of which two are initiated by b quarks accompanying the W's in the top quark decays, and a large net missing momentum in the transverse plane due to the energetic neutrino-antineutrino pair that leave no trace in the detector media. The preliminary result for the measured cross section is:

$$\sigma(t\bar{t}) = 6.9\ ^{+1.2}_{-1.2}\ (\text{stat})\ ^{+0.8}_{-0.7}\ (\text{syst})\ \pm 0.4\ (\text{lumi})\ \text{pb}.$$

This indicates that our finding is consistent with the Standard Model prediction.

Chapter 1

Introduction

The Standard Model (SM) is a quantum field theory that describes all the fundamental particles and their interactions [1]. It is one of the most successful theories in high energy physics in that the model predicted the existence of the top quark before it was discovered in 1995 [2, 3].

The top quark is the heaviest fundamental particle in the SM description of nature. Production of such a heavy particle requires TeV range energy. The Fermilab Tevatron accelerator is designed to collide protons and antiprotons at a center of momentum energy ($\sqrt{s}$) of 1.96 TeV to produce top pair events predicted by the Standard Model.

The precise measurements of top quark pair production rate and other properties allow us to perform precision tests of the Standard Model. On the other hand, models which predict new processes that are not allowed by the Standard Model can also be tested by measuring the top pair production rate.

The top quark pair production rate or cross section is measured in different decay channels. The decay of the top quark into the tau (τ) lepton is especially interesting to study, since it is the heaviest lepton. Any non-standard mass-dependent couplings could produce a very significant effect in this channel. An example is the charged Higgs boson which appears in extensions of the SM Higgs sector (Two-Higgs Doublet Models) [16]. Since the Higgs-fermion coupling is proportional to the fermion's mass, charged Higgs decays to the heavy tau lepton would be much more frequent than those to the lighter leptons (electron e or muon μ).

The measurement in this channel is rather challenging compared to other top decay channels as it is difficult to separate the multijet background events from the signal events. This dissertation performs the test of the SM in the $t\bar{t} \rightarrow \tau + jets$ channel where the final state is characterized by a tau lepton, jets and large missing transverse energy and develops tools for the future search of the charged Higgs bosons. The results obtained represent substantial improvement over those previously shown in the preliminary version of this analysis [91].

The description of the cross section measurement procedure is organized as follows: Chapter 2 gives a brief introduction to the SM along with description of top related physics. Chapter 3 describes the particle accelerator at Fermilab used to produce, accelerate and collide protons and antiprotons and also describes the DØ detector which is used to detect the product of the collisions. Chapter 4 describes how the information from the detector is utilized to form different physics objects connected with the protons and antiprotons collisions. Chapter 5 explains $t\bar{t} \rightarrow \tau + jets$ data event selection procedure and develops tools to model data and background. Finally, chapter 6 explains the procedure for modeling signal events and background events from data with the tools developed in chapter 5. This chapter also describes how the cross section is measured with the selected number of signal events.

Chapter 2

Theory

2.1 The Standard Model

The SM of particle physics is a theory that describes all interactions in terms of fundamental forces and fundamental particles. The SM encompasses three[1] types of interactions: electromagnetic, strong and weak interactions, of which the electromagnetic interaction is the most commonly known. The electromagnetic (EM) interaction (described by Quantum Electrodynamics (QED)) and the weak interaction are combined in the electroweak sector of the SM. The strong interaction is described in the SM by the field theory called Quantum Chromo Dynamics (QCD).

2.1.1 Particles in the Standard Model

In the SM, the fundamental particles are grouped into two categories: sources of fields (spin $\frac{1}{2}$) fermions and the mediators (spin 1 gauge bosons). In terms of interactions, the fermions are subdivided into quarks and leptons. The quarks interact via the strong force while leptons do not experience the strong force at all. However, both quarks and leptons can interact via the electromagnetic and weak forces (uncharged particles like neutrinos do not experience the electromagnetic force). There are six quarks and six leptons and there is an antimatter particle for every matter particle (Table 2.1). Moreover, both quarks and leptons are grouped

[1]Gravitational interaction is not included in the SM.

into three generations. The lightest particles are in the first generation while the heaviest are in the third generation. The grouping of these particles' generations along with their properties are described in Table 2.1. As shown in Table 2.1 both the top quark and the tau lepton belong to the third (or heavy) generation.

The quarks are fermions and have an extra quantum number for strong interactions - color (the strong interaction charge). There are three possible colors for each flavor of the quark. The color-charged quarks are confined in groups with other quarks to form color neutral hadrons. The hadrons are divided into two categories; baryons (qqq) and mesons ($q\bar{q}$). Baryons are fermions while mesons are bosons.

The second type of elementary particles are the spin one mediators or bosons. The bosons are comprised of a photon, γ, which is the exchange particle in electromagnetic interactions, eight gluons which mediate the strong interactions among the quarks, and three weak bosons, $W^{\pm}$ and Z which are the corresponding mediators of the weak interactions as shown in Table (2.2).

However, the SM does not explain why three of these bosons are observed to be massive while one is massless. To allow for massive force carriers, a mechanism called the Higgs mechanism or electroweak symmetry breaking is introduced. This mechanism gives rise to another massive particle called the Higgs boson which has not been observed yet.

	Charge		**Mass in MeV**		**Mass in MeV**		**Mass in GeV**
Quarks	2/3	u	1.5 to 3.3	c	$1.27^{+0.07}_{-0.11}$ GeV	t	$171.3 \pm 1.1 \pm 1.2$
Quarks	- 1/3	d	3.5 to 6	s	105^{+25}_{-35}	b	$4.2^{+0.17}_{-0.07}$
Leptons	0	ν_e	$< 2 \times 10^{-6}$	ν_μ	< 0.19	ν_τ	$< 18.2 \times 10^{-3}$
Leptons	-1	e	.511	μ	105.66	τ	1776.84 ± 0.17 MeV
Generations		I		II		III	

Table 2.1: Three generations of the SM particles [6].

Boson	Charge	Mass in GeV
gluon	0	massless
γ	$< 1 \times 10^{-35} e$	$< 1 \times 10^{-24}$
$W^{\pm}$	± 1	80.398 ± 0.025
Z	0	91.1876 ± 0.0021

Table 2.2: Properties of bosons [6].

2.1.2 Mathematical Structure of the SM

The Standard Model is a quantum field theory based on the gauge symmetry $SU(3)_C \bigotimes SU(2)_L \bigotimes U(1)_Y$ (where $SU(N)$ means special unitary group of dimensions $N \times N$, C stands for color and L stands for left handed and Y stands for weak hypercharge generator of $U(1)$). This gauge group includes the $SU(3)_C$ group of strong interactions and $SU(2)_L \bigotimes U(1)_Y$ for electroweak interactions.

The hypercharge Y is related to the electric charge via the Gell-Mann Nishijima formula $Q = I^3 + \frac{Y}{2}$. Where I^3 represents the 3rd component of the weak isospin quantum number I.

The gauge sector of the SM model is composed of eight gluons that are the gauge bosons of $SU(3)_C$ and four gauge bosons $\gamma, W^{\pm}$ and Z of $SU(2)_L \bigotimes U(1)_Y$. The gluons are massless and carry color quantum numbers. The eight gluons come in eight different net colors. Because of their color they not only interact with quarks but also interact among themselves.

As mentioned earlier, the gauge bosons of the weak interaction, $W^{\pm}$, Z are massive while the gauge boson of the electromagnetic interaction, γ, is massless. In the SM, the fermions and bosons acquire masses via electroweak symmetry breaking. This symmetry breaking is achieved by introducing a scalar field into the model, the Higgs field ϕ, where

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.1}$$

The simplest and most general Lagrangian for the Higgs field, consistent with gauge symmetry is

$$\mathcal{L}_{Higgs} = (\mathcal{D}^{\mu}\phi)^{\dagger}\mathcal{D}_{\mu}\phi + \mu^2\phi^{\dagger}\phi - \lambda(\phi^{\dagger}\phi)^2, \tag{2.2}$$

where the covariant derivative is defined as

$$\mathcal{D}_{\mu} = \partial_{\mu} - igA^a_{\mu}\tau^a - i\frac{1}{2}g'B_{\mu}, \tag{2.3}$$

A^a_{μ} and B_{μ} are the $SU_L(2)$ and $U_Y(1)$ gauge bosons respectively, and both g and g' are the couplings.

The first term in Eq. 2.2 is the Higgs-field kinetic energy and the rest is the potential term. After symmetry breaking the minimum of the potential lies not at zero, but on a circle of minima with radius given by

$$R = \frac{\mu}{\sqrt{2\lambda}} \equiv \frac{v}{\sqrt{2}}, \tag{2.4}$$

where v is the vacuum expectation value. Making the substitution $\phi = (0, v/\sqrt{2})$ in the Higgs Lagrangian one finds that the W and Z boson acquire masses from the interaction of the gauge bosons with the Higgs field. By extracting the coefficients of the terms $A_{\mu}A^{\mu}$ (mass terms) the W and Z masses can be written as

$$M_W = \frac{1}{2}gv \quad M_Z = \frac{1}{2}\sqrt{g^2 + g'^2}\ v. \tag{2.5}$$

The photon does not interact with the Higgs field and remains massless. Similarly, mass can be given to quarks and leptons by invoking the Higgs mechanism of spontaneous symmetry breaking. For example, after symmetry breaking the electron mass term is

$$m_e = \frac{1}{\sqrt{2}}\lambda_e\, v \tag{2.6}$$

where λ_e is a renormalizable coupling[2] and treated as input to the theory.

[2]This means to succeed in adjusting the bare coupling constants to fit the experimentally observed physical coupling constants

2.2 Why the top quark must exist

2.2.1 Anomaly Cancellation

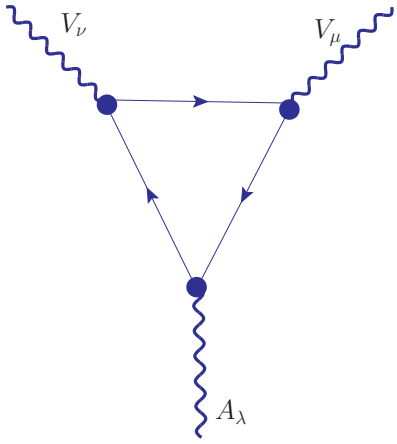

Figure 2.1: An example of a triangular fermion loop causing a chiral anomaly.

In Quantum Field Theory it is possible that gauge invariance holds for tree level (Leading Order) diagrams but not for loop diagrams (Higher Order). These types of violations are called anomalies. Figure 2.1 shows a Feynman diagram producing an anomaly where two vector currents (V) and one axial current (A) are coupled via a fermion triangle. This type of anomaly threatens the renormalizability of the SM.

With leptons only, the SM is not renormalizable. When quarks with three colors are included, the chiral anomaly disappears and renormalizability is restored. For the third generation fermion charges (Table 2.1),

$$\sum Q_L = Q_\tau + Q_{\nu_\tau} + 3 \times (Q_t + Q_b), \tag{2.7}$$

or

$$\sum Q_L = -1 + 0 + 3 \times [(\frac{2}{3}) + (-\frac{1}{3})] = 0. \tag{2.8}$$

Therefore, the top quark with $Q_t = 2/3$ must exist.

2.2.2 b-Quark Decay Properties

Experimental data from $Z \to b\bar{b}$ rule out the hypothesis of an isosinglet b-quark. The Z boson couples to quarks through vector and axial couplings as follows :

$$\sqrt{\frac{G_F m_Z^2}{2\sqrt{2}}}\gamma_\mu[C_V - C_A\gamma_5], \tag{2.9}$$

where,

$$C_V = 2[I^{3L} + I^{3R}] - 4e_q \sin^2\theta_W \tag{2.10}$$

and

$$C_A = 2[I^{3L} - I^{3R}], \tag{2.11}$$

In the high energy limit, the Z boson decay rate [8] can be written as,

$$\Gamma(Z \to b\bar{b}) \approx \frac{G_F m_Z^3}{8\pi\sqrt{2}}\ [C_V^2 + C_A^2]. \tag{2.12}$$

In the SM, the top and bottom quarks form an isospin doublet and their isospin assignments are $I^{3L} = \pm\frac{1}{2}$. In the case of a topless model i.e. isosinglet bottom quark the value of I^{3L} is zero. The isospin assignment for righthanded particles is always zero. For an isosinglet model the branching ratio is,

$$\frac{\Gamma(Z \to b\bar{b})^{I_b^3=0}}{\Gamma(Z \to had)^{I_b^3=-\frac{1}{2}}} = \sum_{q=u,d,c,s,b} \frac{16e_b^2 \sin^4\theta_W}{8(I_q^{3L})^2 - 16e_q I_q^{3L}\sin^2\theta_W + 16e_q^2\sin^4\theta_W} \approx 0. \tag{2.13}$$

On the other hand, the isospin doublet model with top quark yields

$$\frac{\Gamma(Z \to b\bar{b})^{I_b^3=-\frac{1}{2}}}{\Gamma(Z \to had)^{I_b^3=-\frac{1}{2}}} = \sum_{q=u,d,c,s,b} \frac{8(I_b^{3L})^2 - 16e_b I_b^{3L}\sin^2\theta_W + 16e_b^2\sin^4\theta_W}{8(I_q^{3L})^2 - 16e_q I_q^{3L}\sin^2\theta_W + 16e_q^2\sin^4\theta_W} \approx 0.21. \tag{2.14}$$

The experimental value [6] of the branching ratio is

$$\frac{\Gamma(Z \to b\bar{b})}{\Gamma(Z \to had)} = 0.2219 \pm 0.0017. \tag{2.15}$$

The close agreement between the experimental results and the theoretical prediction of branching ratio rules out the isosinglet model of b quark.

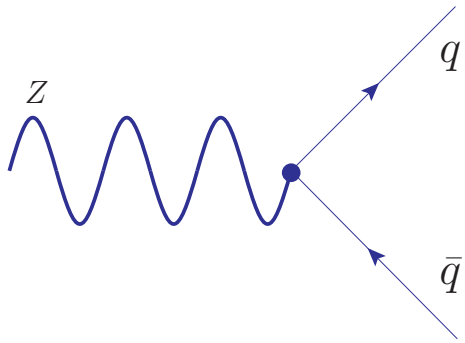

Figure 2.2: Feynman diagram for a Z boson decay.

Forward Backward Asymmetry By counting the number of b quark events in $e^-e^+ \rightarrow b\bar{b}$ production at Z resonance in the angular range $\cos\theta \geq 0$ (forward production) and those where $\cos\theta \leq 0$ (backward production), we can define an asymmetry, called the forward-backward asymmetry:

$$A^{FB} = \frac{3}{4} A_e A_b, \tag{2.16}$$

where

$$A_b \equiv \frac{2C_V^b C_A^b}{(C_V^b)^2 + (C_A^b)^2}. \tag{2.17}$$

Substituting C_V and C_A from Eq. 2.10 and Eq. 2.11 and setting $I^{3R} = 0$ we get

$$A_b = \frac{[2I^{3L} - 4e_b \sin^2\theta_W] \times I^{3L}}{([2I^{3L} - 4e_b \sin^2\theta_W])^2 + (2I^{3L})^2}. \tag{2.18}$$

For a topless model, A_b is zero but experimental measurements of A_b at LEP give approximately 0.883 ± 0.031 [8].

From experimental measurements [7, 8] of the Z boson decay width to b quarks and the value of A^{FB}, the isospin components of the left and right handed b quarks are found to be

$$I_b^{3L} = -0.490^{+0.015}_{-0.012} \rightarrow -1/2 \tag{2.19}$$

$$I_b^{3R} = -0.028 \pm 0.056 \rightarrow 0. \tag{2.20}$$

Therefore, from the non zero value of the b-quark isospin, the top quark can be inferred as the isospin partner of the b quark.

2.2.3 Theoretical Prediction of Top Quark Mass

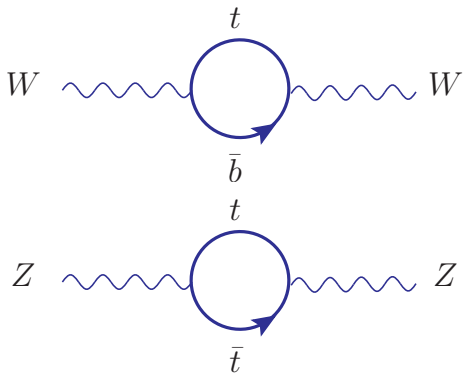

Figure 2.3: Virtual top quark loops contributing to W and Z masses.

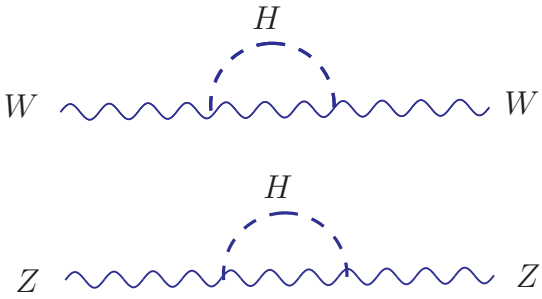

Figure 2.4: Virtual Higgs loops contributing to W and Z masses.

The top quark plays an important role in precision electroweak measurements. For instance in tree level processes, all electroweak measurements depend on precisely

measured quantities α, G_F and M_Z while in higher order processes they depend on the top quark mass (2.3) from which we can estimate an upper limit of its mass.

In the SM for tree level diagrams the value of ρ parameter defined by Eq. 2.21 is exactly one:

$$\rho = \frac{M_W^2}{M_Z^2(1-\sin^2\theta_W)} = 1. \tag{2.21}$$

But for one loop corrections it receives contribution from gauge boson, Higgs boson, and fermion loops. The largest contribution comes from the top quark loop because of its large mass. Calculation of loop diagram in Fig. 2.4 gives,

$$\delta\rho = \frac{3G_F M_t^2}{8\sqrt{2}\pi^2}. \tag{2.22}$$

The experimental value of ρ is equal to $1.00126^{+.00234}_{-.0014}$. The indirect prediction of the top quark mass made from precision measurements [9] is the following:

$$M_t = 170.7 \pm 10.3 \text{ GeV} \quad \text{LEP1/SLD}. \tag{2.23}$$

Higgs mass With the Higgs loop contribution, shown is Fig. 2.4, the correction to ρ is,

$$\delta\rho = \frac{3\ G_F M_t^2}{8\pi^2\sqrt{2}} - \frac{11\sqrt{2}\ \ G_F \sin^2\theta_W}{12\pi^2\ \ \cos^2\theta_W} M_W^2[\ln(\frac{M_H^2}{M_W^2}) + \cdots] + \cdots . \tag{2.24}$$

Thus $\delta\rho$ depends logarithmically on the mass of the Higgs boson. With precisely measured values for both M_t and M_W we can put a limit on the Higgs mass that it is lower than 285 GeV with a 95% confidence level.

2.3 Top Pair Production via Strong Interaction

The Fermilab Tevatron collider produces top quark pairs ($t\bar{t}$) in collisions of protons and antiprotons at the center of momentum energy ($\sqrt{s}$) of 1.96 TeV. The leading

order (LO) Feynman diagrams for top quark pair production via the strong interaction are shown in Fig. 2.5. In $p\bar{p}$ collisions 85% of top quark pairs are produced via $q\bar{q} \rightarrow t\bar{t}$ and 15% are produced via $gg \rightarrow t\bar{t}$ processes.

The theoretical next-to-leading order prediction for $p\bar{p} \rightarrow t\bar{t}$ cross section at $\sqrt{s} = 1.96$ TeV is about 6.8 ± 0.8 pb [15].

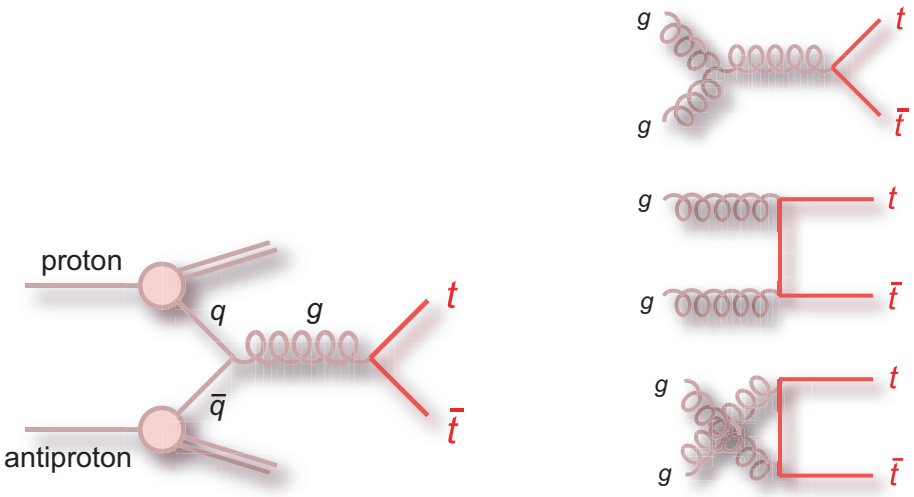

Figure 2.5: Top quark production via the strong interaction at the Tevatron. Quark-antiquark annihilation (left diagram) and gluon fusion (right diagrams).

2.4 Top Decay

2.4.1 Weak Mixing and the CKM Matrix

The couplings of leptons to $W^{\pm}$ take place only within a particular generation (Table 2.1). For example, interactions are of the form $e^- \rightarrow \nu_e + W^-$, $\mu^- \rightarrow \nu_\mu + W^-$, $\tau^- \rightarrow \nu_\tau + W^-$ but no cross-generational coupling of the type $e^- \rightarrow \nu_\mu + W^-$ have been observed. Although the quark generation structure is similar to that of the leptons, the quarks not only interact within their own generation but also with the quarks in other generations via charged weak currents. For instance, within

the first generation, the interaction has the form $d \rightarrow u + W^-$ and the cross-generational interaction has the form $s \rightarrow u + W^-$. This also implies that the quark flavor eigenstates of the weak interactions differ from the mass eigenstates. The Cabibbo-Kobayashi-Maskawa matrix (CKM) [10] relates the weak interaction generations to physical quark states:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{Mass} = \begin{pmatrix} U_{ud} & U_{us} & U_{ub} \\ U_{cd} & U_{cs} & U_{cb} \\ U_{td} & U_{ts} & U_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{Flavor} \tag{2.25}$$

The values of various matrix elements in equation 2.25 are obtained from experimental study of weak interactions. By using the unitarity of the CKM matrix, the assumption of three generations of fermions and the measured values of $|U_{ub}|$ and $|U_{cb}|$, a limit is set on $|U_{tb}|$.

$$1 = |U_{ub}|^2 + |U_{cb}|^2 + |U_{tb}|^2 \tag{2.26}$$

giving,

$$0.9991 < |U_{tb}| < 0.9994. \tag{2.27}$$

The values of the matrix elements $|U_{ts}|$ and $|U_{td}|$ are found to be extremely small (close to zero). The branching ratio [9] is given by

$$\frac{\Gamma(t \rightarrow Wb)}{\Gamma(t \rightarrow Wq)} = \frac{|U_{tb}|^2}{|U_{td}|^2 + |U_{ts}|^2 + |U_{tb}|^2} = 0.94^{+.31}_{-.24} \text{ (CDF)}. \tag{2.28}$$

With the assumption of three generations, the top quarks produced at the Tevatron almost entirely decay into W bosons and b quarks.

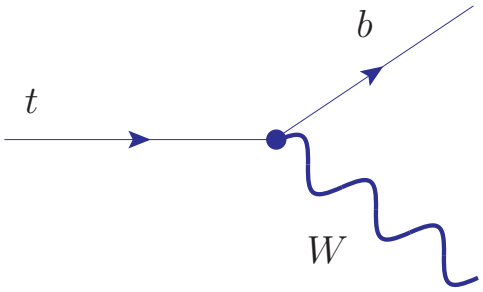

Figure 2.6: Feynman diagram for a top quark decay.

2.4.2 Top Decay via Weak Interaction

The top quark has the same vector-axial charge current weak interaction as all the other fermions. The vertex factor is given by

$$-i\frac{g}{2\sqrt{2}}U_{tb}\gamma^{\mu}(1-\gamma^{5}). \tag{2.29}$$

Figure 2.6 illustrates the Feynman diagram of a top quark decay to a bottom quark via charged weak interaction. As the masses of the b quark and W boson are small compared to the top mass, the decay width can be written as,

$$\Gamma(t \to bW^{+}) \approx |U_{tb}|^{2}\ \frac{G_F M_t^3}{8\pi\sqrt{2}} \approx 1.729\ \text{GeV}. \tag{2.30}$$

Using the decay width, the top quark lifetime can be calculated as follows:

$$\tau \sim 3.8 \times 10^{-25} sec. \tag{2.31}$$

Therefore, because of its very short lifetime the top quark decays before forming a hadronic bound state[3].

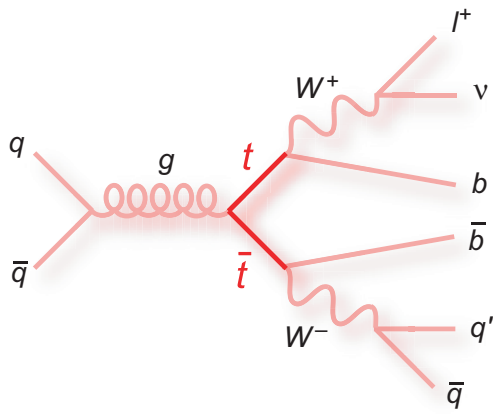

Figure 2.7: The top quark decay in $t\bar{t} \rightarrow lepton(l = e/\mu/\tau) + jets$ channel.

2.4.3 Classification of Top Decay Products

From the calculation of the decay width and the CKM matrix elements, it can be seen that the top quark has a very short lifetime and decays almost entirely into a W boson and a b quark. The process of extraction of the top quark signature in an experiment is based on the W decay mechanisms. For instance, in $t\bar{t}$ pair production when both W bosons decay leptonically, it is defined as the $t\bar{t} \rightarrow$ *dilepton* channel. While if one of the W bosons decays leptonically and the other into jets (explained in Sec. 4.3) it is defined as the $t\bar{t} \rightarrow lepton + jets$ channel. In an event in which both W bosons decay into jets these are defined as the all-jets channel. In summary, an event in which two top quarks are produced should have either

- 2 jets, 2 charged leptons and 2 neutrinos;
- 4 jets, a charged lepton and a neutrino;
- 6 jets.

In all cases, 2 of the jets originate from b quarks. The $t\bar{t} \rightarrow lepton + jets$ channel is further classified into three channels $t\bar{t} \rightarrow e + jets$, $t\bar{t} \rightarrow \mu + jets$ and $t\bar{t} \rightarrow \tau + jets$. Figure 2.7 shows the top decay modes.

[3]Time required for a hadron bound state is about $\sim 10^{-24}$ sec.

Decay channels	Branching
$\boldsymbol{t\bar{t} \rightarrow lepton + jets}$	
$t\bar{t} \rightarrow \tau + jets$	4/27 or 15%
$t\bar{t} \rightarrow e + jets$	4/27 or 15%
$t\bar{t} \rightarrow \mu + jets$	4/27 or 15%
$\boldsymbol{dilepton + jets}$	
$t\bar{t} \rightarrow$ (same lepton) $(\tau\tau/ee/\mu\mu) + jets$	(for each channel) 1/81 or 1.2%
$t\bar{t} \rightarrow$ (different lepton) $(e\tau/\mu\tau/e\mu) + jets$	(for each channel) 2/81 or 2.5%
all-jets	4/9 or 44%

Table 2.3: Branching ratios for different top decay channels.

The branching ratio for different top decay channels can be calculated by using the branching ratio for W decay modes ($W \rightarrow l\nu_l$ (for each lepton types) = 1/9, $l = \tau/e/\mu$ and $W \rightarrow q\bar{q} = 2/3$).

$$BR(t\bar{t} \rightarrow \tau + jets) = 2(\text{for two top quarks})) \times \frac{1}{9} \times \frac{2}{3}$$

The branching ratios for different top decay channels are given in Table 2.3.

Chapter 3

Experimental Apparatus

The results published in this thesis have been attained by using two major high energy physics apparatus: 1. the high energy particle accelerator (Tevatron) and 2. the high energy particle detector (DØ) both of which are located at the Fermi National Accelerator Laboratory, in Batavia, Illinois. The accelerator is used to accelerate, guide and collide the protons and antiproton beams while the detector is used to study the products of the collisions.

In March of 2001 both the (DØ) detector and the Tevatron went through major upgrades. The data taken after the upgrades are referred to as DØ RunII data while the data taken prior to the upgrades are called DØ RunI data. The results of our analyses are obtained by using the DØ RunII data.

3.0.4 The Accelerator

Several acceleration devices operating in different energy ranges are used to produce 980 GeV protons and antiprotons producing collisions with a center of mass energy of 1.96 TeV. Figure 3.1 illustrates the layout of Fermilab accelerator complex.

The acceleration chain starts at the Cockroft-Walton[1] negatively charged hydrogen ions (hydrogen gas is heated to add an additional electron to produce H^-) are

[1]This device uses multistage diode/capacitor for generating high DC voltages at relatively low currents. These types of generators are limited to about 1 MeV before they break down.

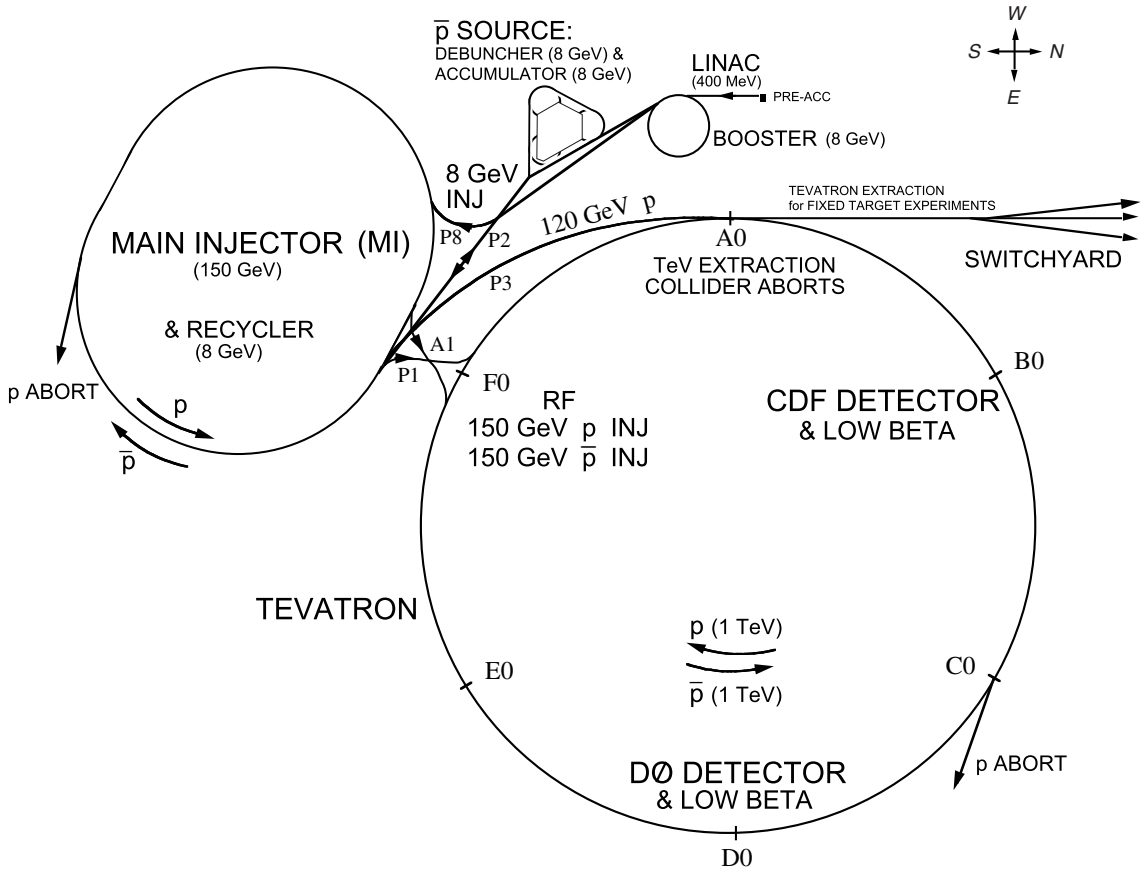

Figure 3.1: The Tevatron

accelerated to 750 KeV through an electric field. The ions from the generator are then transferred to the linear accelerator (Linac) for the second stage of acceleration and for proton extraction.

The Linac consists of a vacuum cylindrical pipe containing a set of metal drift tubes. Fermilab uses a 150 m long Linac to increase the energy of the ions up to 400 MeV. At this stage the ions are passed through a carbon foil, which removes the unnecessary electrons, leaving only the positively charged protons for the next stage of acceleration. The ions are then fed into the Booster, a 75m radius synchrotron accelerator, for the third stage of acceleration. The synchrotron is a cyclic particle accelerator in which the magnetic field is used to turn the particles, and the electric field is used to accelerate the particles.

In the booster the protons are accelerated up to 8 GeV using 18 Radio Frequency (RF) voltage cavities. The 8 GeV protons are then injected into the Main Ring which is a large (1000 m in radius) multi functional synchrotron. In the Main Ring the protons are accelerated up to 150 GeV and then injected to the Tevatron for the final stage of acceleration. The Tevatron uses RF cavities in the same manner as the Booster to accelerate protons to reach 980 GeV and uses magnetic fields to confine the proton beams the central circular orbit.

Besides injecting 150 GeV protons into the Tevatron, the Main Ring also directs 120 GeV protons onto a fixed nickel target for antiproton production. The interaction between the proton and nickel nucleus produces antiprotons (roughly about 10^{-5} antiprotons per incident proton). The antiprotons are then separated from the production target using a lithium magnetic lens that focuses the antiprotons and defocuses the protons. As they are produced, the antiprotons have a spread in their phase space which needs to be minimized. The process of minimization of this spread of antiprotons before injection into the main ring takes place in the Debuncher and Accumulator.

The Debuncher modifies the input antiprotons into a continuous band with a lower momentum spread. It also performs stochastic cooling, in which deviations from the ideal orbit (central beam) are measured and correction signals are applied.

Next, the antiprotons are transferred to the Accumulator which stacks and continuously cools the antiprotons. Antiproton beams from the Accumulator are transferred back to the Main Injector and are accelerated to 150 GeV for the final stage of acceleration at the Tevatron.

The Tevatron then accelerates these antiprotons up to 980 GeV in the direction opposite to the proton beam. Finally, the proton and antiproton beams with total center of mass energy of 1.96 TeV are guided to collide at different detectors locations such as DØ, CDF, etc., located at the Tevatron.

The beams in the Tevatron are discrete. There are 36 bunches of proton and antiprotons $p\bar{p}$ which cross every 396 ns. The amount of interaction/data recorded over time at DØ is described in terms of integrated luminosity (fb^{-1}).

3.1 Top Detection at DØ

The top quark has a very short lifetime, of the order of $10^{-24}s$, therefore its existence can only be detected by analyzing its decay products in the detector. Detectors for high energy particles consist of various particle detection devices. The tracking devices are installed close to the interaction point. At this tracking level a magnetic field is used to bend the motion of the charged particles from the original path. The radius of curvature of the trajectory provides the measurement of the momentum of the charged particles coming from the interaction point. After tracking the particles, energy is measured in a calorimeter. Typically a calorimeter is designed such that it will absorb all the incident particle's energy in a shower except for Muons and Neutrinos. Muons do not deposit much energy in the calorimeter so they are detected outside the calorimeter with a special detector. Neutrinos are not detected directly but are accounted for by the imbalance left in the total detected momentum transverse to the beam.

The next sections describe different components and functions of the DØ detector and the chapter that follows explains how different physics objects such as electrons, muons, taus, quark jets, $\not\!\!E_T$ (Missing Transverse Energy) are identified from the measured quantities at the detector.

3.1.1 The DØ Detector

The DØ has a donut shaped cylindrical structure, the beamline passes through the axis of the detector and the collision takes place close to the center of the cylindrical volume.

The detector has three major detection components. The components are Central Tracking Detectors, Calorimeter, and Muon Detector. Only the Central Tracking Detectors and Calorimeter are described in the following sections. A rigorous description of the detector can be found in Ref. [17].

In order to describe the collision point and physics objects, the DØ detector uses a right handed Cartesian coordinate system (x, y, z) with the positive z-axis is defined along the proton direction (south), the positive y-axis is defined upward, and the positive x-axis points east. As $p\bar{p}$ collisions are boosted along the z-axis, the detector uses a pseudo-spherical coordinate system in which the polar angle θ is replaced with the pseudorapidity η to define the froward and central regions. The coordinate η is defined as follows:

$$\eta = -ln\left[tan\left(\frac{\theta}{2}\right)\right], \tag{3.1}$$

which approximate the true rapidity $y = \frac{1}{2}ln\left[\left(E + p_z c\right) / \left(E - p_z c\right)\right]$ for finite angles in the limit $(mc^2/E) \rightarrow 0$.

3.1.2 Tracking Detectors

There are two central tracking detectors in the DØ detector, the silicon microstrip tracker (SMT) and the central fiber tracker (CFT) located close to the primary interaction point and surrounded by a 2 tesla magnetic field generated by a super conducting solenoid magnet. These detectors are used to identify the location of the primary interaction vertex, to measure charged particle momentum, and to identify of b-quark decay vertex.

The main building blocks of the SMT are doped silicon based semiconductor devices (typically 50 μm in pitch) which operate under reverse bias voltage. When an energetic charged particle enters the device, it creates electron-hole pairs. Under the influence of the bias voltage, the charge drifts towards the sensors and sends signal to the readout system.

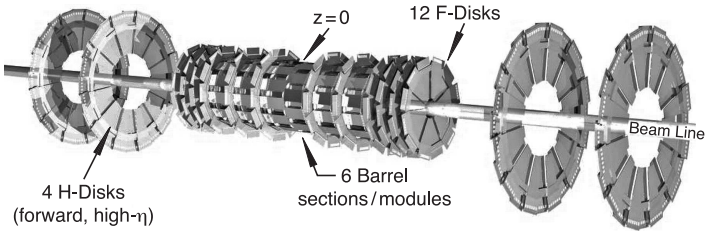

Figure 3.2: The Silicon Microstrip Tracker.

The SMT is arranged in such a way that it maximizes the efficiency of the tracking with a given number of silicon devices. There are 6 barrel modules concentric to the $p\bar{p}$ beam line, each with 4 layers of silicon detectors to measure positions of charged particles in the transverse direction. Figure 3.2 shows the schematic of the silicon microstrip tracker which provides tracking coverage for $|\eta| < 3$ and Table 3.1 shows the specifications of the SMT.

There are 12 F disk assemblies perpendicular to the beam line, 6 of which are interspersed by the barrel module to measure positions of centrally produced charged particles. The barrels provide coverage for $|\eta| < 1.1$.

The SMT has 4 H disk modules in order to detect the charged particles in both longitudinal and transverse directions for high η regions.

	Barrels	F-Disks	H-Disks
Si area	1.3 m^2	0.4 m^2	1.3 m^2
Inner radius	2.7 cm	2.6 cm	9.4 cm
Outer radius	9.4 cm	10.5 cm	26.0 cm
Max $\|Z\|$	38.4 cm	54.8 cm	120.0 cm

Table 3.1: Physical attributes of the SMT.

The CFT is located immediately outside the SMT and is organized into 8 layers of scintillating fibers made out of organic scintillating compounds surrounded by a thin layer of cladding designed to provide total internal reflection inside the fiber. Each layer of fibers is oriented in two directions one parallel to the beam line and the other rotated $\pm 3^0$ with respect to the beam line.

When a charged particle interacts with the scintillating material of the fiber, it emits photons which are sent via wave guide to the visible light photon counters (VLPC) located outside the detector.

3.1.3 Calorimeter and Inter-Cryostat Detector

A calorimeter is a device that measures energy. When an electron enters the calorimeter with several GeV of energy, it radiates its energy by emitting an energetic photon while passing through the calorimeter material (also known as Bremsstrahlung). These energetic photons then initiate electron-positron pair production and thus create a shower of secondary electrons and photons of lower energy. Similarly, an incident photon with several GeV of energy initially produces an electron positron pair and finally produces a shower of increasingly lower energy electrons and photons. The energy loss of these electromagnetic particles in passing through a specified amount of material is determined by the material's radiation length X_o. A radiation length is related to the energy of the incident particle which can be illustrated by the following equation,

$$\langle E \rangle = E_o e^{-x/X_o}. \tag{3.2}$$

Where, $\langle E \rangle$ represents the mean energy of incident beam as a function of position x and E_o is the energy of the incident beam. Thus the radiation length X_o is defined as the thickness of the medium that reduces the mean energy of a beam by $1/e$.

The principles of hadronic shower energy measurements are similar to those of the electromagnetic. In this case the incident particle collides inelastically with a nucleus in the absorber medium, producing a number of secondary hadrons that

may also collide inelastically and produce a shower. The hadronic shower mostly contains Pions and a small fraction of Kaons, Lambdas, Deltas etc. The hadronic shower also contains a EM shower component because π^os decay to photons. A measurement of a hadronic shower is given in terms of the nuclear interaction length λ for the material. Starting from the interaction point, the hadronic shower penetrates deeper inside the calorimeter. So, the layers of the calorimeter closest to the interaction point are optimized for measurement of EM shower and the layers outside the EM are optimized for hadronic shower measurement.

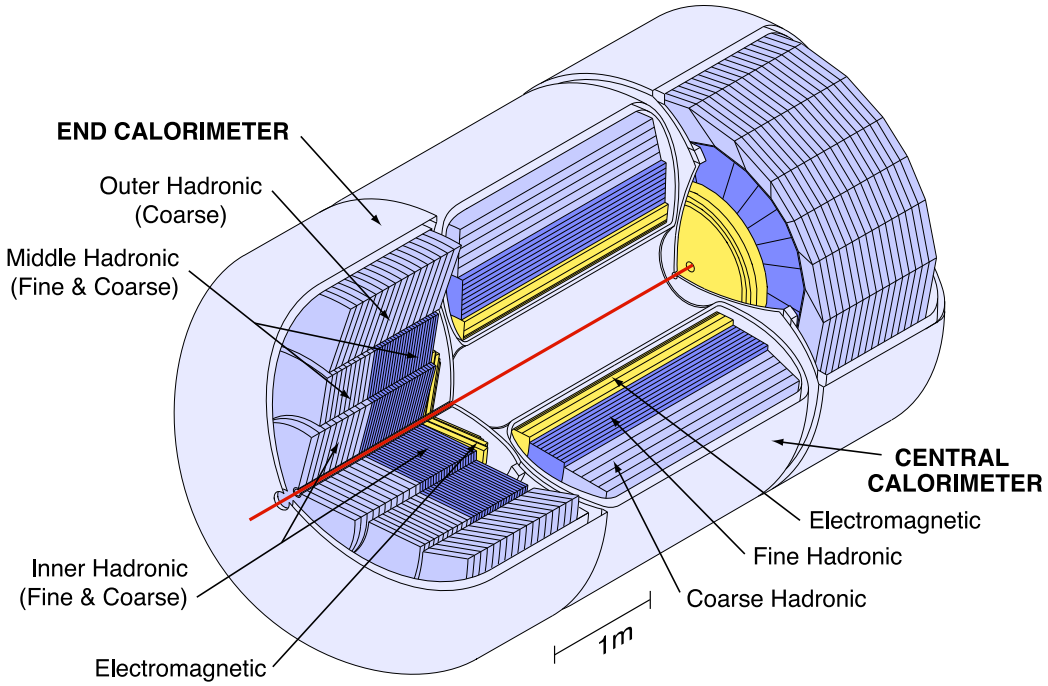

Figure 3.3: The DØ calorimeter.

The DØ calorimeter, shown in Fig. 3.3, consists of layered structure of alternating denser and lighter material. Uranium or copper or stainless steel is used as the absorber material while argon is used as the ionization medium. The cells consists of an absorber plate and a signal board and the gap between the two is filled with liquid argon. The absorber plate is grounded and a positive voltage of 2 KeV – 2.5 KeV is applied to the signal board. When a particle enters the calorimeter it showers inside the absorber plate and the surviving particles from the absorber

ionize the argon atoms. The electrons drift towards the signal board under the applied field producing a current in the unit cell and the ionized atoms drift in the opposite direction. This current induces charges on the copper signal board and are sent to the electrical readout for recording.

The calorimeter has three units : The central calorimeter and two End-cap calorimeters. The central calorimeter consists of three concentric cylindrical layers. The innermost layers measure the EM shower energy. The thickness of the EM modules is optimized [2] to absorb total energy of an electromagnetic shower. The layers outside the EM calorimeter are designed for hadronic shower measurements. The thickness of hadronic module is also optimized to absorb the total energy of an hadronic shower. The hadronic module is segmented in two sections. The fine hadronic module [3] (FH) lie immediately after the EM module and course hadronic module[4] (CH) sits right after FH.

3.1.4 Trigger and Data Acquisition System

The goal of the trigger system is to record events that have the potential for improving the Standard Model results and is capable of finding the Standard Model Higgs particle and also new physics beyond the Standard Model. This section describes how the DØ detector filters and records the products of $p\bar{p}$ collisions using information from different components of the detector.

There are three independent trigger levels at DØ and at each level the data flow rate is reduced such that the final rate is 50 Hz, the maximum acceptance rate required by the DØ collaboration. As $p\bar{p}$ collide every 396 seconds, the level 1 has very little time to reconstruct physics objects like electrons, muons, taus, jets etc. and make trigger a decision. For this time constraint it uses energy deposited in the calorimeter trigger towers (all the calorimeter cells within $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ space) and number of hits at CFT or Muon system to achieve two important goals:

[2]For an EM module the Uranium absorber plate thickness is 3-4 mm

[3]It has 6 mm thick Uranium-niobium absorber plate and provides detailed information of shower shape.

[4]A copper plate of thickness 46.5 mm is used as absorber. Its primary purpose is to absorb the entire energy of a hadronic shower

1) accept events that are interesting and 2) keep the acceptance rate below 1800 Hz.

In order to achieve these two goals, a set of optimized requirements are implemented at level 1. For instance, a multijet trigger (v14.0-v15.0 & JT2_4JT20 in Table 5.2) that was used to record data for this result tells the level 1 trigger system to select an event and pass it for a level 2 decision if there are at least three level 1 calorimeter trigger towers with $E_T > 5$ GeV, at least three level 1 calorimeter trigger towers with $E_T > 4$ GeV and $|\eta| < 2.6$ and at least one level 1 calorimeter trigger towers with $E_T > 7$ GeV and $|\eta| < 1.8$ in a $p\bar{p}$ event.

The level 2 trigger system collects information both from the level 1 trigger and from different components of the detector. Unlike level 1, level 2 uses sub-detector specific processing nodes and a global node to perform simple physics object reconstruction such as electrons, jets, $\not\!\!E_T$, H_T (the scalar P_T sum of all jets in an event), as well as make trigger decisions based on a set of optimized conditions. For example, the multijet trigger (v14.0-v15.0 & JT2_4JT20 in Table 5.2), an event can pass the level 1 trigger condition and if the event has at least three level 2 jets with $E_T > 6$ GeV and $H_T > 75$ GeV (jets with $E_T > 6$ GeV and $|\eta| < 2.6$ then the event is sent to the level 3 trigger decision. The maximum allowed data rate at this level is 850 Hz.

The level 3 trigger is a completely software based fast algorithm. It has two parts level 3 data acquisition system (DAQ) and level 3 filters. The DAQ uses a collection of processing nodes to completely build physics objects and sends it to a level 3 filter which makes the final trigger decisions on an event. Depending on the physics objects, level 3 applies more sophisticated criteria such as object isolation, matching, neural network cuts etc. in addition to standard kinematic selections. As the objects are more visible, number of objects, kinematic thresholds and $|\eta|$ cuts are usually higher at this level. For instance, the conditions for an event to pass the level 3 trigger (v14.0-v15.0 & JT2_4JT20 in Table 5.2) are at least four level 3 jets with $E_T > 10$ GeV and $|\eta| < 3.6$, out of the four jets at least two level 3 jets with $E_T > 20$ GeV and $|\eta| < 3.6$.

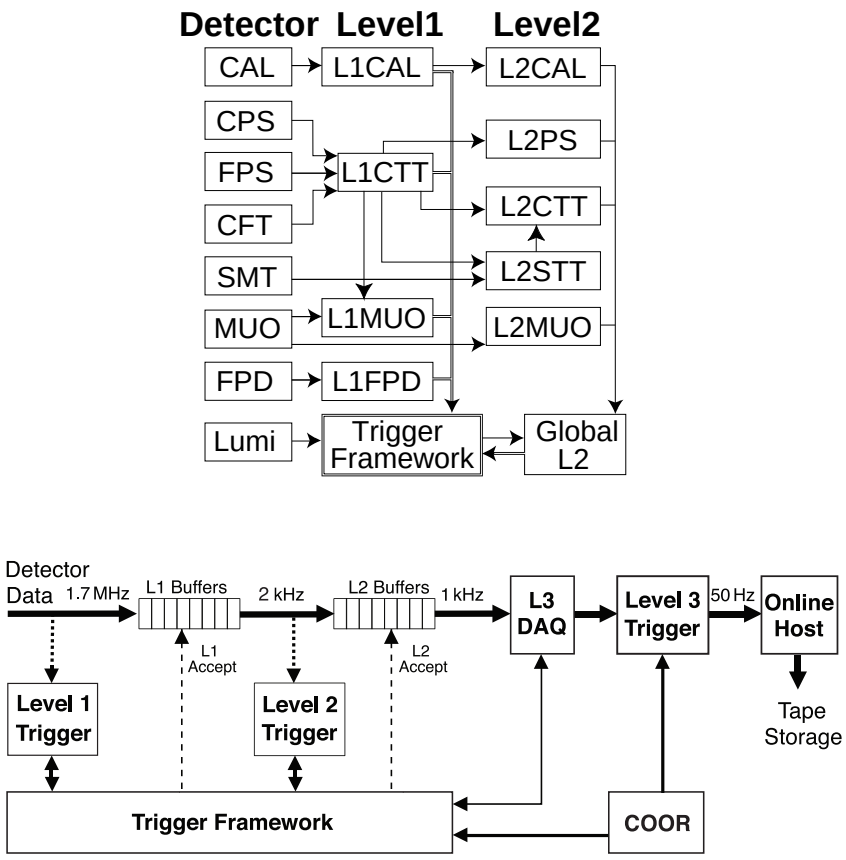

Figure 3.4: The DØ trigger system.

If an event passes all three levels of the trigger then the event is sent to Feynman Computing center to be stored as raw data. The next chapter is devoted to reconstruction and identification of physics events and objects from the raw data. Figure 3.4 shows a schematic of the combined trigger system.

Chapter 4

Object Identification And Optimization

This analysis uses $t\bar{t}$ decay products such as taus, jets, b-jets, $\not\!\!E_T$ and their properties to separate $t\bar{t} \rightarrow \tau + jets$ events from multijet/QCD events. Therefore, only the reconstruction of these decay products are discussed in this chapter. Since electrons and muons are rejected (vetoed), their reconstruction algorithms is not discussed here [1]. The track and vertex information is shared among all the object reconstruction procedures, and are briefly discussed.

The raw data collected by the acquisition system are processed offline by a set of software algorithms called D0reco [31]. This data primarily contains digitized pulse-heights resulting from accumulated charge by the SMT or calorimeter, the photon counts in CFT scintillators, and the time differences from the muon drift chambers.

The algorithms first unpack the raw data and use information from different components of the detector to determine charged particle tracks. The algorithms then construct useful physics objects such as electrons, muons, taus, jets, $\not\!\!E_T$ and vertices. Along with the reconstruction procedure, the algorithms also apply a set of optimization techniques and selection criteria to improve the purity and efficiency of the physics objects.

An equivalent reconstruction algorithm is also used in reconstruction of Monte Carlo events so that they look like or become representative of data events.

[1] A good explanation of electron and muon reconstruction can be found in the Ref. [18, 30].

4.1 Track Reconstruction

A track object is defined by the path of a charged particle in three dimensional space. To reconstruct tracks, D0reco first uses the Track Hit Clustering (THC) algorithm. The THC algorithm starts clustering hits in the SMT if a silicon strip registers charge accumulation above a threshold due to the ionization by a charged particle from $p\bar{p}$ collisions. The THC algorithm then looks for other hits adjacent to the first one. If it finds more than one adjacent hit, the hits are combined together into one. In case of the CFT, a hit is reconstructed if two fibers in each super layer register photon emissions due to the interaction between a charge particle and the fibers. After the hit reconstruction, two complementary algorithms: Alternative Algorithm (AA) [33] and Histogramming Track Finding Method (HTF) [32] are utilized to link the hits together and create charged particle tracks.

The Alternative Algorithm selects a pool of track candidates by selecting a hit cluster in a layer of the tracking detector and by constructing a track and incrementally adding hits from more layers of the SMT and CFT detectors. A set of well studied selection criteria is applied to the pool of track candidates to select the final tracks for later analysis. The selection criteria are as follows :

- track radius of curvature is greater than 30 cm (i.e., $P_T > 180$ GeV)
- χ^2 fit is less than 16

The HTF algorithm transforms (x-y) hit clusters from the SMT or CFT into another plane defined by radius of curvature ρ $(= qB/P_T$, where q and P_T are the charge and transverse[2] momentum of the particle, and B is the magnetic field) and azimuthal angle (ρ, ϕ) and fills histograms. In an event all the hits from the same particle will produce a peak in the (ρ, ϕ) plane. The peak is then parametrized to define a set of candidate tracks. A set of selection criteria and a two-dimensional Kalman [34] filter which can solve mathematical problems of determining optimal track parameters with errors from a set of measurements, is also used to improve

[2]To the magnetic field.

the quality of the tracks. The track information is then shared to reconstruct physics objects that can be associated with a track(s).

4.2 Primary Vertex Reconstruction

The main task of the vertex reconstruction algorithm [35] is to locate the origin of tracks in $p\bar{p}$ hard-scatter interaction. The algorithm starts selecting tracks with $P_T > 0.5$ GeV and at least two SMT hits. The tracks are then clustered in the z-direction in such a way that the z-distance between the track and the cluster is less than 2 cm. After clustering the tracks, vertex fitting is performed for each of the z-clusters in a 2-pass approach:

In the first pass, the Kalman Filter vertex fitting algorithm is used to determine the location and width of the beam by fitting all selected tracks within each z-cluster into a common vertex.

In the second pass, tracks in each z-cluster are first pre-selected according to their distance of closest approach to the beam spot.

Finally, the algorithm computes the probability of a vertex to be a minimum bias vertex as a function of vertex track P_T distribution. The probability is shown in Eq. 4.1

$$P(P_T) = \frac{\int_{\log_{10}(P_T)}^{\infty} F(x)dx}{\int_{\log_{10}(0.5)}^{\infty} F(x)dx} \tag{4.1}$$

where, $F(x)$ is the distribution of $x = log_{10}(P_T)$ for tracks from minimum bias events.

As hard scatter vertices show a harder P_T spectrum for the associated tracks, the probability for each primary vertex is estimated such that it corresponds to a minimum bias vertex (a vertex from an inelastic collision) in $p\bar{p}$ collisions and not to a hard scatter vertex. The vertex which has the smallest probability to be a minimum bias vertex is selected as the primary vertex [36, 37].

4.3 Jet Reconstruction

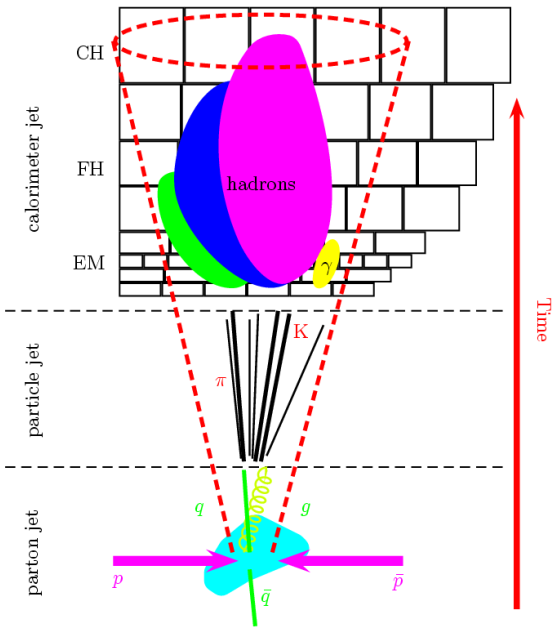

Figure 4.1: Cartoon of different stages of a jet formed by a hard scattering event.

A jet is defined as a spray of particles that fly out in a relatively narrow cone from $p\bar{p}$ collisions. As they travel away from the collision point, they emit gluons, which split into even more gluons resulting in a relatively narrow cascade of particles also known as parton-jets. In the last stage of jet production, partons hadronize to form hadrons such as protons, pions, and kaons. At this stage jets are labeled as particle-jets (before entering the detector). These hadrons will deposit most of their energy in the layers of cells of the calorimeter part of the detector and become calorimeter jets. However, charged hadrons can deposit very little energy in the central tracker which are identified as track-jets. The main goal of the jet reconstruction algorithm is to properly measure the jet energy and direction in order to determine the original parton energy and momentum.

In RunII, DØ uses the legacy jet reconstruction cone algorithm [38]. The jets are reconstructed in the following two steps:

First the algorithm uses a noise suppression procedure and selects a list of calorimeter cells containing energy at least $2.5\sigma_{ped}$ (the rms value of the calorimeter noise) but less than $4\sigma_{ped}$[41]. The algorithm also uses NADA (New Anomalous Deposit Algorithm) [39] to calculate the total energy contained in the neighboring cells and removes them according to the energy threshold cuts listed in the Ref. [40]. At this stage the algorithm constructs towers[3] in $(\eta - \phi)$ space with cells that passed the noise suppression scheme and calculates the four-vector (E, p_x, p_y, p_z) for every tower. For instance, a tower four-vector is calculated by summing up four-momenta for all the selected/good cells in a tower.

Second, the cone algorithm makes a list of all towers with transverse momenta $(E_T = Esin(\theta))$ greater than 0.5 GeV and then defines "seed towers" (cluster of towers) by requiring the sum of transverse energy to be greater than 1 GeV within a cone of $\Delta R = 0.3$. The algorithm then clusters calorimeter towers in a cone of $R = 0.5$ around the seed tower and defines this as the jet candidate or proto-jets if it has a sum of tower transverse energy greater than 6 GeV. The axis of the jet is assigned by the E_T weighted midpoints of each calorimeter tower. This procedure is repeated throughout the detector until the position of jet central axis does not change (i.e forms a stable jet cone axis) from one iteration to another.

4.3.1 Jet Merging And Splitting

The jet found in the clustering algorithm can share a common tower energy. In the final stage, the algorithm merge or split jets based on a fraction of the tower transverse energy shared between two adjacent jets. The jets are merged or split depending on the following criteria:

- Two jets are merged if their overlapping energy in a tower is *more* than 50% of the individual jet energies.
- Two jets are merged if their overlapping energy in a tower is *less* than 50% of the individual jet energies.

[3] A calorimeter tower represents stacks of cells that have the same η and ϕ.

The merging and splitting algorithms are repeated until all the jet candidates have stable axes.

4.3.2 Jet Quality

After the jet reconstruction, a set of quality criteria are applied to reduce the number of fake jets.

- To remove fake jets created by electromagnetic particles such as photons or electrons, all jets are required to have between 5% and 95% of their energy deposited in the hadronic layers of the calorimeter and isolated from the electromagnetic clusters of the calorimeter by $\Delta R > 0.5$.
- To remove fake jets created by calorimeter noise, jets are required to have more than 60% of their energy in the fine hadronic section of the calorimeter.
- The fraction of the jet energy deposited in the coarse hadronic layer of the calorimeter must be less than 40% as this part of the detector is more sensitive to noise due to large readout system. In terms of detector design it is unlikely that jets will deposit most of their energy in this part of the detector.
- Electrons, photons and noisy towers tend to deposit most of the energy within a single tower. Jets created by such processes are removed by removing single calorimeter towers that contain more than 90% of the jet energy.
- The ratio of the total energy in Level 1 trigger towers associated with a jet to the reconstructed energy of a jet must be at least 0.4. This criteria is used to remove the noise from the precision readout system.

4.3.3 Jet Energy Corrections

The jet energy measured after the quality cuts certainly differ from the original particle-jet energy. The difference can be accounted for by three major factors:

- The presence of additional absorbing materials in front of the calorimeter

- Various gaps and cracks in the calorimeter and
- Fixed size jets cone which may not always include all of the constituent particles that originated from particle-jets

So the aim of the jet energy correction is to apply a correction factor to the reconstructed jets after the jet quality cuts such that their energy match with the parton energy. The corrections are then applied to both the data and the Monte Carlo jets. Details of the energy corrections can be found in the Ref. [44, 45, 46]. The corrected energy of a jet is given by

$$E_{corrected} = \frac{E_{uncorrected} - O}{F_\eta \times S}, \tag{4.2}$$

where $E_{corrected}$ is the energy of the particle-jet, $E_{uncorrected}$ is the energy found by the jet cone algorithm, O is the offset energy correction, F_η is the relative response correction, and S is the showering correction. Description of all the terms are summarized below:

The jet energy calculated by the cone algorithm is offset by energy contributions from the electronic noise, noise due to uranium decay, energy from the previous collisions known as pile-up effect, multiple interactions in the same beam, and from underlying events[4]. Therefore, a sum of all offset energy contribution is calculated using minimum bias events[5] as a function of η and for different primary vertices (Fig. 4.2) and subtracted from the $E_{uncorrected}$ Eq. (4.2) for the jet offset energy correction.

The DØ calorimeter is uniform within the central calorimeter (CC) and end calorimeter (EC) cryostat(s). The gap between both cryostats ($0.5 < |\eta| < 1.8$) is not as well instrumented, causing a non-uniformity in response as a function of pseudo-rapidity and energy loss in the cracks between calorimeter modules. These effects are modeled by the relative response correction factor F_η. Measurement of this

[4]An underlying event is an event in which there is an elastic interaction or a quasi elastic interaction.

[5]A minimum bias event contains energy from the underlying events. The events are recorded by a minimum bias trigger which is defined by the hit on luminosity counters located on both sides of the point of interaction.

correction is done using Missing Transverse Energy Projection Fraction (MPF) technique (Fig. 4.3) on $\gamma + jets$ and dijet data events. The MPF tags photon/jet in the central calorimeter and probes the jets. It then calculates momentum imbalance in the transverse plane (Missing Transverse Energy or $\not\!\!E_T$ projected in the tag object direction (Fig.4.3). This correction is determined as a function of tag object transverse momentum for $|\eta| < 2.5$ (Fig. 4.5).

The showering term S, either corrects for missed energy outside the jet cone size or deposits additional unwanted energy inside the cone. The showering correction only corrects for detector instrumental effects. Both data and parton level Monte Carlo events (without detector simulation) are utilized to measure this quantity. For instance, in data events, the energy measurement is affected by both physics and detector effects, while in Monte Carlo events the energy measurement represents physics effects only. By taking the ratio between the effects of physics+detector and physics, the showering correction for the detector is determined. Figure 4.4 shows the showering correction as function of jet transverse energy for central and forward regions of the calorimeter.

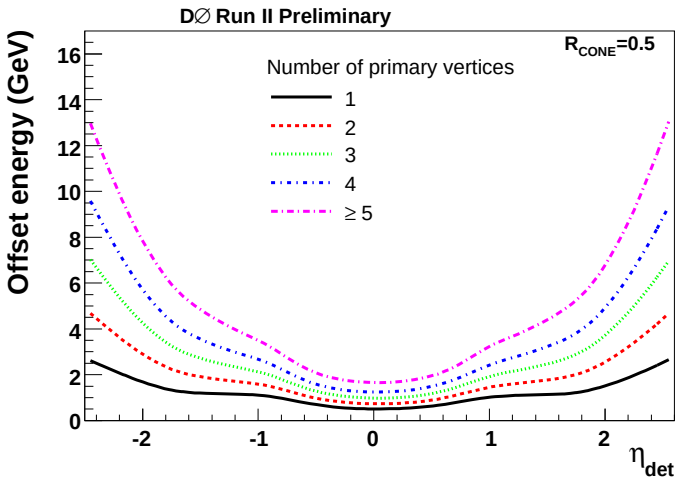

Figure 4.2: Plot of offset energy as a function of pesudorapidity for different primary vertices.

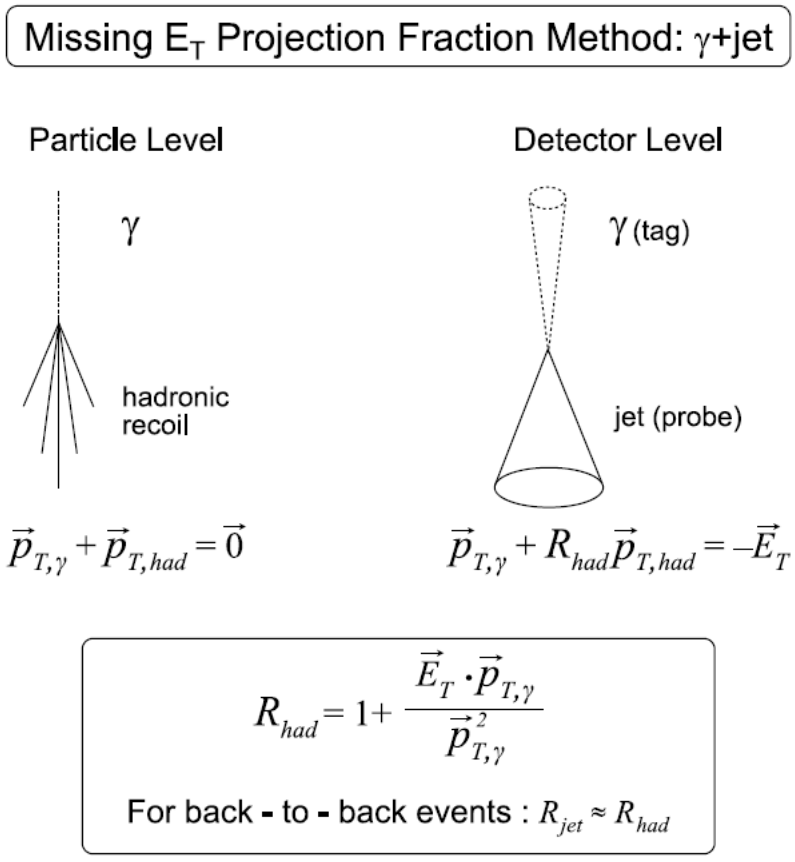

Figure 4.3: Illustration of the MPF method used in the relative response correction.

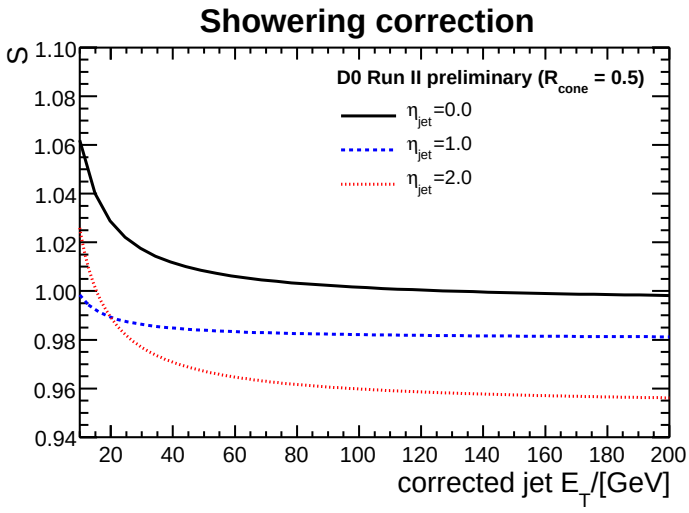

Figure 4.4: Showering correction for jets in data as a function pesudorapidity of corrected jet transverse energy.

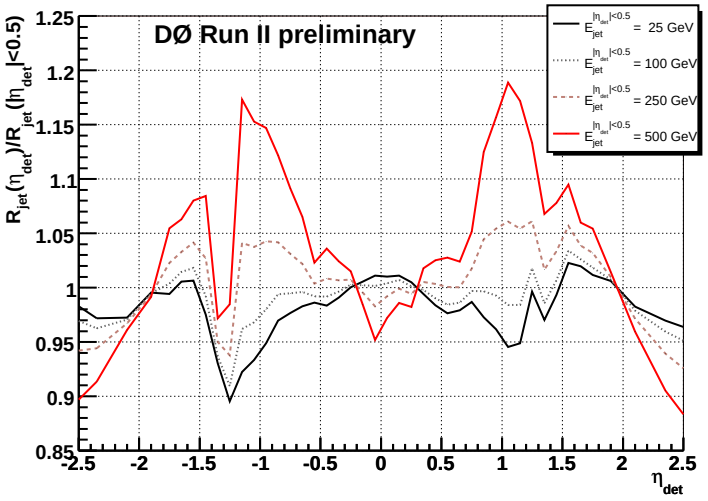

Figure 4.5: Relative response correction in data as function of pesudorapidity.

4.3.4 Muon correction for Jets

If a muon is found inside a jet cone, the jet energy needs to be corrected for that muon. The muon correction for jets uses the momentum of the muon from the muon system and tracker, and add the energy of the muon from the jet energy. This correction is known as JESMU and is used the analysis.

4.3.5 Jet Energy Resolution

Jet energy resolution is applied to both data and Monte Carlo to improve jet energy measurements [49, 48]. Data events from $\gamma + jet$ and dijet processes are used to estimate the jet energy resolution parameter. For jets with $P_T > 50$ GeV, dijets samples are used and a symmetry variable is defined by jet P_Ts as follows:

$$A_{P_T^{jet}>50} = \frac{|P_T^{jet1} - P_T^{jet2}|}{P_T^{jet1} + P_T^{jet2}}, \tag{4.3}$$

where P_T^{jet1} and P_T^{jet2} are the transverse momentum of two jets. The jet resolution for a jet with $P_T > 50$ GeV is given by the following relation:

$$JER_{P_T^{jet}>50} = \frac{\sigma_{P_T}^{jet}}{P_T^{jet}} = \sqrt{2}\sigma_A, \tag{4.4}$$

where σ_{P_T} is the uncertainty of jet P_T measurement and σ_A is the width of $A_{P_T^{jet}>50}$ distribution obtained by a Gaussian fit.

A similar technique is applied to determine jet energy resolution for jets with $P_T < 50$ using $\gamma + jet$ data events. Finally, the jet resolution is found using both techniques together to be finally applied to the jets.

4.3.6 b-jet Tagging

The $t\bar{t}$ final state contains two b-jets, while jets in QCD and W+jets events originate most often from light quarks or gluons. Requiring at least one jet in an event to be $b-$tagged is a very powerful method for background rejection.

The goal of the b-jet tagging algorithm is to identify b-jets originating from b-quarks. The b-quarks hadronize into B-hadrons (a bound state of a b quark and one or two light quarks) which have longer lifetime compared to lighter hadrons. Due to this long lifetime a B-hadron has a displaced decay vertex (secondary vertex) from the primary interaction vertex (PV) (Fig. 4.6, 4.7). The typical decay length, which is the distance from the primary to the secondary vertex, is a few millimeters. Using the property of displaced vertex and tracks from decaying B-hadron, the tagging algorithm estimates the probability of a jet originating from a b-quark.

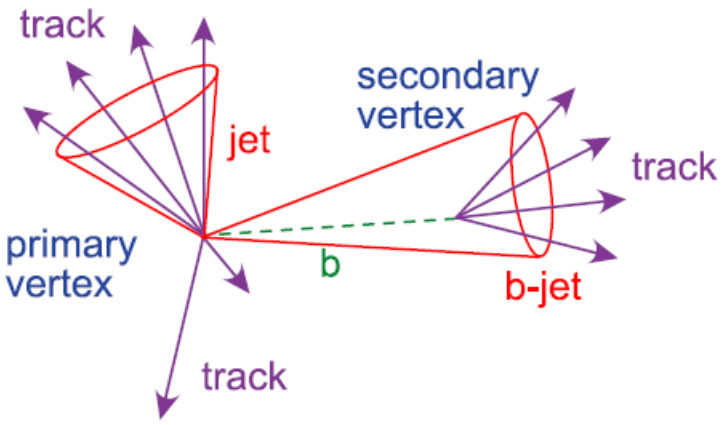

Figure 4.6: Illustration of secondary vertex in a jet cone [19].

The $b-$tagging algorithm used in this measurement is a Neural Network[6] tagging algorithm ($NNbtag$), which combines seven characteristic variables based on properties of secondary vertices within the jets, impact parameter significance, probability of the jet to originate from the primary vertex etc. listed in Table 4.1. The network is trained on $Z \rightarrow b\bar{b}$ Monte Carlo events and the output of this is a discriminant with value between 0 and 1, where most of the b-jets have a $NNbtag$ value close to 1. Therefore, a $NNbtag$ cut is applied to select b-jets i.e. a jet is tagged as a b-jet if the $NNbtag$ value is greater than 0.775 (also known as the tight operating point).

[6] A good explanation of Neural Network Training can be found in the Ref. [20]

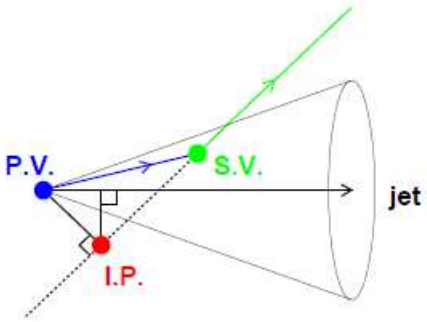

Figure 4.7: Illustration of impact parameter [19].

Rank	**Short Description**
1	Decay Length Significance of the Secondary Vertex (SV)
2	Weighted combination of the track's impact parameter significance
3	Probability that the jet originates from the primary vertex (PV)
4	Chi Square per degree of freedom of the SV
5	Number of tracks used to reconstruct the SV
6	Mass of the SV
7	Number of SV found in the jet

Table 4.1: NNbtag input variables ranked in order of power.

For data events, the tagging algorithm is applied directly to jets in the events. For Monte Carlo events, the tagging algorithm is not applied directly but rather the algorithm assigns a probability value to each jet in an event that may have originated from a b-jet. This probability value is measured from data and is often referred to as the tag rate function (TRF). For example, for n jets in an event, the probability of having at least one b-tagged jet is given by $P_{event} = 1 - \prod_{j=1}^{n}(1 - P_j)$, where, P_j is TRF-derived probability to tag a jet.

The reason for treating $b-$jets differently in Monte Carlo events is that when the tagging algorithm is applied directly to Monte Carlo events, it gives an efficiency that is higher than data [50]. In order to account for this factor, the TRF has been parameterized on $t\bar{t} \rightarrow \mu + jets$ Monte Carlo and $t\bar{t} \rightarrow \mu + jets$ data. The $t\bar{t}$ Monte Carlo sample is chosen over $Z \rightarrow \mu\mu$ Monte Carlo sample in order to optimize the TRF for the top quark related physics.

The algorithm can only be applied to the jets that contain tracks. Such jets are called "taggable" jets. Taggability of a jet is different for data and Monte Carlo. This difference is due to imperfect simulation of the tracking system in the Monte Carlo. To account for this difference, taggability is measured and parameterized in data and is applied to Monte Carlo events as a weight. This taggability weight is applied on top of the TRF probability in an event [51].

4.4 Tau Candidates Reconstruction

4.4.1 Tau Identification

The tau lepton is the heaviest particle among all the leptons in the SM and has a lifetime of $\sim 290.0 \times 10^{-15}$ *sec*. The tau decay modes can be grouped into two flavors: leptonic (e/μ) and hadronic (Table 4.2 summaries the dominant decay modes of tau leptons).

The decay length of the tau is ~ 87 μm which is beyond the SVT tagging resolution of the DØ detector making it difficult to separate the secondary tau vertices from

the primary one. Therefore, the SVT algorithm cannot be used to identify tau leptons.

At the DØ detector, only the hadronic tau decay modes are used for identification as it is almost impossible to separate leptonic decay modes of a tau from the leptonic decay modes of its parent particle. For instance, in a data event an electron from $W \rightarrow \tau\nu_\tau \rightarrow e\nu_e$ process may have similar kinematic properties as that of an electron from $W \rightarrow e\nu_e$ which makes it harder to distinguish one from the other. Therefore, only the hadronic tau identification is discussed in this section.

Leptonic Decay Modes	**Branching ratios**
$\tau \rightarrow e\bar{\nu_e}\nu_\tau$	$17.85 \pm 0.05\%$
$\tau \rightarrow \mu\bar{\nu_\mu}\nu_\tau$	$17.36 \pm 0.05\%$
Hadronic Decay Modes	
Type I (one prong decay)	
$\tau \rightarrow \pi^-\nu_\tau$	$10.91 \pm 0.07\%$
Type II (one prong decay)	
$\tau \rightarrow \pi^-\pi^0\nu_\tau$	$25.51 \pm 0.09\%$
$\tau \rightarrow \pi^-2\pi^0\nu_\tau$	$9.29 \pm 0.11\%$
$\tau \rightarrow \pi^-3\pi^0\nu_\tau$	$1.04 \pm 0.07\%$
Type III (three prong decay)	
$\tau \rightarrow \pi^-\pi^+\pi^-\nu_\tau$	$9.32 \pm 0.07\%$
$\tau \rightarrow \pi^-\pi^+\pi^-\pi^0\nu_\tau$	$4.61 \pm 0.06\%$

Table 4.2: Branching ratios for dominant tau decay modes.

The jets from quarks or gluons are the major background for hadronic taus in that these jets can mimic taus in the calorimeter. It can be seen from Table 4.2 that hadronic taus predominantly decay into pions and these pions (only charge pions leave tracks in the tracking detector) deposit energy in the electromagnetic and hadronic layers of the calorimeter similar to jets.

However, hadronic taus can be distinguished from jets by means of energy deposited in the jet cone, number of tracks, track energy etc. For instance, hadronic taus tend to be energetic and confined in a narrower cone than jets as they are usually a product of a W or Z heavy particle decay and are boosted along the boost

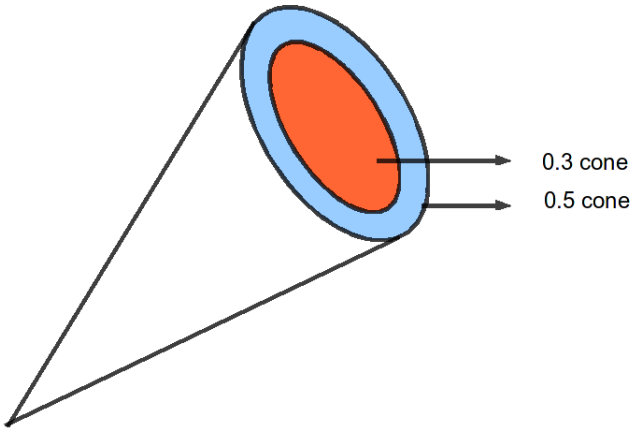

Figure 4.8: Illustration of tau reconstruction cones.

directions of the parent particles. In terms of the number of tracks, hadronic taus have fewer number of tracks $(1-3)$ in comparison to jets. The maximum number of tracks that can be attached to a tau is three and the minimum is one. If additional tracks are found that are also matched to a tau, the sum of the energy of additional (extra) track is expected to be small and for a jet it is large. These are just a few examples to show that it is possible to distinguish taus from jets. In practice, in a tau reconstruction procedure, a number of discriminating variables (Sec. 4.4.2) are defined to find isolated calorimeter energy clusters that have associated track(s) and at the end tau Neural Networking Training (NN_τ) is used to separate taus from jets.

The tau reconstruction and optimization algorithm can be explained in the following five steps:

1. Building tau clusters in the calorimeter
2. Building electromagnetic sub cluster
3. Matching track(s) to tau candidates
4. Defining tau types
5. Separating taus from the background

4.4.1.1 Building tau clusters in the calorimeter

The list of hadronic tau clusters are built using the same jet cone algorithm used in the jet reconstruction algorithm (Sec. 4.3). The tau clusters are built with calorimeter towers around a seed tower within a cone of $\Delta R = 0.3$ with transverse energy $E^{\tau}_{Tcore} > 4$ GeV, and requiring shower shape (defined in tau NN_{τ} variables) (rms) to be less than 0.25. An isolation cone of $\Delta R = 0.5$ around the tau core cone with $E^{\tau}_{T} > 5$ GeV is also defined.

4.4.1.2 Building electromagnetic sub cluster

The algorithm forms the electromagnetic sub clusters (third layer of the electromagnetic calorimeter) by the neutral pions π^0 from taus decay into two photons and deposit energy in the electromagnetic section of the calorimeter. To form electromagnetic sub clusters the nearest-neighbor algorithm is utilized. The way the nearest-neighbor (CellNN) algorithm works is, it reconstructs calorimeter in floor-level clusters using four neighboring cells, (cells with an edge in common are neighbors) each cell with an energy threshold of 200 MeV (Fig. 4.9). The reconstructed floor-clusters are then combined into global-clusters. A detailed description of the CellNN algorithm is in Ref. [55]. If such a cluster is found, then it is attached to EM cells in other layers and preshower hits. The EM subcluster energy is used to define tau NN_{τ} variables.

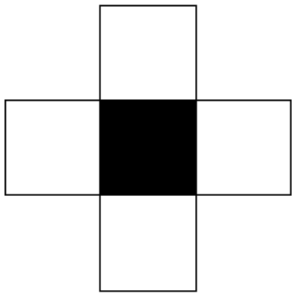

Figure 4.9: Floor-level EM neighboring cells.

4.4.1.3 Matching track(s) to tau candidates

Tracks are associated with all tau decay modes. The tracks are matched to hadronic tau clusters as track(s) are associated with all tau decay modes. The conditions for matching track(s) to taus are listed below:

1. All the tracks with $P_T^{track} > 1.5$ GeV in a cone of $\Delta R = 0.5$ about the centroid of the calorimeter cone cluster are preselected as candidates for the tau tracks.

2. Among all the preselected tracks only the highest P_T track is attached to the tau cluster.

3. Up to two more tracks can be assigned to the tau cluster if their z-positions are within 2 cm from the first track located at distance of closest approach.

4. A second track is added to the cluster if the invariant mass of the first and the second track is less than 1.1 GeV.

5. A third track is associated to the cluster if the invariant mass of the three tracks is less than 1.7 GeV and the total sum of their charges is either +1 or 1.

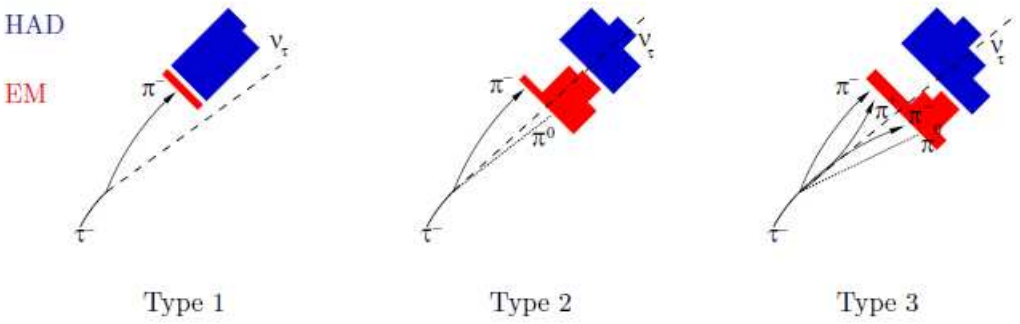

Figure 4.10: Illustration of three types of taus as seen by the detector [24].

4.4.1.4 Define tau types

The algorithm defines three types of hadronic taus in terms of tracks and calorimeter clusters (Fig. 4.10).

- **Type I** One track and hadronic cluster
- **Type II** One track and both EM and Hadronic clusters
- **Type III** Two or three tracks and both EM and Hadronic clusters

4.4.2 Separate taus from background

The taus are separated from hadronic jets using a tau Neural Network Training NN_τ. Three different sets of discriminating variables are used in NN_τ training of three different types of taus. These variables are defined in terms of hadronic and electromagnetic calorimeter cluster energies, shower shape, the detector geometry, tracks, combination of tracks and clusters. A tau rich $Z \rightarrow \tau^+\tau^-$ signal Monte Carlo sample and jets back-to-back to non-isolated muons from data as a background sample is used to train the NN_τ. Each tau in an event is given a NN_τ value between 0 and 1: Good taus have NN_τ value close to one and fake taus have value around zero. For this analysis, in an event tight/good taus are defined by $NN_\tau > 0.95$ and medium (loose-tight) taus are define by $0.3 < NN_\tau < 0.9$.

The complete list of the variables are given below:

1. *caliso* is defined by the ratio of the transverse energy in the annular region between the cones of $\Delta R = 0.3$ and $\Delta R = 0.5$ (Fig. 4.8), to the total transverse energy of the hadronic tau cluster, or $(E_T^\tau - E_{Tcore}^\tau)/E_{Tcore}^\tau$, where, E_{Tcore}^τ is the transverse energy within a cone of $\Delta R = 0.3$ and E_T^τ is the transverse energy of the hadronic tau cluster within a cone of $\Delta R = 0.5$. This is used in NN_τ training for all types of taus.

2. *EM12isof* is defined by the ratio of energy deposited in the first two layers of the electromagnetic calorimeter to the total energy of the hadronic tau

cluster, or $(E^{EM1} + E^{EM2})/E^{\tau}$, where E^{EM1} and E^{EM2} are the energies deposited in the first two layers of the EM calorimeter with in a cone of $\Delta R = 0.5$. This is used in NN_{τ} training for tau Type I only.

3. ET_o_sum is defined by the ratio of transverse energy of the tau cluster to the sum of transverse energy of the tau cluster and all track momenta associated with the tau, or $E_T^{\tau}/(E_T^{\tau} + \sum P_T^{\tau_{track}})$. This is used in NN_{τ} training for all types of taus.

4. $profile$ is defined by the ratio of sum of the first two energetic tower to the total transverse energy of the hadronic tau cluster, or $(E_T^{T1} + E_T^{T2})/E_T^{\tau}$, where, E_T^{T1} and E_T^{T2} are energies of the two most energetic towers within the cone of $\Delta R = 0.5$. This is used in NN_{τ} training for all types of taus.

5. $prf3$ is defined by the ratio of transverse energy of highest EM sub cluster to the transverse energy deposited in the third electromagnetic layer of the calorimeter within a cone of $\Delta R = 0.5$ or $E_T^{EM}(leading)/E_T^{EM3}$. This is used in NN_{τ} training for only tau Type II.

6. $emET\_o\_ET$ is defined by the ratio of the transverse energy of the EM sub clusters to the transverse energy of the tau cluster. This is used in NN_{τ} training for tau Type II and III.

7. ett_o_ETiso is defined by the ratio of transverse momentum of the leading track that is matched to a tau to the transverse energy deposited in a isolation cone of $\Delta R = 0.7$, or $P_T^{\tau}(leading\ tau\ track)/E_T(iso)$. This is used in NN_{τ} training for only tau Type III.

8. $\Delta\alpha$ is defined by

$$\frac{\sqrt{(\frac{\Delta\phi}{sin(\theta)})^2 + (\Delta\eta)^2}}{\pi},$$

where $\Delta\eta$ and $\Delta\phi$ are the pseudorapidity and azimuthal differences in between the vector sum of the tau tracks and the vector sum of all electromagnetic sub clusters and θ is the polar angle of the vector sum of the tau track momenta. The variable is used to find the opening angle between the vector

sum of tau track(s) directions and the vector sum of EM clusters. This is used in NN_τ training for tau Type II and III.

9. *trkiso* is defined by the ratio of sum of transverse momentum of the tracks not associated with the tau candidate with in a cone of $\Delta R = 0.5$ to the sum of the transverse momentum of tracks matched with tracks or $\sum P_T^{track}(non\ tau)/\sum P_T^{\tau track}$. This is used in NN_τ training for all types of taus.

10. fhf (fine hadronic fraction) is defined by the fraction tau cluster transverse energy deposited in the fine hadronic layers of the calorimeter or E_T^{FH}/E_T^τ. This is used in NN_τ training for all types of taus.

11. *rms* is defined by
$$\sqrt{\frac{\sum_{i=1}^{n}((\Delta\phi_i)^2+(\Delta\eta_i)^2)E_{T_i}}{E_T^\tau}},$$
where $\Delta\eta$ and $\Delta\phi$ are the pseudorapidity and azimuthal differences between calorimeter tower and centroid of the tau calorimeter cluster, i is the index of calorimeter towers associated with a tau cluster, E_T^τ is the total transverse energy of the tau candidate and E_{T_i} is the transverse energy of tower i. This variable is used in NN_τ training for all types of taus.

12. Detector pesudorapidity of a tau cluster η_d. This is used in NN_τ training for all types of taus.

4.5 Missing Transverse Energy Reconstruction

Neutrinos produced in $p\bar{p}$ collisions leave the detector without detection. However, their existence can be correlated with a large transverse momentum imbalance in collisions. As collisions occur along the beam axis (in the z direction), the net momenta perpendicular to the beam line (in x-y plane) is zero. The missing transverse energy ($\not\!\!E_T$) is calculated from the vector sum of the transverse energies

of all the good cells in the calorimeter except for some noisy Coarse Hadronic calorimeter cells.

For events with both electromagnetic objects and jets, the $\not\!\!E_T$ needs to be corrected for corrected JES. This correction is applied by subtracting JES correction from the raw $\not\!\!E_T$ vector.

After the object reconstruction procedure all the objects are stored event by event in a root file format for further analysis.

4.6 Tau Triggers: Special Topic

This analysis uses RunIIa data taken with ALLJET triggers but not tau triggers. Tau triggers are a new addition to RunIIb data (after 2006) and are designed with an aim to be used for future $t\bar{t} \rightarrow \tau + jets$ related analysis. Hence, a description of tau triggers are provided for prospective researchers.

During the RunIIb upgrade in 2006, DØ completely redesigned its calorimeter trigger [56] system to meet two major goals: 1) to cope with the increase in instantaneous luminosity delivered by the Tevatron and 2) to efficiently trigger on interesting physics objects such as electromagnetic objects etc.. To meet these goals a new algorithm called Sliding Window algorithm is implemented. There are many advantages of this algorithm [57], one of which is that it allows DØ trigger system to trigger on objects like hadronic taus. In the following sections DØ hadronic tau triggers are described.

4.6.1 Level 1 Tau Trigger

The RunIIa level 1 trigger system sees the total cylindrical volume of the detector as an array of 1280 Trigger Towers or 40×32 TTs in the region of $(4 \times 2\pi)$ in $(|\eta|, \phi)$ space (Sec. 3.1.4). For a jet, level 1 trigger decision is made by total transverse energy information from the trigger towers in a region of (0.2×0.2) in $(\Delta\eta \times \Delta\phi)$ space.

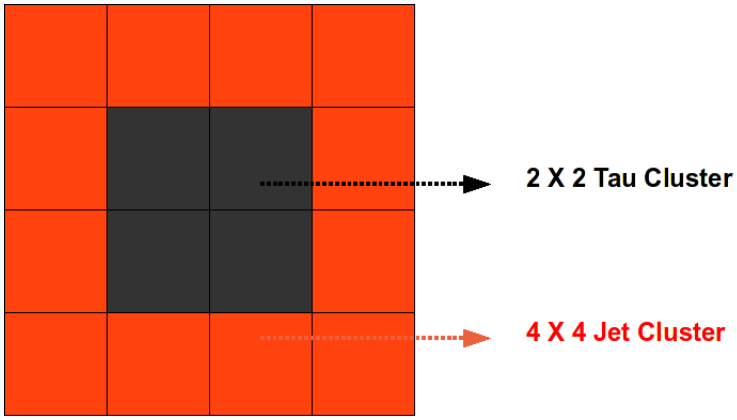

Figure 4.11: Illustration of Sliding Window algorithm for hadronic tau clusters.

On the other hand, RunIIb Calorimeter Sliding Window algorithm (CSW) performs a parallel search for local maxima of transverse energy deposition among trigger towers of the calorimeter using a fixed "window" (Trigger Tower Cluster) made up of a group of adjacent trigger towers. Local maxima are determined by sliding the window and searching for a window whose sum of trigger tower transverse energy is higher than all of its nearest neighbor windows. The windows are defined by (2×2) Trigger Towers, corresponding to a (0.4×0.4) region in $(\Delta\eta\times\Delta\phi)$ space. Local maxima are found by requiring the sum of transverse energies in a window to exceed those in all other windows in a (5×5) Trigger Tower region about a candidate tower.

A jet object is defined by the total transverse energy in the (4×4) (corresponding to a region of (0.8×0.8) in $(\Delta\eta \times \Delta\phi)$ space) window surrounding the maximum window.

As described in Sec. 4.4 the hadronic tau objects appear in a narrow energetic cone in the calorimeter compared to jets which appear in a wide cone . This led to the idea of designing hadronic tau triggers by requiring a large ratio of the transverse energy in the (2×2) window to that in the (4×4) jet window (Fig. 4.11) [57, 58]. The higher the ratio the higher the probability of triggering tau objects. For example, the tau trigger TAUJT_MET15T10TK is designed based on

CSW algorithm to trigger on events that will improve both RunIIb $t\bar{t} \rightarrow \tau + jets$ cross section results and limits on Charges Higgs searches. At level 1 the trigger definition for TAUJT_MET15T10TK consists of a number of trigger terms,

$$\text{tau} - \text{term} \times \text{two} - \text{jet} - \text{terms} \times \text{track} - \text{term} \times \not{E}_T - term.$$

The tau trigger logic term (tau-term) is defined by CSWTA(1,10.,2.4,2) to trigger on events with tau like objects and the rest of the trigger terms are used to reduce the trigger rate. The CSWTA(1,10.,2.4,2) term tells the level 1 trigger system to search for an event that has a tau object within $|\eta| < 2.4$, above 10 GeV trans verse energy threshold, and with a minimum core energy (energy in the (2×2) region) of 8 GeV that passes a ratio (defined by $(E(2 \times 2)/E(4 \times 4))$ cut greater than 0.75 (in the logic term 2 refers to 7.5). If the CSWTA condition is satisfied along with other trigger terms then the event is send to level 2 trigger for further decisions.

4.6.2 Level 2 Tau Trigger

Similar ratio cuts are implemented at level 2 in order to trigger events containing hadronic tau objects. A ratio is defined by a seed tower and the neighboring tower with the highest $E_T(seed + nearest\ neighbor)$ to the E_T sum in the (3×3) tower region (around the seed tower). For instance, level 2 tau term of the trigger TAUJT_MET15T10TK is defined by L2TAUTRK(1,5,8,0.75,TTK). The L2TAUTRK(1,5,8,0.75) logic part of the term tells Level 2 trigger to find objects in an event which contains one tau with $E_T > 8$ GeV and ratio $(E(seed + nearest - neighbor/E(3 \times 3)) > 0.75$ and the other part requires a track.

4.6.3 Level 3 Tau Trigger

The level 3 hadronic tau algorithm is similar to the tau reconstruction algorithm. The notable part of the level 3 tau trigger is that it uses Neural Network (NNTauL3) algorithm to reject background events and keep the trigger rate within the band width. The NNTauL3 is trained with five variables that can separate taus

from background. The NNTauL3 variables are $profile$, $caliso$, $EM12isof$, $EtaD$ and $width/rms$. The value of the NNTauL3 cuts ranges from $0-0.4$ and the cuts are utilized on some tau triggers to keep the trigger rate within the acceptable range. The level 3 triggers usually trigger events with $E_T > 10$ GeV and cannot distinguish between tau types.

Chapter 5

Samples and Event Selection

This chapter explains the procedure for the selection of $t\bar{t} \rightarrow \tau + jets$ signal events which is one of the most important steps for cross section measurement. The data selection procedure is based on signal observables such as the number of jet objects, tracks, missing transverse energy and narrow calorimeter clusters. Once the total number of $t\bar{t} \rightarrow \tau + jets$ events, the total integrated luminosity and the detector efficiency are known, the cross section should be measurable. However, even with strict selection criteria, the data contain mostly background events. The main reason behind this is that the detector and its components are far from perfect. The sources of detector inefficiencies arises from the trigger system all the way to the object reconstruction procedure and needs to be taken into consideration for the cross section measurement. Moreover, the background can arise from electroweak physics processes which can mimic the signal like events. Therefore, it is necessary to accurately evaluate the number of signal events and background events in the dataset.

This chapter is divided into two parts, the first part deals with data selection procedure, which is based on signal observables from a large set of experimental data (Sec. 5.2), while in the second part, tools are developed for estimating the number of signal events, signal selection efficiency (Sec. 5.3) as well as background events.

5.1 Samples and Event Selection Strategy

The event selection strategy is based on final state products of $t\bar{t} \rightarrow \tau + jets$ decay events, which includes four quarks, a tau lepton and a neutrino. In the detector, these particles are identified by the energy they deposit in the calorimeter, tracks they leave in the tracking system and the missing energy in the transverse plane.

The final state products of signal events have the following event characteristics in the detector:

1. A narrow calorimeter jet with associated track(s), a characteristic of a hadronic tau decay,
2. Four or more jets, two of which are initiated by b-quarks while the other two jets are from W decay and
3. A large net missing transverse energy ($\not\!\!E_T$) due to energetic neutrinos mostly from W and some from tau decays.

Therefore, the data sample selection and the trigger selection procedures are designed based on the final state observables of $t\bar{t} \rightarrow \tau + jets$ signal events in the detector. The observables are later tuned in the selection process to maximize the probability of detecting signal events in the data.

In addition, rejection of background events that mimic signal event observables in the detector is also a powerful mechanism for enhancing the signal like events. Therefore, tools that have background rejection power are also utilized in the selection procedure.

5.2 Data Sample Selection

This analysis utilizes DØ RunIIa data recorded over the time period of 2002 to 2006 inclusive. The recorded data sample contains a huge number of background

events such as events with one jet, two jets and three jets with very low P_T. In order to reduce the number of background events, a set of skim cuts are imposed (Table 5.1). After skim cuts, the data sample contains approximately 115 million events and is known as "3JET" data skim.

Leading Jet (in P_T)	P_T (GeV)	$\|\eta\|$
1st Jet	> 20	< 2.6
2nd Jet	> 15	< 2.6
3rd Jet	> 15	< 2.6

Table 5.1: List of all 3JET data skim cuts applied to preselected RunIIa data.

Trigger Versions	Trigger Names	Luminosity $\mathcal{L}$ (pb^{-1})
V09.00-V10.00	4JT10	24.07
V10.00-V11.00	4JT10	10.22
V11.00-V12.00	4JT10	64.68
V12.00-V13.03	4JT12	210.52
V13.03-V13.30	JT2_4JT12L_HT	49.40
V13.30-V14.00	JT2_4JT12L_HT	257.39
V14.00-V14.80	JT2_4JT20	189.83
V14.80-V15.00	JT2_4JT20	143.55
T O T A L		974.19

Table 5.2: Description of RunIIa ALLJET triggers with their respective luminosity and versions.

5.2.1 Trigger Selection

The final state of $t\bar{t} \rightarrow \tau + jets$ consists of four jets, missing transverse energies and a narrow calorimeter cluster with tracks. If triggers could record data events with such objects, it would be ideal. However, since there are no such triggers available

in DØ RunIIa data, triggers are selected on the basis of jet characteristics of signal events. The set of triggers chosen is called ALLJETS triggers (detailed description of all the triggers can be found in Sec. 3.1.4 and in appendix A), which record data events from $p\bar{p}$ collisions during the RunIIa data recording period. From collisions, these triggers select data events containing at least four jets above an energy threshold and within a pesudorapidity region in an event. The total amount of data recorded by these triggers is expressed in terms of integrated luminosity which is $974.2pb^{-1}$ ($\approx 1fb^{-1}$). Table 5.2 lists a short description of the triggers used for this analysis with their respective luminosity and versions.

5.2.2 Data Quality Cuts

The DØ detector occasionally encounters problems while recording data due to the temporary malfunction or underperformance of certain parts of the detector. The data recorded under such conditions are not acceptable for physics analysis. To remove these "problematic" data, a set of quality criteria is imposed to check for the proper operation of all of the detector components. These quality checks are implemented in terms of run[1] quality basis and luminosity block quality[2] basis.

The run quality ensures that all the detector components are properly functional during a run. This check is important for track quality which is used for lepton identification. The luminosity block quality removes calorimeter noise such as *ring-of-fire* (a ring of energy along the ϕ direction) which causes a large missing energy in an event.

[1] A run is a data-taking time period which approximately lasts for two to four hours with fixed detector configuration.

[2] A luminosity block represents a much shorter time period than that of a run and last for approximately one minute.

Selection Cuts	***Number of Data Events***
Initial Number of Events:	115832858
Trigger Selection:	31063568
Data Quality Cuts:	22036392
JESMU Jet Energy Scale Corrections: **Jet Selection Cuts:** At least four jets with $\|\eta\| < 2.5$, three leading (in P_T) jets with P_T s , 35, 25, and 25 GeV, and all other jets with $P_T > 13$ GeV	15653441
Veto Electrons and Muons:	15650306
Primary Vertex Selection Cuts: $\|z\| < 60$ cm, at least 3 tracks	15006194
$\not\!\!E_T$ **Selection Cuts:** The $\not\!\!E_T$ cut range is 15 GeV $\leq \not\!\!E_T \leq$ 500 GeV	6579088
$\not\!\!E_T$ **Significance Selection Cut:** $\not\!\!E_T$ Significance > 3.0	498476
Loose Tau Selection: At least one tau with $NN_\tau > 0.3$ and $P_T > 10$ GeV	191341
Final Selections Cuts :	
Final Jet Selection All jets with $P_T > 20$ GeV, $\|\eta\| < 2.5$, ***b*-jet Identification:** At least one *b*-tag jet **Tight Tau Selection Cuts:** $NN_\tau > 0.95$ with $P_T > 10$ GeV and $\|\eta\| < 2.5$ **Tau Type Selection Cut:** Select tau Type I and Type II together	414

Table 5.3: Data event selection cut flows for tau Types I & II.

5.2.3 Data Event Selection

In order to measure the cross section, it is essential to obtain the number $t\bar{t} \rightarrow \tau + jets$ events in the experimental data. Based on the characteristics of $t\bar{t} \rightarrow \tau + jets$ like events, the event selection cuts are designed as follows:

- **Jet Selection:** Jet selections are crucial for selecting $t\bar{t} \rightarrow \tau + jets$ like events. The jet selection procedure requires at least 4 jets within $|\eta| < 2.5$. In addition, the P_T of each jet in an event is required to be above 13 GeV. Furthermore, three leading (in P_T) jet's transverse momentum are selected as follow; $P_T > 35$ GeV for the first jet and $P_T > 25$ GeV for the second and the third jets.

- **Veto Electrons and Muons:** In order that the $t\bar{t} \rightarrow \tau + jets$ data be orthogonal to the lepton related top analyses, data such as $e + jets$ data, $\mu + jets$ data and $dilepton + jets$ data, events with isolated electrons and isolated muons are rejected [65, 67].

- **Primary Vertex Selection:** The primary vertex position is selected to be within ± 60 cm from the center of the detector and is reconstructed with at least three tracks.

- $\not\!\!E_T$ **Selection:** The missing transverse energy $\not\!\!E_T$, is selected within the range of 15 GeV $\leq \not\!\!E_T \leq$ 500GeV. The cut of 500 GeV is chosen to avoid unphysical events.

- $\not\!\!E_T$ **Significance Selection:** The $\not\!\!E_T$ significance is a likelihood type variable. The variable computes the probability distribution for $\not\!\!E_T$ based on resolutions of each event's particular objects such as jets, muons, electrons and unclustered energy.

  These resolutions are predicted by measuring and parameterizing the object resolutions in data. For example, the jet resolution is obtained by measuring the momentum imbalance in dijet events. A good description of this variable can be found in Ref. [83, 78, 79, 80].

The $\not\!\!E_T$ significance variable is a powerful tool for separating real $\not\!\!E_T$ from fake $\not\!\!E_T$. For instance, the major background for this analysis is multijet events as they may contain fake $\not\!\!E_T$ and can pass signal selection criteria. Therefore, to reject the multijet events a requirement of $\not\!\!E_T$ significance $>$ 3 is used. Figure 5.1 shows the background discrimination power of $\not\!\!E_T$ significance.

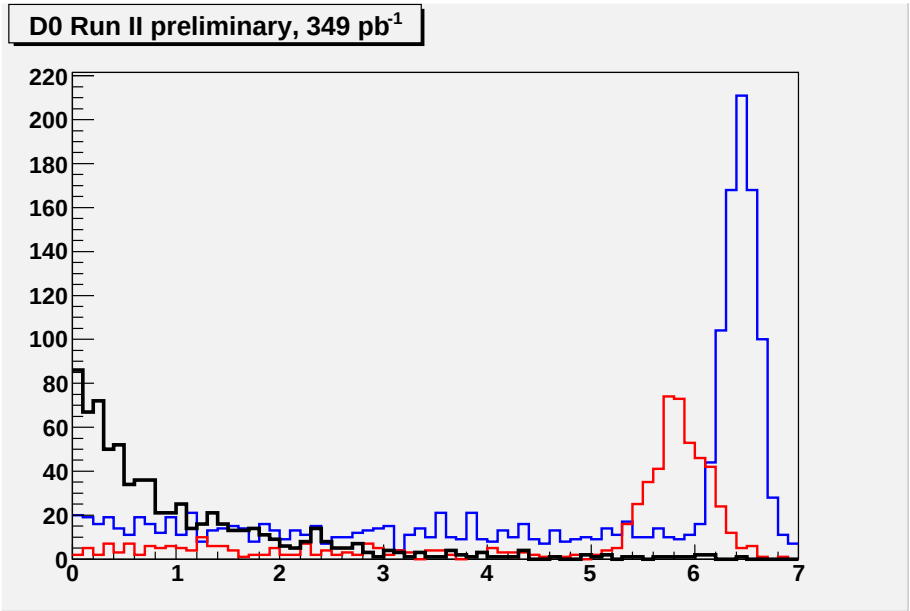

Figure 5.1: Plot for $\not\!\!E_T$ significance. The $\not\!\!E_T$ significance for multijet dominated data are ploted in black, $W + jets$ simulated events are plotted in blue and $t\bar{t} \rightarrow \tau + jets$ simulated events in red. It is clear from the plot that a $\not\!\!E_T$ significance cut of 3 removes most of the background.

- **Loose Tau Selection:** The first tau selection starts with a loose tau selection requirement. The loose tau selection requires at least one tau with $NN_\tau > 0.3$ and with $P_T > 10$ GeV. This selection removes a sizable number of background events and reduces computing time.

- **Veto all-jets:** In order to avoid overlapping between $t\bar{t} \rightarrow \tau + jets$ data and the data used for $t\bar{t} \rightarrow$ all $-$ jets cross section measurements, a veto condition is imposed. The veto condition is designed on the basis of the all-jets event selection criteria [66, 83]. The all-jets selection criteria are as follows:

- At least 5 jets with $P_T > 15$ GeV and η <2.5, with 3 leading (in P_T) jets having $P_T > 40$ GeV
- All 5 jets have to be taggable
- Two jets have to be b-tagged with a loose tagging operating point.

- **b-jet Identification:** The b-jet requirement in an event increases the probability of selecting top events. Therefore, the b-jet tagging algorithm described in Sec. 4.3.6 is used to identify b-jets. At least one b-jet in an event is required for this analysis. Jets matched to tau candidates in ΔR are not tagged.

- **Tight Tau Selection:** Taus with high NN_τ are considered good taus as explained in Sec. 4.4. The tight selection requires at least one tau with $NN_\tau > 0.95$ and $P_T > 10$ GeV and within $|\eta| < 2.5$ in an event.

- **Tau Type Selection:** After all these selections are applied, the selected data sample is split into two samples in terms of tau types. The first one is called tau Types I & II data sample and it includes events with Type I and Type II taus. The second sample is called tau Type III data sample, which includes events with only Type III taus.

Finally after all these selection cuts, the tau Types I & II data sample contains 414 events and tau Type III data sample contains 1115 events out of a population of 115 million events. Even after tight selections, these two data samples still contain a large number of background events. At this stage of the analysis it is impossible to determine what fraction of the data events are signal like events and what fraction are background like events in each dataset. Therefore, in the next step, signal like events are simulated using a Monte Carlo technique as shown in Sec. (5.3) and background events are modeled using both background like data and simulated background events to estimate the signal content both in tau Types I & II data and in tau Type III data.

5.3 Monte Carlo Tools for Data Modeling

In Sec. 2.4 it is argued that the Standard Model predicts the production of the top quark and its decay into different stable particles in proton antiproton collisions. Therefore, to check the validity of the Standard Model predictions, events are generated using Monte Carlo technique to compare it with the data. The events thus generated are called Monte Carlo events or simulated events and are widely used in high energy physics.

5.3.1 Monte Carlo Sample Selection

The top pair production signal Monte Carlo samples are simulated for two different top decay channels: the $t\bar{t} \rightarrow lepton + jets$ channel and the $t\bar{t} \rightarrow dilepton$ channel. The $t\bar{t} \rightarrow lepton + jets$ channel is defined by the final state product of top quark decay events that contains one lepton, one neutrino, and four quarks. This happens when the two top quarks decay into two W bosons, one that decays into two quarks while the other decays into a lepton. In an event, when both W bosons decay into leptons it is referred to as a $t\bar{t} \rightarrow dilepton$ channel.

The major source of background present in both tau Types I & II and tau Type III data are multijet events produced by the strong interaction process. From a multijet process a narrow energetic jet can often fake tau like objects and pass the data selection criteria. Moreover, although QCD processes do have neutrinos, QCD/multijet events can have fake $\not\!\!E_T$ arising from detector gaps or bad cells and may pass the signal selection criteria. As there are no well modeled multijet simulation method available for this analysis, such contributions are modeled from data with loose-tau selection as cuts explained in Sec. 5.3.6. The electroweak processes such as $W \rightarrow \tau\nu$, can pass the data selection cuts and appear as background. The electroweak processes are simulated by Monte Carlo technique and the number of electroweak events is estimated by the signal selection criteria (Sec. 5.3.6).

Signal Samples	***Background Samples***
$t\bar{t} + 0lp \rightarrow l\nu jjb\bar{b} + 0lp$	$W + Nlp, \quad N = 0 - 5$
$t\bar{t} + 1lp \rightarrow l\nu jjb\bar{b} + 1lp$	$Wb\bar{b} + Nlp,\ N = 0 - 3$
$t\bar{t} + 2lp \rightarrow l\nu jjb\bar{b} + 2lp$	$Wc\bar{c} + Nlp,\ N = 0 - 3$
$\boldsymbol{t\bar{t} \rightarrow lepton + jets}$	$\boldsymbol{W + jets}$
$t\bar{t} + 0lp \rightarrow l\nu l\nu b\bar{b} + 0lp$	$Z + Nlp, \quad N = 0 - 3$
$t\bar{t} + 1lp \rightarrow l\nu l\nu b\bar{b} + 1lp$	$Zb\bar{b} + Nlp,\ N = 0 - 3$
$t\bar{t} + 2lp \rightarrow l\nu l\nu b\bar{b} + 2lp$	$Zc\bar{c} + Nlp,\ N = 0 - 3$
$\boldsymbol{t\bar{t} \rightarrow dilepton + jets}$	$\boldsymbol{Z + jets}$

Table 5.4: List of Monte Carlo samples generated.

5.3.2 Monte Carlo Simulation Process

The simulation of hard scattering processes in $p\bar{p}$ collisions at the center of mass energy of 1.96 TeV involves the usage of different generators with different simulation capabilities. The simulation process is accomplished in multiple phases which include parton level event simulation, hadronization and showering, and detector level event simulation which are explained below:

5.3.2.1 Parton Level Simulation and Hadronisation

In order to simulate Monte Carlo events for hard scattering processes in hadronic collisions, an event generator known as ALPGEN is used. The generator consists of a collection of codes for a leading order calculation of matrix elements for a large set of parton-level (quark/gluons) processes of interest. In the simulation process, a parton distribution function (CTEQ6.1M [68]) is used to model the quarks/gluons inside protons.

To make these partonic processes useful for the experimental data, the output of the ALPGEN generator is fed into another generator called PYTHIA [60]. PYTHIA simulates transformation of the partons into observable hadrons (hadronisation process) as well as the development of partonic cascades (shower development process). However, the generator PYTHIA cannot simulate decays of tau and B hadrons properly. Therefore, in the next stage, two simulators TAUOLA [61] and EVTGEN [62] are used to simulate tau and B hadron decays respectively.

5.3.2.2 Detector Level Simulation

The events produced by the ALPGEN and the PYTHIA interface do not have any detector information and need to be processed by the detector simulator. To include detector information such as tracking hits, calorimeter cell energy and muon hits to these parton-level Monte Carlo events such that they represent data like events, a DØ detector based simulation package D0gstar [69] is used. The D0gstar simulation package is built on GEANT [63] which models the interaction

between particles and material or electromagnetic fields by using detector geometry as well as material definitions and the magnetic field map.

For simulation of the digitization process, the electronics noise and pile-up effects (effects from the previous collisions), a separate software package, known as D0sim [70] is used. To include these type of detector effects in the Monte Carlo, D0sim utilizes real data events recorded without any trigger requirements. These data are also known as zero bias data and mostly contain information of electronics noise and energy deposited by cosmic muons.

5.3.2.3 Trigger Modeling For Monte Carlo

The simulation of detector trigger conditions are implemented by a probabilistic interpretation method. The ALLJET trigger is performed by a probabilistic interpretation method based on the jet object P_T turn-on curve (P_T vs efficiency curve). There are three levels of trigger and the probability of any event passing all these three trigger levels can be written as the product of conditional probabilities

$$P_{event} = P(L1) \times P(L2|L1) \times P(L3|L1L2), \tag{5.1}$$

where $P(L1)$ is the probability of an event passing level 1 trigger condition, $P(L2|L1)$ is the probability of an event passing level 2 trigger, conditional on the fact that it has already passed the level 1 trigger condition and $P(L3|L1L2)$ is the probability of an event passing the level 3 trigger conditional on the fact that it has already passed the previous two trigger levels.

The level 1 trigger probability for passing at least one jet for an event with N jets is given by

$$P(L1) = 1 - P^0 = 1 - \prod_{i=1}^{N}(1 - P_i(P_T)) \tag{5.2}$$

where P_i is the ith jet efficiency as a function of ith jet P_T.

The trigger 4JT10 (Table 5.2 and appendix A) can be used as an illustrative example to explain Eq. 5.2. The level 1 trigger condition requires at least 4 jets

with $P_T > 5$ GeV. So the level 1 trigger probability is calculated by Eq. 5.3 given that the turn-on curve for single jet P_T is known (jet turn-on curves are derived by the DØ jet ID group and used for this analysis).

$$\begin{aligned} P(L1) &= 1 - P^0 - P^1 - P^2 - P^3 \\ &= 1 - \prod_{i=1}^{N}(1 - P_i(5)) \\ &- \sum_{i=1}^{N} P_i(5) \prod_{i\neq j}^{N}(1 - P_j(5)) \\ &- \sum_{i=1}^{N} P_i(5) \sum_{j=1}^{N} P_j(5) \prod_{i\neq j,k}^{N}(1 - P_k(5)) \\ &- \sum_{i=1}^{N} P_i(5) \sum_{j=1}^{N} P_j(5) \sum_{k=1}^{N} P_k(5) \prod_{i\neq j,k,l}^{N}(1 - P_l(5)) \end{aligned} \tag{5.3}$$

Trigger probabilities for level 2 and level 3 are calculated in a similar fashion for all ALLJET triggers.

In addition, a level 4 trigger probability is used for Monte Carlos in order to account for the effect of the 3JET skim cuts in Table 5.1 which is implemented before the jet energy scale corrections. The problem with skim cuts before applying the jet energy corrections can be explained with an example data event which consists of at least three jets with leading (inP_T) jet P_T s 24 GeV, 16 GeV, and 16 GeV with $|\eta| < 2.6$. This event will fail the skim cut. However, if the 3JET skim cuts are applied after the jet energy corrections which changes the value of the leading jet P_T to 26 GeV, then skim cuts will allow the event to pass. The best way to simulate this effect in Monte Carlo events is by considering the cuts as a level 4 trigger which is calculated in the same manner as explained above.

5.3.2.4 ALPGEN Weighting Factor

The ALPGEN weighting factor is used to remove duplicate events in the Monte Carlo generated events and is known as the MLM jet-parton matching scheme [82]. As explained in the section on the event simulation process (Sec. 5.3.2.1), the events are generated with the ALPGEN generator and then interfaced with the PYTHIA generator for hadronisation and showering simulations. Moreover, samples are generated in flavors and in number of partons (Table 5.4). Because of this interface, some events are double counted when the samples are combined. For instance, when a Wgg light parton event which is generated by the ALPGEN generator passes to PYTHIA for hadronisation and showering, there is the possibility that the gluon will split into heavy quarks $g \rightarrow b\bar{b}$ leading to double counting of $Wb\bar{b}$ events. Therefore, to address the issue of double counting, a factor is derived and used in this analysis for combining Monte Carlo samples generated with the ALPGEN and the PYTHIA interface. For $t\bar{t} \rightarrow \tau + jets$ signal Monte Carlo sample, the calculated weighting factor is 0.0065 (Table B.2).

5.3.3 Monte Carlo Normalization

The ALPGEN Monte Carlo samples for signal and background in Table 5.4 are normalized to the integrated luminosity using Eq. 5.4, where σ is the cross section and $\mathcal{L}$ is the total integrated luminosity which is $974.2pb^{-1}$ ($\approx 1fb^{-1}$).

$$Number\ of\ events\ generated\ = \mathcal{L} \times \sigma \tag{5.4}$$

k-factor The cross section σ, used in 5.4 for simulated event normalization is a leading order process cross section which does not represent data. In order to correct for this, a scaling factor known as k-factor is derived to scale the cross section so that these simulated events represents data like events. The k-factor is derived by taking the ratio of leading order (LO) theoretical cross section to the NNLO cross section [15].

W and Z Scaling Factors The scaling factor scales the W+jets and Z+jets Monte Carlo samples to make them comparable to the data. The background Monte Carlo samples W+jets and Z+jets are generated in a similar fashion as the signal Monte Carlo samples, that is by first using the ALPGEN generator and then using the PYTHIA interface for hadronization. The Monte Carlo events are generated in to two parts: the heavy flavor events and the light parton (lp) events. The heavy flavor events consist of W$b\bar{b}$+jets or Z$b\bar{b}$+jets and W$c\bar{c}$+jets or Z$c\bar{c}$+jets and the light parton events consist of W+Nlp or Z+Nlp events.

However, the heavy flavor samples show poor agreement when compared to data. This disagreement is attributed to using a lower order calculation for heavy flavor event generation. Therefore, a correction factor or heavy flavor scaling factor is derived in order to achive good agreement between data and Monte Carlo [72, 73, 74] (Table 5.5).

For W+jets light flavor events the scaling factor is derived from the EMinclusive data set (described in the Ref. [65]) with the following selection cuts:

- $\not\!\!E_T > 15$ GeV
- One tight electron with $P_T > 15$ GeV
- The e+jets trigger [65], matched to this electron,
- At least one jet

With these cuts, the W peak is clearly observed on the plot of transverse mass of the electron and $\not\!\!E_T$ ($m_T = \sqrt{2\ \not\!\!E_T E_e(1 - cos(\Delta\phi(e, \not\!\!E_T)))}$) plot (Fig. 5.2). As can be seen from Fig. 5.2, the Monte Carlo sample needs to be scaled up by a factor of about 1.50±0.01 (stat). Figure 5.3 shows the exact fit for this value. After a $\not\!\!E_T$ significance cut of 3 (as used for this anlysis), the scaling factor is closer to 1.3±0.1 (stat) (Fig. 5.4). This is the value that is used throughout the analysis to normalize W+Nlp samples.

However, the selection cuts used to derive the scaling factor can pass $t\bar{t}$ events and contaminate W+Nlp samples. No measures have been taken to remove $t\bar{t}$ contamination since this factor is comparable to the factor used in $t\bar{t} \rightarrow lepton + jets$

[65] analysis that uses Matrix Element method to remove $t\bar{t}$ contamination from W+Nlp samples.

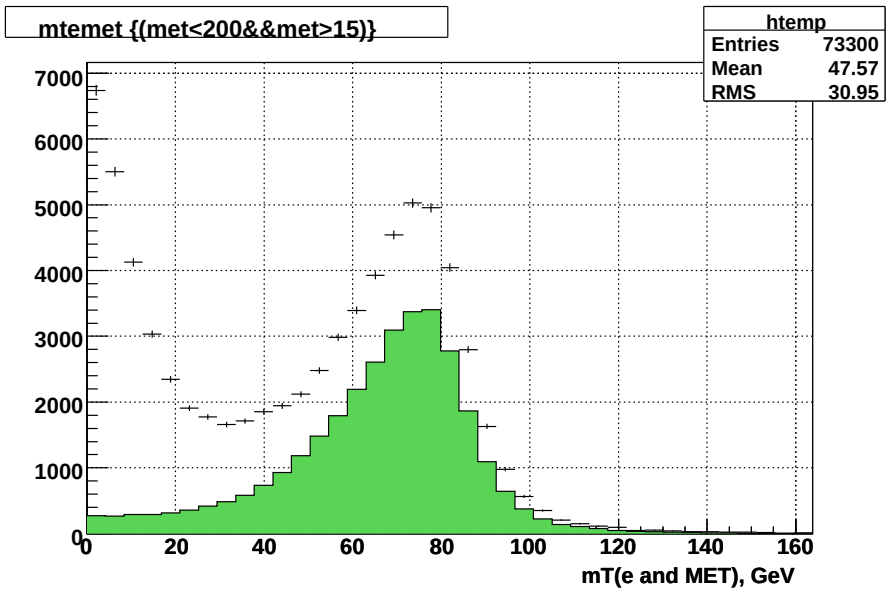

Figure 5.2: The transverse mass of electron and $\not\!\!E_T$ for data (points) and Monte Carlo ($W \rightarrow e\nu$) (histogram).

5.3.4 Monte Carlo Corrections

After the simulation process, a number of correction factors are needed to have good agreement between data and Monte Carlo. The correction factors are described in the following sections and are implemented during the Monte Carlo event selection process.

5.3.4.1 Luminosity Reweighting

As described in Sec. 5.3.2.2, the detector simulator D0sim uses zero bias data events to simulate detector effects such as electronics noise, pile-up etc. in the Monte Carlo sample. The problem of using zero bias data is that it contains low

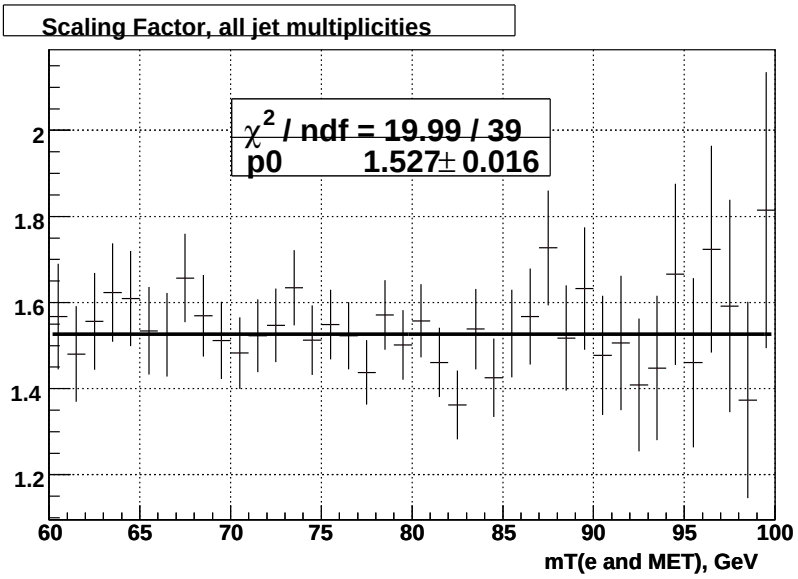

Figure 5.3: The fit for scaling factor inside the "W-window" of $m_T(\not\!\!E_T, e)$

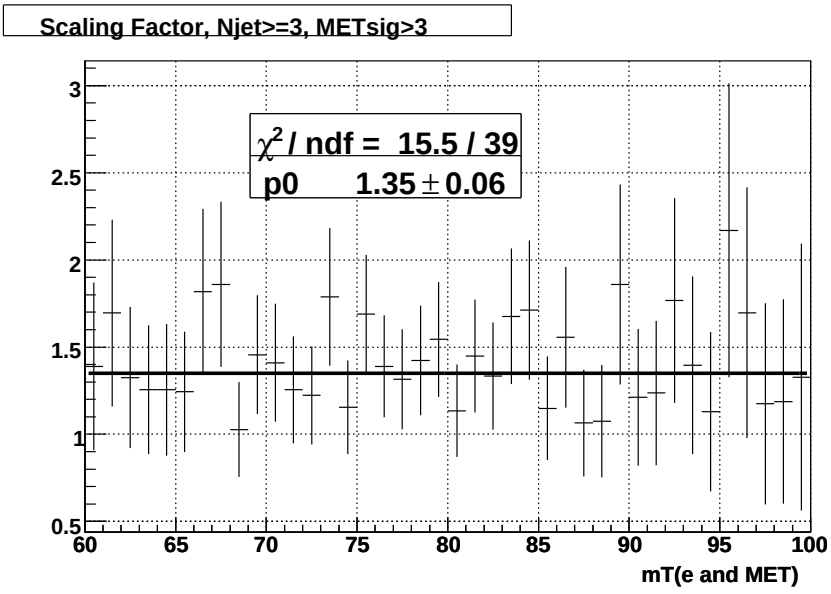

Figure 5.4: The fit for scaling factor inside the "W-window" of $m_T(\not\!\!E_T, e)$ after $\not\!\!E_T$ significance cut of 3

Signal Samples	k-fator		***Background Samples***	k-factor	HF factor
$t\bar{t}+0lp \rightarrow l\nu jjb\bar{b}+0lp$	1.2		$W+Nlp$	1.2	
$t\bar{t}+1lp \rightarrow l\nu jjb\bar{b}+1lp$	1.2		$Wb\bar{b}+Nlp$	1.2	1.7
$t\bar{t}+2lp \rightarrow l\nu jjb\bar{b}+2lp$	1.2		$Wc\bar{c}+Nlp$	1.2	1.7
$\boldsymbol{t\bar{t} \rightarrow lepton+jets}$	1.2		$\boldsymbol{W+jets}$		
$t\bar{t}+0lp \rightarrow l\nu l\nu b\bar{b}+0lp$	1.2		$Z+Nlp$	1.23	
$t\bar{t}+1lp \rightarrow l\nu l\nu b\bar{b}+1lp$	1.2		$Zb\bar{b}+Nlp$		1.35
$t\bar{t}+2lp \rightarrow l\nu l\nu b\bar{b}+2lp$	1.2		$Zc\bar{c}+Nlp$		1.35
$\boldsymbol{t\bar{t} \rightarrow dilepton+jets}$	1.2		$\boldsymbol{Z+jets}$		

Table 5.5: List of k-factors and heavy flavor factors used for Monte Carlo sample scaling.

luminosity events and the luminosity distribution or luminosity profile is constant, whereas in the data sample the instantaneous luminosity profile varies over time. Therefore, the simulated events do not reflect data events. In order to account for this, an event weight is derived by taking the ratio of normalized instantaneous luminosity plots for data and Monte Carlo. [71]. The effect due to the luminosity reweight is $\approx 0.01\%$ of the total number of $t\bar{t} \rightarrow \tau + jets$ signal events.

5.3.4.2 Primary Vertex Reweighting

The position of the interactions along the beam line (z-axis) can fluctuate due to mechanical issues and temperature changes. However, in the Monte Carlo simulation processes the interaction region is always kept fixed. Therefore, Monte Carlo events are reweighted for differences in z-position of the primary vertex so that they represent the data. The effect due to the primary vertex is $\approx 0.03\%$ of the total number of $t\bar{t} \rightarrow \tau + jets$ signal events.

5.3.5 Signal Monte Carlo Event Selection

The event selection cuts for the Monte Carlos are the same as for the data selection cuts. In addition all the above correction factors are applied to the Monte Carlo samples so that they represent data like events. A total of 375408 $t\bar{t} \rightarrow lepton{+}jets$ events and a total of 532750 $t\bar{t} \rightarrow dilepton + jets$ events are generated. Moreover, for the selection process, the $t\bar{t} \rightarrow lepton{+}jets$ sample is split into three orthogonal lepton channels (Table 2.3) using particle selectors that can select events using parton information stored in the Monte Carlo events. The purpose of this sample splitting is to calculate the $t\bar{t} \rightarrow \tau_{hadronic} + jets$ cross section later in the analysis.

$t\bar{t} \rightarrow \tau_{hadronic} + jets$ *Event Selection*:

The $t\bar{t} \rightarrow \tau_{hadronic} + jets$ is defined by the final state product of top quark decays in the event that contains one tau with a hadronic tau decay mode, one neutrino, and four quarks. The event yields after the selection cuts are shown in the Table (B.2).

$t\bar{t} \rightarrow e + jets$ *Event Selection*:

The $t\bar{t} \rightarrow e + jets$ is defined by the final state product of top quarks decay in the event that contains one electron, one neutrino, and four quarks. In addition this channel contains events from top quark decay events that consists of one tau with a leptonic (e) tau decay mode, one neutrino, and four quarks. The event yield after the selection cuts is expected to be comparable to the $t\bar{t} \rightarrow \tau_{hadronic} + jets$ channel as electrons can pass Tight Tau Selection cuts and are shown in the Table (B.3).

$t\bar{t} \rightarrow \mu + jets$ *Event Selection*:

The channel is defined by the final state product of top quark decays in the event that contain one muon, one neutrino, and four quarks. In addition, this contains events from top quark decays comprising of one tau with a leptonic (μ) tau decay mode, one neutrino, and four quarks. The event yield after the selection cuts is expected to be small as muons rarely passes tight tau Selection cuts as shown in the Table (B.4).

$t\bar{t} \rightarrow dilepton + jets$ *Event Selection*:

The channel is defined by the final state product of top quarks decay in the event that contains two leptons, two neutrinos, and two quarks. The sample is not split in terms of leptons as the event yield after the selection cuts is expected to be small and are shown in Table (B.5).

5.3.6 Electroweak Event Selection

The electroweak processes are physics background and may pass the signal selection criteria. To model these physics backgrounds in data, the electroweak samples in Table 5.4 are passed through with the same signal selection (Table B.2) cuts and similar corrections that are applied for signal Monte Carlo samples. In addition, light flavor and heavy flavor factors are used for respective samples as described in Sec. 5.3.3. The electroweak event yields for different channels are listed in Table B.1.

5.4 QCD/Multijet Background Event Selection

The multijet background is extracted from data and selection cuts are the same as those for data selection cuts. The only exception is the tight tau Selection cut. For background events, the tau selections are designed such that they look like medium taus. Therefore, multijet events are selected with a random tau with $NN_\tau > 0.3$ and $NN_\tau < 0.9$ which is also known as loose-tight tau selection. Also the multijet sample have been split into two parts after selections. The part that is used for estimating multijet fraction of the data events is shown in Table B.6. The other part is used for Neural Network training and is explained in Sec. (6.1.4).

5.4.1 Electroweak Subtraction

The algorithm used for tau identification algorithm is not 100% efficient. Therefore, the multijet events in Table B.6 may contain a tiny amount of $t\bar{t}$ events and a small amount of electroweak events. These events are estimated by the same selection criteria as shown in Table B.2. The only difference is that instead of tight tau event selection, loose-tight tau events are selected. Table B.7 represents the estimated number of electroweak and $t\bar{t}$ events present in the multijet sample. These electroweak events are subtracted from the multijet events to remove contributions of electroweak processes in multijet events.

5.5 Summary of Event Selection

Table 5.6 lists the results of event selection procedure for data, signal Monte Carlo and background events. These events will be used in the next chapter for template fitting and for extracting the number of top pair events present in the data.

Templates	*Events Type I & II*	*Events Type III*
Data	414 ± 20.35	1115 ± 33.39
Signal Monte Carlos		
$t\bar{t} \rightarrow \tau + jets$	27.20 ± 0.38	21.79 ± 0.34
$t\bar{t} \rightarrow e + jets$	25.87 ± 0.36	5.85 ± 0.17
$t\bar{t} \rightarrow \mu + jets$	1.92 ± 0.10	3.63 ± 0.13
$t\bar{t} \rightarrow dilepton + jets$	5.44 ± 0.07	2.65 ± 0.05
Background		
multijet		
Electroweak		
$W + jets$	13.48 ± 0.31	5.92 ± 0.21
$Z + jets$	3.35 ± 0.37	1.93 ± 0.06

Table 5.6: Summary of tau Types I & II and tau Type III event selections.

Chapter 6

Cross Section Measurement

This chapter is divided into two parts: The first part describes the method for extracting $t\bar{t} \rightarrow \tau + jets$ events from the data which is necessary for the cross section measurement and the second part explains the cross section measurement procedure.

6.1 Signal Event Extraction by Fitting

The data obtained after the event selection procedure contain both signal and background events. As described in Table 5.6, the signal $t\bar{t}$ Monte Carlo predicts approximately 60 signal tau Types I & II events and 34 signal tau Type III events. Even though, this only accounts for approximately 16% of the data for tau Types I & II and 3% of the data for tau Type III, it is still possible to extract the number of signal events. In the following sections, methods are developed for extracting tau Types I & II signal events, which will also be used for extracting tau Type III signal events.

6.1.1 Fitting Procedure

The method used to extract signal events from the data involves fitting templates of signal Monte Carlo and background events to the data. There are three templates involved in the fitting procedure: the signal template, the multijet template, and

the electroweak template (Table 5.6). The signal template is comprised of all the $t\bar{t}$ Monte Carlo samples, the multijet background template contains selected loose-tight tau events with subtracted electroweak (explained in Sec. 5.4.1) events and the electroweak template is composed of the events that pass $t\bar{t} \rightarrow \tau + jets$ selection cuts (Table 5.6). Moreover, the $t\bar{t}$, electroweak, and multijet templates are normalized to the number of events in the data.

To explain the general fitting procedure for these templates, it is convenient to use a characteristic variable X, where X is a top like event variable such as $\not\!\!E_T$ significance. The shape of the X distribution is given by the templates and the relative normalization is fitted.

The data events are expected to contain a small fraction of $t\bar{t} \rightarrow \tau + jets$ signal, electroweak background and a large fraction of multijet events. Therefore, the number of events in the ith bin of the X distribution for the data template can be written in terms of its specific components:

$$X(i)_{data} \equiv X(i)_{electroweak} + X(i)_{multijet} + X(i)_{t\bar{t}\rightarrow\tau+jets} \tag{6.1}$$

where, $X(i)_{electroweak}$ and $X(i)_{multijet}$ are the ith bins of the X distribution of electroweak and multijet templates and $X_{t\bar{t}_i}$ is the ith bin of X distribution of $t\bar{t} \rightarrow \tau + jets$ template. Equation 6.1 can be rewritten in terms of templates as follows:

$$\begin{aligned} X(i)_{data} \equiv\ & F(electroweak)X(i)^{template}_{electroweak} + F(multijet)X(i)^{template}_{multijet} \\ & + F(t\bar{t} \rightarrow \tau + jets)X(i)^{template}_{t\bar{t}\rightarrow\tau+jets} \end{aligned} \tag{6.2}$$

In order to fit signal and background fractions, a negative Log Likelihood function has been used. During the fitting procedure, a constraint $F(electroweak) +$

$F(multijet) + F(t\bar{t} \rightarrow \tau + jets) = 1$ and the electroweak components are kept fixed[1]. The negative Log Likelihood function is defined as

$$\begin{aligned} L(F(t\bar{t} \rightarrow \tau + jets)) &= -log\left(\prod_i \frac{\tilde{N}_i^{N_i^{obs}}}{N_i^{obs}!} e^{-\tilde{N}_i}\right) \\ &= -\sum_i \left(-N_i^{obs} log(\tilde{N}_i) + \tilde{N}_i\right) + constant \end{aligned} \tag{6.3}$$

where $\tilde{N}_i = F(electroweak)X(i)^{template}_{electroweak} + F(t\bar{t} \rightarrow \tau + jets)X(i)^{template}_{t\bar{t}\rightarrow\tau+jets} + (1 - F(electroweak) - F(t\bar{t} \rightarrow \tau + jets))X(i)^{template}_{multijet}$ is the number of events predicted in the ith bin and N_i^{obs} is the actual number of events observed in that bin.

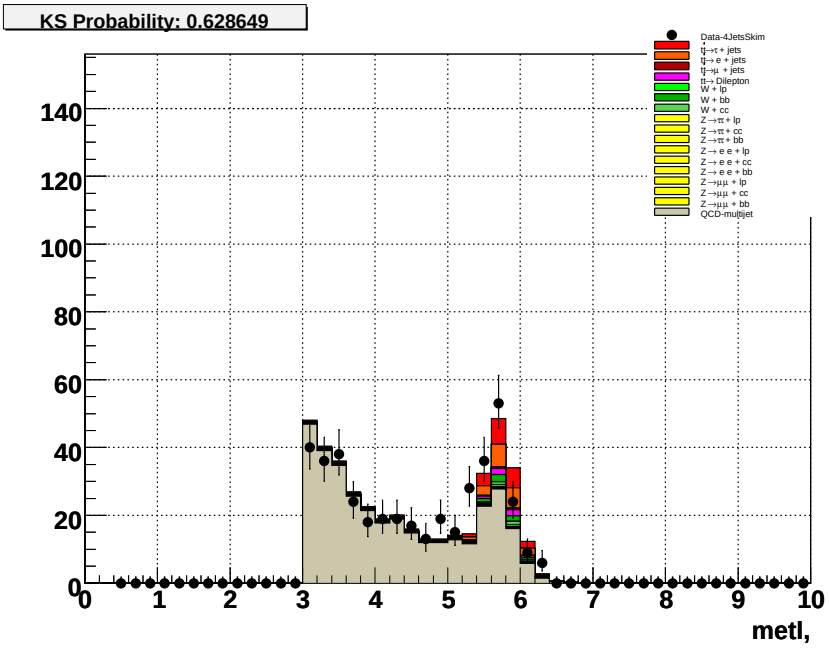

Figure 6.1: The fit result of multijet and $t\bar{t} \rightarrow \tau + jets$ templates to data candidate events using $\not\!\!E_T$ significance variable for tau Types I and II. Out of 414 data events, the number of $t\bar{t} \rightarrow \tau + jets$ events observed is 53.85 $^{+16.42}_{-15.93}$.

[1]The Monte Carlo electroweak events are scaled and a systematic uncertainty is included (Sec. 5.3.3).

Figure 6.1 represents a typical result of the fitting procedure which uses the $\not\!\!E_T$ significance as a $t\bar{t} \rightarrow \tau + jets$ variable. The uncertainty in the number of $t\bar{t} \rightarrow \tau + jets$ events obtained using this variable is fairly large ($\approx 30\%$) in comparison to the uncertainty in cross section measurement in all-jets channel which is approximately $\approx 17\%$ [76]. This is because the chosen variable does not separate signal from background very effectively. Moreover, a number of variables (as described in Sec. 6.1.2) are used for fitting but none perform well in extracting signal events from the data.

Several studies at DØ have shown that the Neural Network training can be very effective in separating signal events from background events and may drastically improve the precision of the measurement. Therefore, the fitting procedure is performed on templates made out of topological Neural Network Training output variables.

6.1.2 Selecting Variables

A number of top pair event variables are chosen to create the Neural Network Training input variables. For training, the variables that have strong discriminating power in separating signal from background are considered. In addition, the variables that use tau information are not selected as studies have shown that the tau related variables do not show good agreement with the data (appendix C.12). Therefore, only the variables that have strong discriminating power and do not use any tau information are selected. The selected variables and their definitions are listed below:

- ***H_T*** **:** This variable is defined as the scalar P_T sum of all jets and the tau candidates in an event.
- ***Sphericity and Aplanarity :*** These variables are formed from the eigenvalues of the normalized momentum tensor of the jets in the event. The value of these variables are expected to be larger in a top pair event than in a typical multijet event [66].

- ***Centrality :*** This variable is defined by $\frac{H_T}{H_E}$, where H_E is the sum of energies of all the jets in an event. The $t\bar{t} \rightarrow \tau + jets$ signal events are expected to be more central than the multijet events.

- ***Top and W mass likelihood (L) :*** This is a χ^2 like variable and is defined by $L \equiv \left(\frac{M_{3j}-m_t}{\sigma_t}\right)^2 + \left(\frac{M_{2j}-M_W}{\sigma_W}\right)^2$, where m_t and M_W are the top quark and W boson masses (170 GeV and 80 GeV respectively) and σ_t and σ_W are the top quark and W boson mass resolutions (45 GeV and 10 GeV respectively)[66]. The terms M_{3j} and M_{2j} are the jet invariant masses composed of the jet combinations. The combination that minimizes L is chosen.

- ***$\not\!\!E_T$ significance :*** The $\not\!\!E_T$ significance [77] is a measure of the likelihood of $\not\!\!E_T$ arising from physical sources (jets, muons, electrons and unclustered energy) rather than fluctuations in $\not\!\!E_T$ measurements itself as defined earlier in Sec. 5.2.3.

- ***Cos(θ*)*** **:** This variable is defined as the angle between the proton beam axis and the highest P_T jet in the rest frame of all the jets in the event.

- $\sqrt{(}$***S)*** **:** This variable is defined by the invariant mass of all jets the tau candidates in an event.

- ***ktminp :*** This variable is defined by $\Delta R^{min}_{j_1 j_2} P_T^{min}$, where $R^{min}_{j_1 j_2}$ is the minimum separation between a pair of jets out of all the jets in $\eta - \phi$ space, P_T^{min} is the lowest jet P_T of the $R^{min}_{j_1 j_2}$ pair.

6.1.3 Test of Multijet Modeling

As explained in the event selection procedure, the signal and the multijet modeling differ by a tau selection cut. The signal rich sample is selected by a tight tau selection cut and the multijet background is selected by loose-tight selection cut. Therefore, before doing such a modeling it is important to test whether the loose-tight sample is really dominated by multijet events.

To perform this test, a control sample dominated by multijet events is derived by keeping all the selection cuts the same as in Tables 5.3, B.2, B.1, but requiring

a tight b-tag veto instead of b-tag. The veto ensures that the data sample is dominated by non top like events. Table 6.1 demonstrates that such a sample is indeed dominated by multijet events (total number of expected $t\bar{t} \rightarrow \tau + jets$ event is approximately 40 and total number of expected electroweak events is approximately 304).

Samples	# passed scaled
data	4803±69.30
$t\bar{t} \rightarrow \tau + jets$	17.44 ± 0.26
$t\bar{t} \rightarrow e + jets$	18.16 ± 0.28
$t\bar{t} \rightarrow \mu + jets$	1.25 ± 0.07
$t\bar{t} \rightarrow l + l$	3.56 ± 0.05
$Wbb + jets \rightarrow l\nu + bb + jets$	10.5 ± 0.27
$Wcc + jets \rightarrow l\nu + cc + jets$	34.7 ± 0.97
$Wjj + jets \rightarrow l\nu + jj + jets$	159.6 ± 2.83
$Zbb + jets \rightarrow \tau\tau + bb + jets$	1.78 ± 0.05
$Zcc + jets \rightarrow \tau\tau + cc + jets$	5.54 ± 0.39
$Zjj + jets \rightarrow \tau\tau + jj + jets$	77.01 ± 1.45
$Zbb + jets \rightarrow ee + bb + jets$	0.32 ± 0.04
$Zcc + jets \rightarrow ee + cc + jets$	0.80 ± 0.13
$Zjj + jets \rightarrow ee + jj + jets$	12.58 ± 0.68
$Zbb + jets \rightarrow \mu\mu + bb + jets$	0.02 ± 0.01
$Zcc + jets \rightarrow \mu\mu + cc + jets$	0.07 ± 0.03
$Zjj + jets \rightarrow \mu\mu + jj + jets$	1.20 ± 0.15

Table 6.1: b-tag veto sample composition for Types I and II taus. The $t\bar{t} \rightarrow \tau + jets$ samples are scaled to 7 pb.

As mentioned earlier, for fitting, the variables listed in Sec. 6.1.2 are chosen as they are independent of tau information and have strong discriminating power. Figure 6.2 is a plot of the H_T distribution for the b-veto control samples. All other variable distributions for b-veto samples are in appendix C.2. Some of these plots give slight disagreement between loose-tight tau sample and the tight tau sample. In order to offset this disagreement, an associated systematic uncertainty for multijet modeling is incorporated in this analysis (Sec. 6.3).

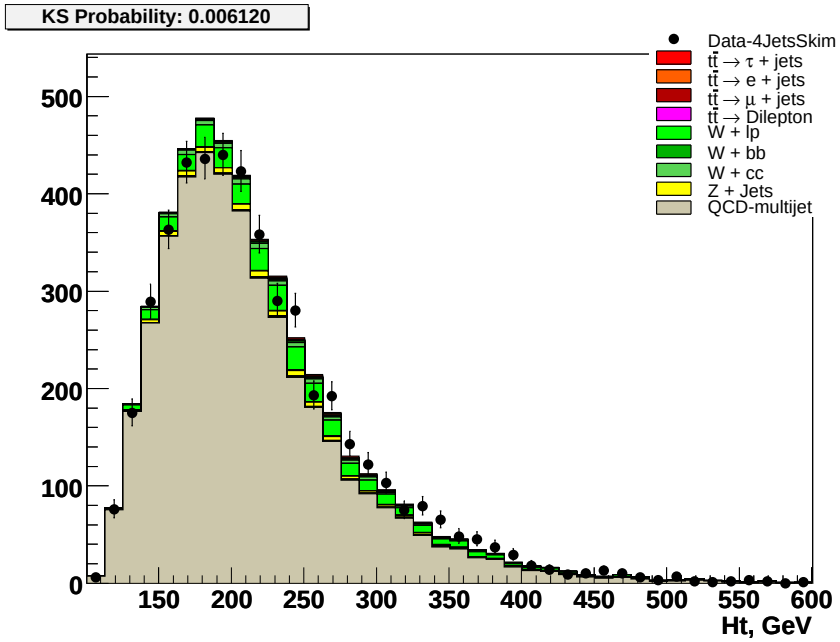

Figure 6.2: Histograms showing multijet domination of the b-veto samples. The H_T variable in "loose-tight" tau sample (the grey histogram) should agree with the tight tau sample (data points). Distributions of other variables are shown in appendix C.2.

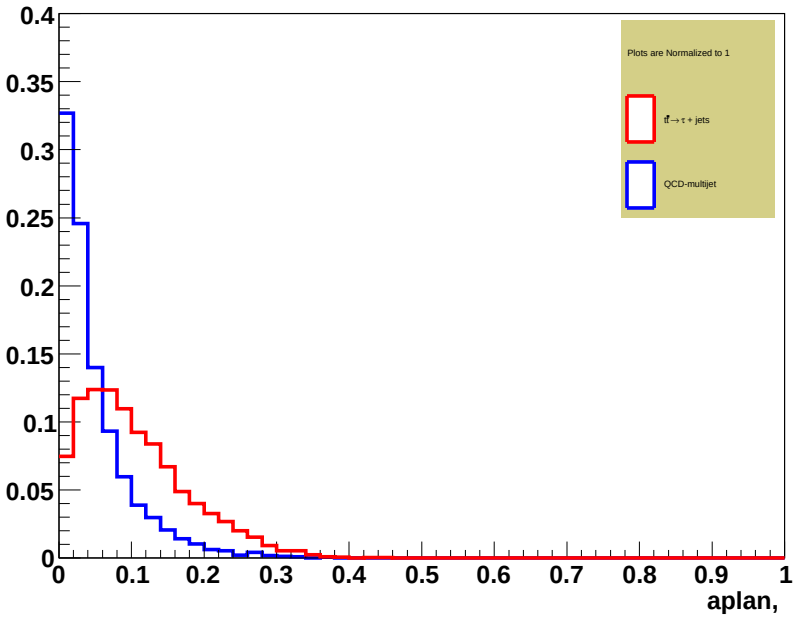

Figure 6.3: One of the Neural Network Training variables. The red plot is for $t\bar{t} \rightarrow \tau + jets$ Monte Carlo sample and the blue plot is for multijet background. The plot illustrates the discriminative power of the "aplan" variable. Other variables are shown in the appendix C.1.

6.1.4 Topological NN

As described in Sec. 6.1.1, in order to improve the estimate of signal events over a single variable fitting, a topological Neural Network (NN) Training is used. The NN incorporates the discriminating power of a set of top event variables and creates a new variable called NN_{topo}. For NN Training, the Multilayer Perceptron algorithm is used [85].

Training requires a separate background and a separate signal sample. For tau Types I & II training input samples, the part of the data that is not used for multijet modeling, (5.4) is used as background. A separate PYTHIA Monte Carlo sample (not used in signal template) of $t\bar{t} \rightarrow \tau + jets$ with signal selection cuts according to the Table B.2, is used as signal. Figure 6.3 represents the discriminating power of one of the training input variables. For tau Type III training inputs, the same data and Monte Carlo is used but with Type III selection instead of Types I & II samples.

Figure 6.4 shows the relative effect of each of the five NN_{topo} input variables ($\not\!\!E_T$ significance (metl), H_T, top mass likelihood (topmassl), Aplanarity (aplan) and $\sqrt{(}S)$), for this particular training, on the final output for tau Types I & II. Figures 6.5 and 6.6 represent NN_{topo} outputs for tau Types I & II and tau Type III respectively. In these figures, the high values of Topological Neural Network (NN_{topo}) correspond to signal-enriched region. For instance, a cut of 0.6 in NN_{topo} output predominantly selects $t\bar{t} \rightarrow \tau + jets$ events. This type of cut is not used for the cross section measurement in Sec. 6.2 but rather to check the agreement between data and prediction in the signal dominated region (appendix C.4 through C.6). In fact, to get a better statistical significance, the total range of the NN_{topo} is used for the cross section measurement.

6.1.5 NN optimization

As explained earlier, the variables are chosen in a way such that they are uncorrelated with the taus. However, this is not the only important factor for choosing the

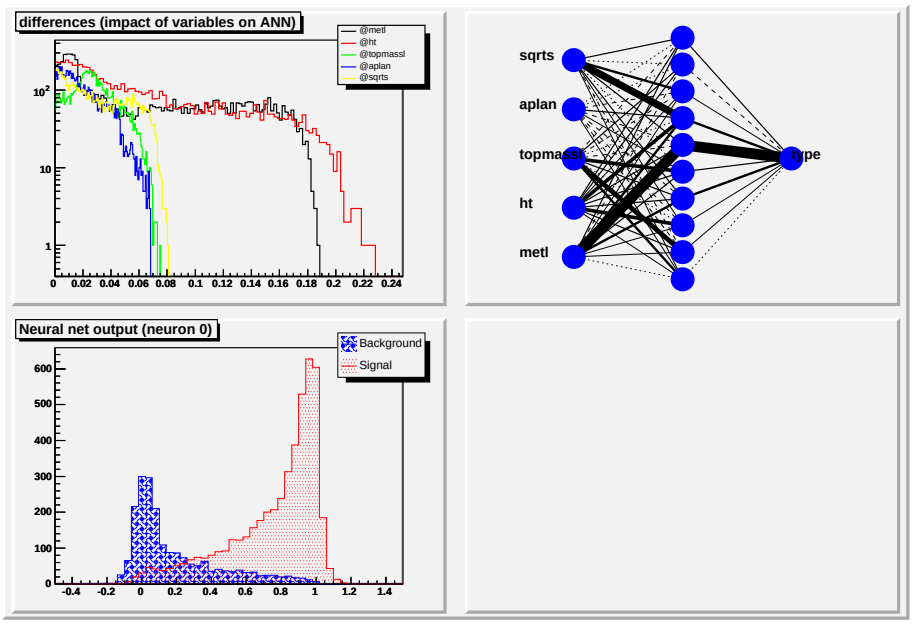

Figure 6.4: Training of Topological Neural Network output for tau Types I & II. The upper left plot demonstrates the relative impact of each of the input variables. The upper right plot shows the topological structure and the lower left plot illustrates signal-background separation strength.

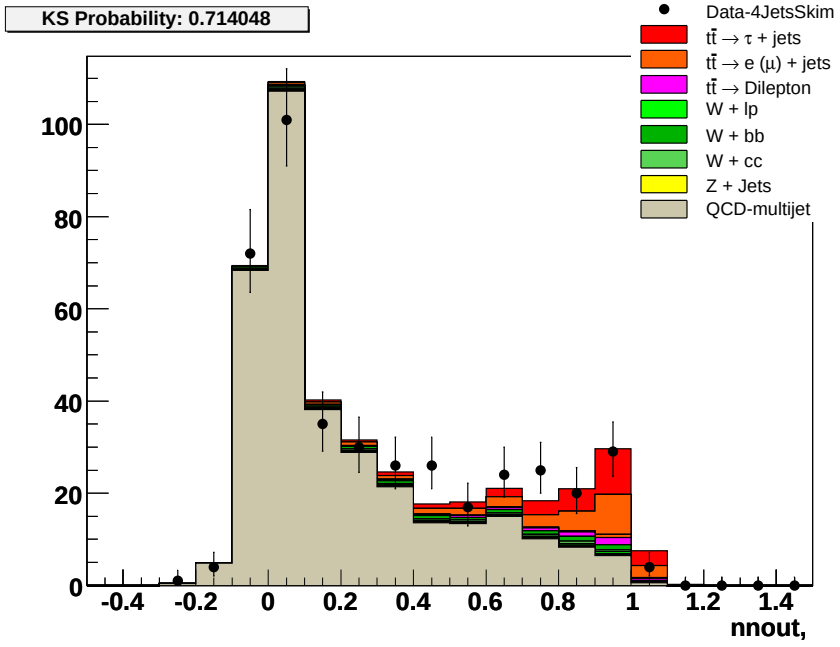

Figure 6.5: The Topological Neural Network NN_{topo} output for tau Types I & II.

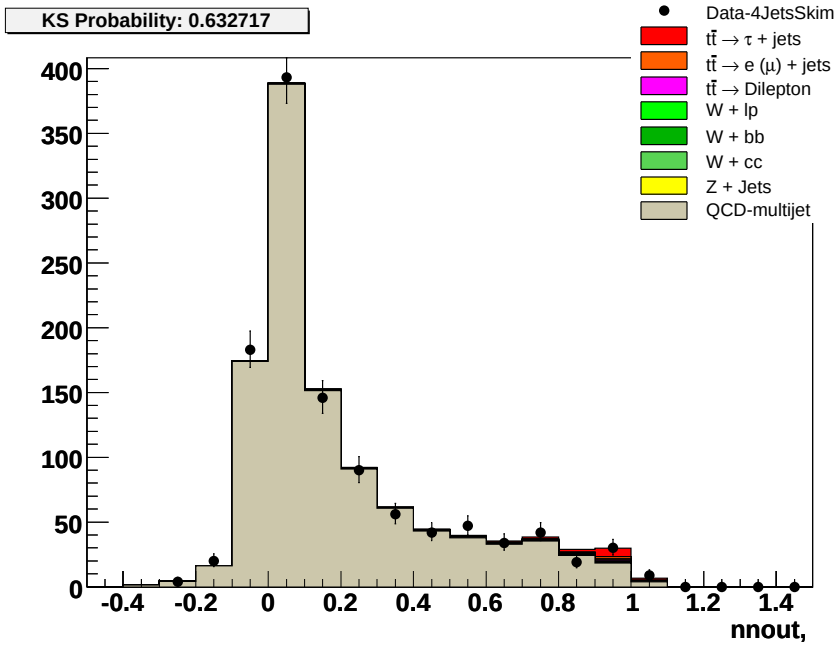

Figure 6.6: The Topological Neural Network NN_{topo} output for tau Type III.

topological Neural Network input variables. The variables that have large separation power between signal and background are also considered. Figure 6.3 and all the figures in appendix C.1 illustrate how different variables look for $t\bar{t} \rightarrow \tau + jets$ and multijet events.

For topological Neural Network training three sets of variables are created based on the following selection criteria:

- variables with high discriminating power in a set.
- choose the total number of variables in a set to five.
- $\not\!\!E_T$ significance (or *metl*) variable for all sets, since it is the most powerful discriminating variable.
- variables that are not highly correlated.
- only non-tau based variables.

Based on the above criteria, three sets of variables are chosen. They are as follows:

- ***Set I*** : *metl* ($\not\!\!E_T$ significance), H_T, top mass likelihood, aplanarity, sqrts ($\sqrt{S}$).
- ***Set II*** : metl, sphericity (spher), costhetastar ($cos(\theta^*)$), aplanarity (aplan), centrality (cent).
- ***Set III*** : metl, spher , sqrts, topmassl, ktminp.

In order to make the decision on which of the three sets of variables is the optimal, an ensemble of 1000 pseudo-datasets has been created each containing events randomly selected from multijet, electroweak and $t\bar{t} \rightarrow \tau + jets$ templates but with a set number of $t\bar{t} \rightarrow \tau + jets$ events. These data sets are considered as real data and the templates used for fitting these data sets are the same as described in Sec. 6.1.1. However, each of these pseudo-datasets is fitted in NN_{topo} variable instead of a single variable such as $\not\!\!E_T$ significance. The fitting mechanism is the same as explained earlier; the sum of fractions is set to unity ($F(electroweak) + F(multijet) + F(t\bar{t} \rightarrow \tau + jets) = 1$) and the electroweak components are kept fixed.

Figure 6.7 shows the fitted value of $t\bar{t} \rightarrow \tau + jets$ content obtained for 1000 different pseudo-datasets. The fitted value of $t\bar{t} \rightarrow \tau + jets$ events are in good agreement with the initially set numbers of signal events in the pseudo-datasets. The measured number of signal events are spread nicely around the predicted 59 events. The signal event prediction distribution is also a well fitting Gaussian distribution with the width corresponding to the statistical uncertainty of the $t\bar{t} \rightarrow \tau + jets$ measurement. Moreover, Fig. 6.8 shows the calibration curve for 1000 ensembles and the agreement is uniform for all the ensembles. The "pull", defined by $(N_{obs} - N_{preset})/\sigma(obs)$, in Fig. 6.9 has a well shaped Gaussian distribution which indicates that fit uncertainties are well behaved in the ensembles as well.

This ensemble test is used to compare different sets of topological variables. The decision on variable sets is made by measuring the "mistake" made in estimating the $t\bar{t} \rightarrow \tau + jets$ content in the pseudo-datasets. The "mistake" is defined by the difference between the observed and the initially preset number of top pairs. After comparing the mistakes for three different sets of variables (Figs. 6.10, 6.11

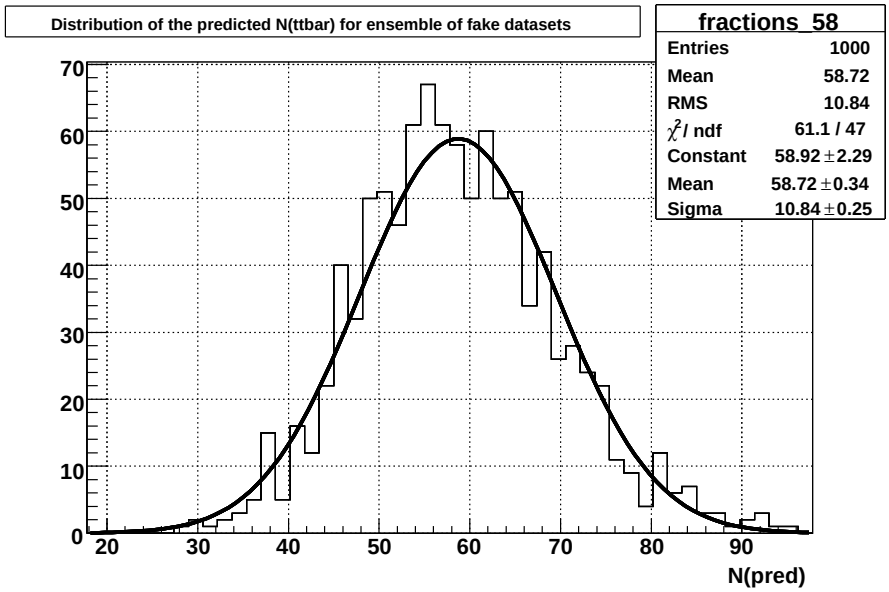

Figure 6.7: Distribution of the output of the ensemble test. The number of predicted $t\bar{t} \rightarrow \tau + jets$ events for tau Types I & II is plotted along the $x - axis$.

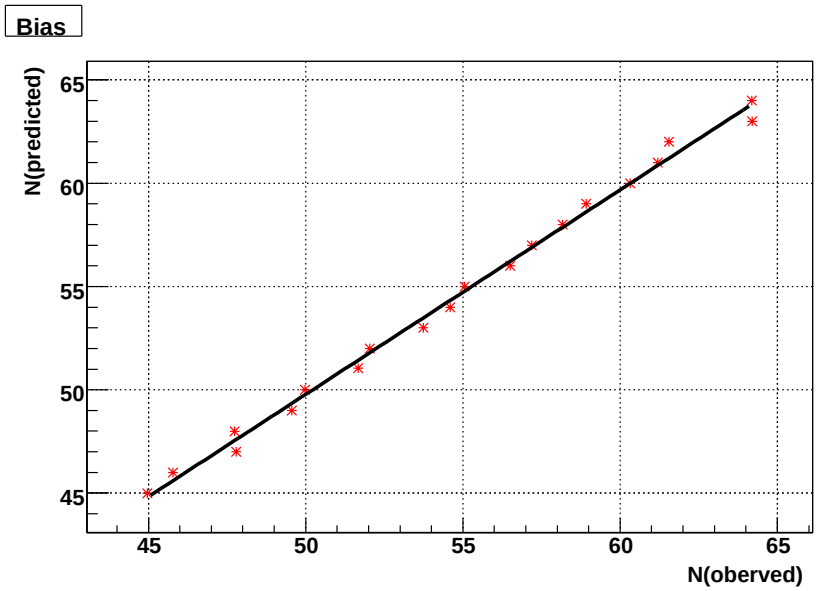

Figure 6.8: The ensemble test's bias plot. The number of observed events are plotted along the $x - axis$ and the preset number of $t\bar{t} \rightarrow \tau + jets$ events in the ensemble are plotted along the $y - axis$.

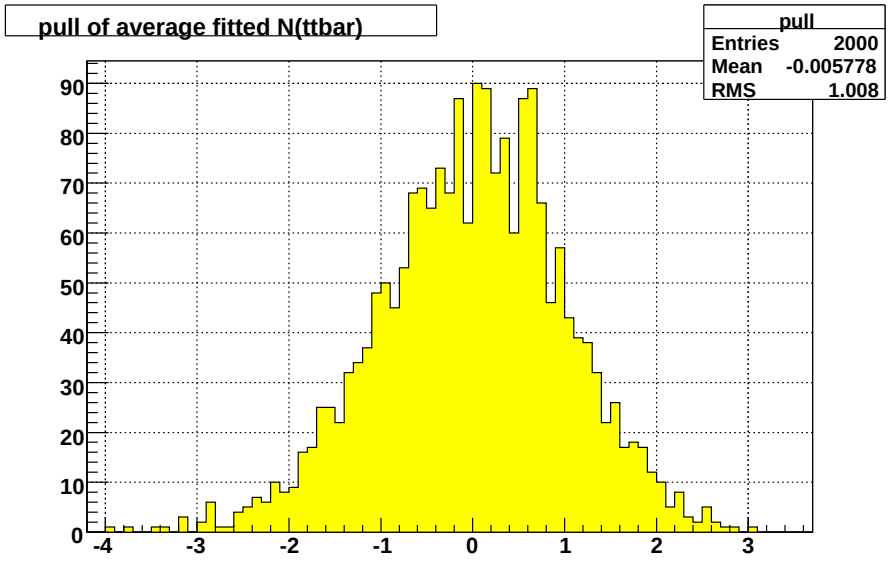

Figure 6.9: The ensemble test's pull $((N_{obs} - N_{preset})/\sigma(obs))$ distributions along $x - axis$.

and 6.12), it is found that set I has the smallest uncertainty. Therefore, set I variables are chosen for the analysis and the output NN_{topo} is used to perform the measurement of $t\bar{t} \rightarrow \tau + jets$ cross section.

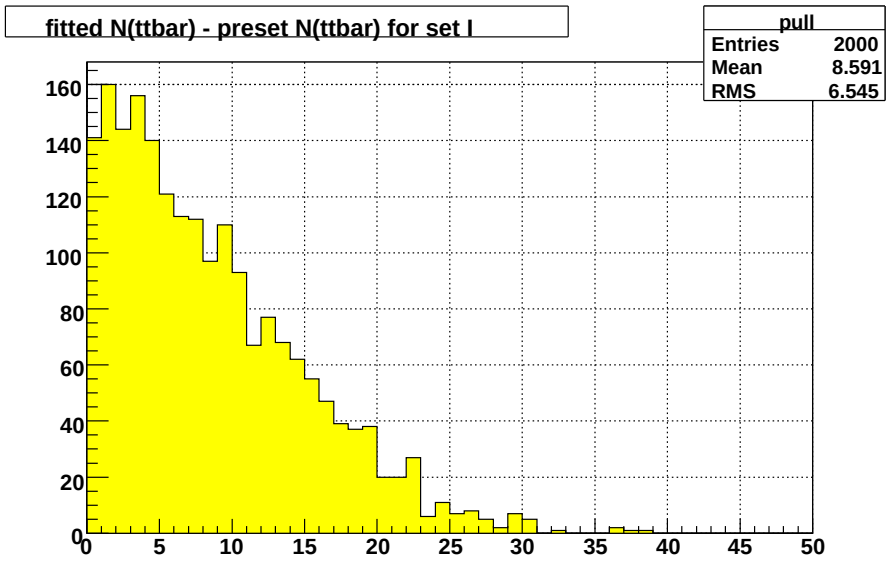

Figure 6.10: Set I average "mistake" which is defined by absolute value of the difference between observed and initially set number. The plot is for tau Types I and II.

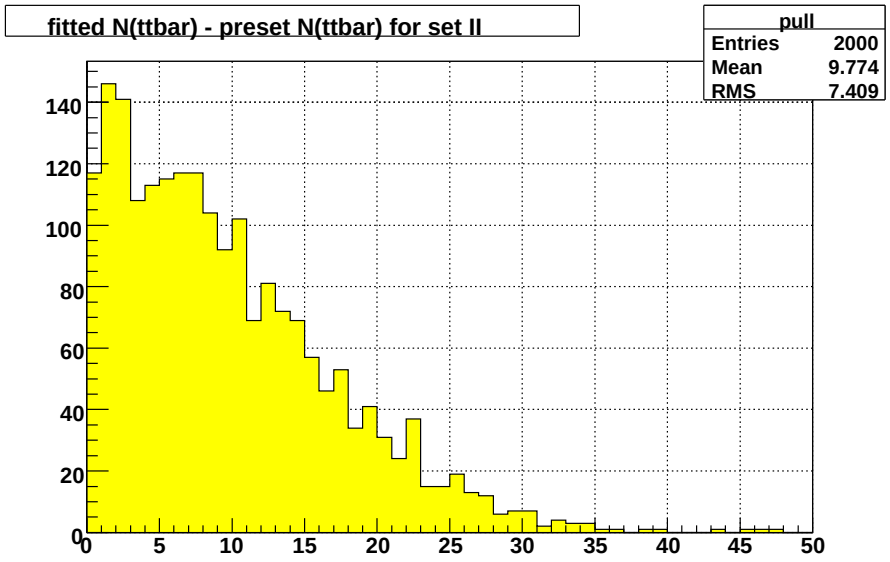

Figure 6.11: Set II average "mistake" (plotted along $x-axis$) which is defined by the absolute value of the difference between observed and initially set number of events. The plot is for tau Types I and II.

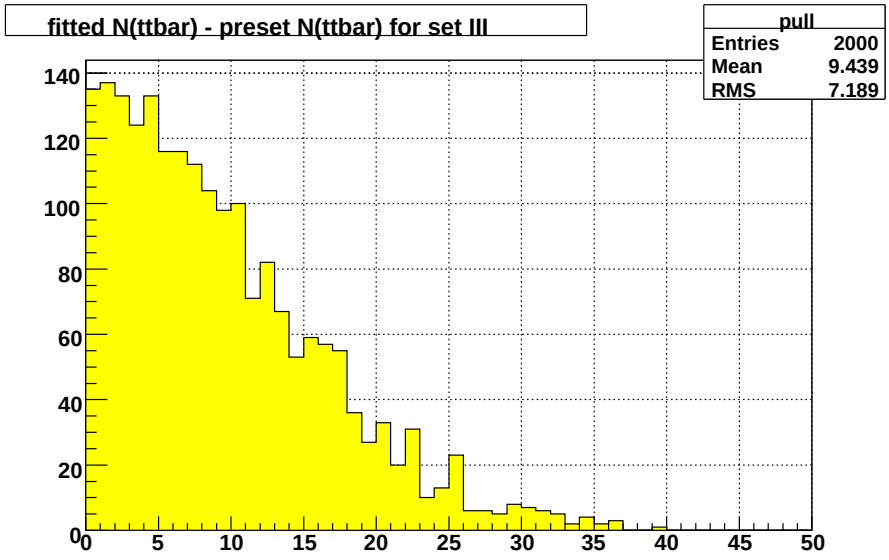

Figure 6.12: Set III average "mistake" (plotted along $x - axis$) which is defined by the absolute value of the difference between observed and initially set number of events. The plot is for tau Types I and II.

6.2 The Cross Section Measurement

In the previous section, all the tools necessary for the cross section measurement have been developed. Data have been selected, the multijet contribution has been modeled and Monte Carlo samples (both $t\bar{t}$ and electroweak) have been processed to extract the number of $t\bar{t} \rightarrow \tau + jets$ events which is used for the cross section measurement. The cross section is defined as

$$\sigma(t\bar{t} \rightarrow \tau + jets) = \frac{Number\ of\ signal\ events}{\epsilon(t\bar{t} \rightarrow \tau + jets) \times BR(t\bar{t} \rightarrow \tau + jets) \times \mathcal{L}} \tag{6.4}$$

where the input parameters are explained as follows:

- ***Number of signal events :***

  In order to extract the number of signal events, the Monte Carlo templates, the electroweak template and the multijet template are fitted in data using the same fitting procedure as explained in Sec. 6.1.1 but using the NN_{topo} variable instead of a single variable. The fit constraints are $F(electroweak) + F(multijet) + F(t\bar{t} \rightarrow \tau + jets) = 1$ and fixed electroweak components. The result of fitted number of signal ($t\bar{t} \rightarrow \tau + jets$) events are shown in the Table 6.2.

- $\boldsymbol{\epsilon(t\bar{t} \rightarrow \tau + jets)}$: The selection efficiency $\boldsymbol{\epsilon(t\bar{t} \rightarrow \tau + jets)}$ is calculated as the ratio of the initial number of scaled events to the final number of scaled events (before scaling)(Tables B.2, B.3, B.4, B.5).

$$\begin{aligned}
\boldsymbol{\epsilon(t\bar{t} \rightarrow \tau + jets)} &= \frac{Final\ Number\ of\ Scaled\ Events}{Initial\ Number\ of\ Scaled\ Events} \\
&= \frac{\boldsymbol{t\bar{t} \rightarrow \tau/e/\mu/dilepton + jets}}{\boldsymbol{((t\bar{t} \rightarrow lepton + jets) + (t\bar{t} \rightarrow dilepton + jets))}} \\
&= \frac{\boldsymbol{(27.20 + 25.87 + 1.92 + 5.44)}}{\boldsymbol{1.2 \times (375408 \times 0.0065 + 532750 \times 0.001586)}} \\
&= \boldsymbol{0.0153295}
\end{aligned} \tag{6.5}$$

Templates	***Events Types I & II***	***Events Type III***
Data	414 ± 20.35	1115 ± 33.39
Signal $t\bar{t} \rightarrow \tau + jets$ ***Events***		
Total Fitted	$\mathbf{60.48 \pm 11.84}$	$\mathbf{23.96 \pm 11.37}$
Efficiency, $\boldsymbol{\epsilon(t\bar{t} \rightarrow \tau + jets)}$	0.0153295	0.00860355
Background		
Multijet		
Total Fitted	$\mathbf{336.68 \pm 11.84}$	$\mathbf{1083.19 \pm 11.37}$
Electroweak		
$\boldsymbol{W + jets}$	13.48 ± 0.31	5.92 ± 0.21
$\boldsymbol{Z + jets}$	3.35 ± 0.37	1.93 ± 0.06

Table 6.2: Summary of the fitted number of events and the corresponding efficiency for cross section measurement in the $t\bar{t} \rightarrow \tau + jets$ channel.

- $\boldsymbol{BR(t\bar{t} \rightarrow \tau + jets)}$ **:** The branching ratio (BR) for this channel is calculated from the reference [6] and is shown in Table 6.3.

$\boldsymbol{BR(t\bar{t} \rightarrow \tau + jets)}$ **channel**	
$\boldsymbol{t\bar{t} \rightarrow \tau_{had} + jets}$	0.097
$\boldsymbol{t\bar{t} \rightarrow e + jets}$, *includes* $\boldsymbol{t\bar{t} \rightarrow \tau \rightarrow e + jets}$	0.171
$\boldsymbol{t\bar{t} \rightarrow \mu + jets}$, *includes* $\boldsymbol{t\bar{t} \rightarrow \tau \rightarrow \mu + jets}$	0.171
$\boldsymbol{t\bar{t} \rightarrow dilepton + jets}$	0.106
Total	0.545

Table 6.3: Branching Ratios for different top decay channels.

- $\boldsymbol{\mathcal{L}}$ **:** The total integrated luminosity calculation is shown in Table 5.2 and is $\boldsymbol{974.2 pb^{-1}}$.

With these inputs the cross section can be measured. For example, for $t\bar{t} \rightarrow \tau + jets$ cross section for tau Types I & II is

$$\sigma(t\bar{t} \rightarrow \tau + jets) = \frac{60.48}{0.0153295 \times 0.545 \times 974.2} \approx 7.4 pb \tag{6.6}$$

However, for precision, the cross section is measured by utilizing the entire range of NN_{topo} output and by minimizing the sum of the negative Log Likelihood functions for each bin of both tau Types I & II and tau Type III. These Log Likelihood functions are used to perform fits shown in Figs. 6.5 and 6.6. For the cross section measurement the function is expressed in terms of $\sigma(t\bar{t} \rightarrow \tau + jets)$:

$$L(\sigma; \tilde{N}_i, N_i^{obs}) \equiv -log(\prod_i \frac{\tilde{N}_i^{N_i^{obs}}}{N_i^{obs}!} e^{-\tilde{N}_i}) \tag{6.7}$$

where $\tilde{N}_i = \sigma \times BR \times \mathcal{L} \times \epsilon_i(t\bar{t} \rightarrow \tau + jets) + N_{bkg}$ is the number of events predicted in the ith bin of the data in NN_{topo} distribution and N_i^{obs} is the actual number of events observed in that bin. The Log Likelihood functions are displayed in Figs. 6.13 and 6.14. The minimum value along the horizontal axis in each of the graphs is the measured cross section.

The cross section results are summarized as follows:

$t\bar{t} \rightarrow \tau + jets$ **tau Types I & II cross section:**

$$\sigma(t\bar{t}) = 7.4 \ ^{+1.4}_{-1.4} \ (\text{stat}) \ ^{+0.9}_{-0.8} \ (\text{syst}) \ \pm 0.4 \ (\text{lumi}) \ \text{pb}$$

$t\bar{t} \rightarrow \tau + jets$ **tau Type III cross section:**

$$\sigma(t\bar{t}) = 5.2 \ ^{+2.5}_{-2.2} \ (\text{stat}) \ ^{+1.0}_{-1.0} \ (\text{syst}) \ \pm 0.4 \ (\text{lumi}) \ \text{pb}$$

The combined cross section is obtained from the product of the likelihoods in Figs. 6.13 and 6.14. The combined cross section for $t\bar{t} \rightarrow \tau + jets$ with tau Types I, II, & III is

$$\sigma(t\bar{t}) = 6.9 \ ^{+1.2}_{-1.2} \ (\text{stat}) \ ^{+0.8}_{-0.7} \ (\text{syst}) \ \pm 0.4 \ (\text{lumi}) \ \text{pb}. \qquad (6.8)$$

6.3 Systematic Uncertainties

A number of factors contribute to systematic uncertainties in the cross section measurement. This section describes the different sources of uncertainties related to the cross section measurement. For most of the cases the uncertainty is calculated by varying the source by $\pm 1\sigma$. The effects of the systematic uncertainties on the final measurement are listed in the Table 6.4.

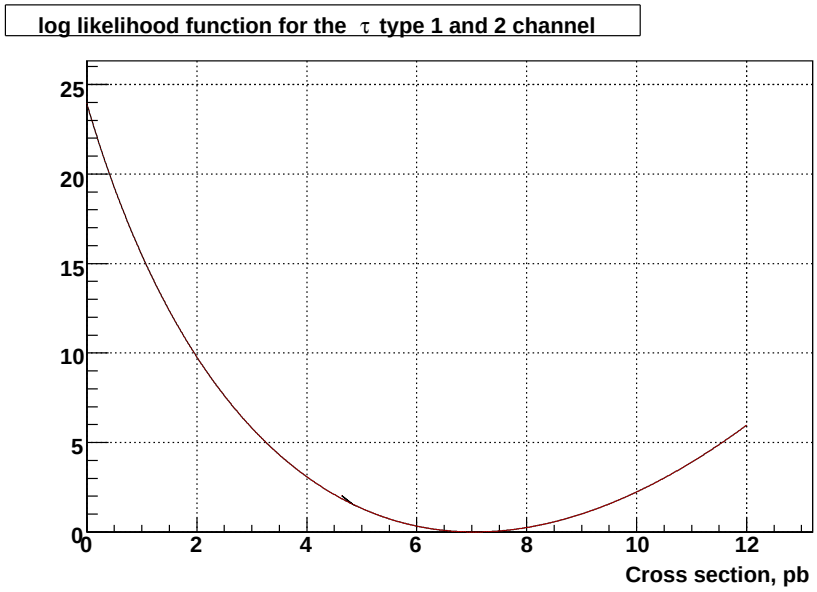

Figure 6.13: The log likelihood function for tau Types I & II.

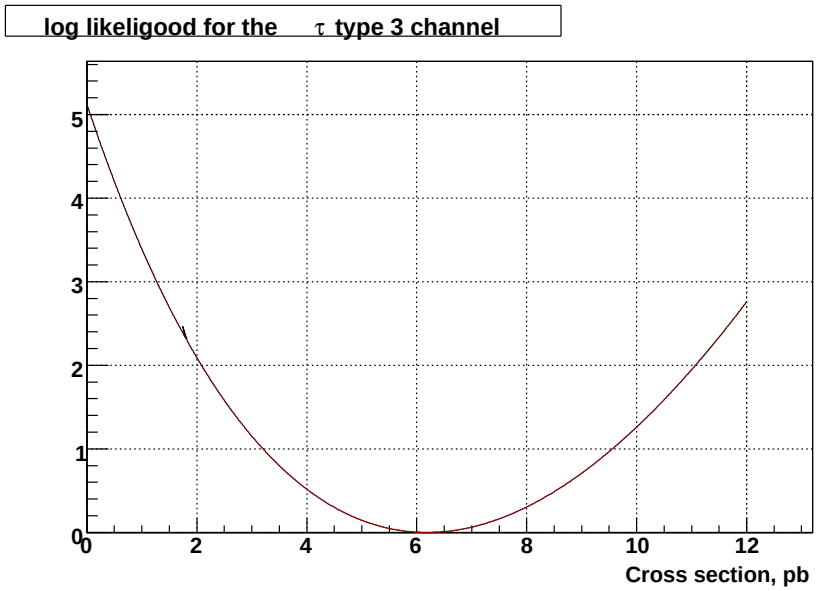

Figure 6.14: The log likelihood function for tau Type III.

JES The jets energy scale (JES) uncertainty is measured by shifting the jet energy scale by $\pm 1\sigma$ in all Monte Carlo samples.

JER The jet energy resolution (JER) systematic uncertainty is obtained by varying JER by $\pm 1\sigma$.

Jet ID The jet identification efficiency is slightly higher in the Monte Carlo than in the data. This is because the calorimeter clusters that fail the standard jet criteria are randomly removed from the Monte Carlo. To account for this difference the Jet ID systematic is introduced.

Trigger Uncertainties associated with the turn-on curves which are used for trigger probability calculation are propagated to the trigger systematic uncertainty calculation [89].

Tau-jet Triggering The ALLJETS triggers used for this analysis are designed for jets. But it is quite possible for a tau to pass a jet trigger. This may lead to a difference in the jet trigger turn-on used for this analysis. However, as demonstrated in appendix C.7 the difference for this in the trigger turn-on is found to be small. A systematic uncertainty is estimated in appendix C.7 to account for this difference.

***b*-tagging** The b-tagging uncertainty effects are taken into account by varying the systematic and statistical errors on the Monte Carlo tagging weights. These errors are computed using the standard DØ b-ID group tools [50].

Tau ID The uncertainty associated with taus is found by fluctuating each of the NN_τ input variables within the statistics and observing the effect it has on the NN_τ output. This procedure is detailed in note [87], according to which the resulting systematic uncertainty is about 5%.

QCD systematics As explained in the Sec. 6.1.3, there is a slight disagreement between the loose-tight tau sample and the tight tau sample . In order to account for this disagreement a systematic uncertainty is introduced by scaling the multijet template (loose-tight tau) such that it matches the tight

tau data exactly (with electroweak backgrounds subtracted). The variable used for this scaling is NN_{topo} variable. Figure 6.15 shows that this scaling is close to one, as expected, since the multijet template models the multijet dominated data very well. This scaling is applied to the signal sample's multijet template and measures the cross section. This gives an estimate of the systematic uncertainty on the multijet background modeling (Table 6.4).

W scale factor and Heavy Flavor (HF) scale factor In Sec. 5.3.3, the scaling factor and the heavy flavor scale factors for the electroweak backgrounds are described. These scaling factors have uncertainties that are propagated to the cross section measurement.

Template statistics The multijet template used in the fitting procedure to extract the $t\bar{t} \rightarrow \tau + jets$ signal events as well as the multijets events from data, have limited statistics (1529 events for Types I & II tau and 4609 events for Type III tau). A systematic uncertainty calculated by varying the content of each bin of the multijet template distribution within its uncertainty to account for this limited statistics.

PDF As mentioned in Sec. 5.3 the parton distribution function (PDF) set CTEQ 6.1M [68] is used in Monte Carlo simulation. These functions have associated uncertainties that have to be included as systematic uncertainty in the cross section. However, as demonstrated in [65] (which uses the same Monte Carlo samples as used in this analysis and have very similar kinematic distributions) these uncertainties are less then 0.5%. Thus a conservative estimate of 0.5% is applied.

Luminosity For the luminosity measurement uncertainty, the DØ standard value of 6.1% is used for this analysis [65].

Table 6.4 summarizes all of the sources of uncertainties and show their effect on the cross section measurement.

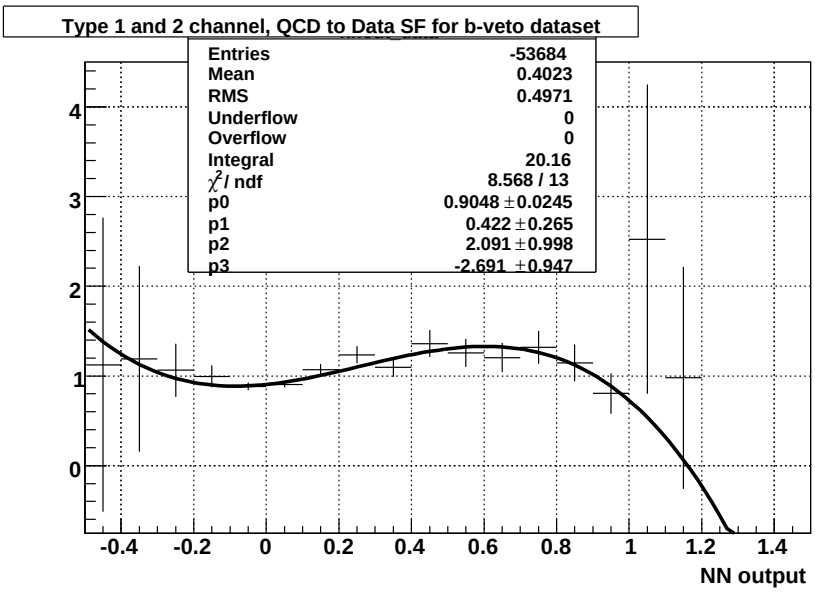

Figure 6.15: The QCD/multijet scaling factor.

Channel	τ+jets Types I & II	τ+jets Type III	Combined
Jet Energy Scale	$-0.396,+0.408$	$-0.276,+0.264$	$-0.372,+0.372$
Jet Energy Resolution	$+0.000,-0.000$	$+0.000,-0.012$	$+0.000,-0.000$
Jet ID	$+0.180,-0.000$	$+0.132,-0.000$	$+0.168,-0.000$
b-tag	$+0.420,-0.372$	$+0.264,-0.240$	$+0.384,-0.336$
Trigger	$-0.012,+0.180$	$-0.012,+0.108$	$-0.012,+0.168$
QCD Modeling	$+0.084,-0.084$	$+0.012,-0.024$	$+0.060,-0.060$
τ ID	$+0.372,-0.372$	$+0.252,-0.264$	$+0.336,-0.348$
τ triggering	$+0.192,-0.000$	$+0.120,-0.000$	$+0.168,-0.000$
W Scale Factor	$+0.072,-0.072$	$+0.036,-0.048$	$+0.072,-0.072$
HF scale factor	$+0.108,-0.108$	$+0.060,-0.072$	$+0.096,-0.108$
Template statistics	$+0.401,-0.401$	$+0.870,-0.870$	$+0.361,-0.361$
PDF	$+0.036,-0.036$	$+0.024,-0.036$	$+0.036,-0.036$

Table 6.4: Source of systematic uncertainties on $\sigma(t\bar{t} \rightarrow \tau + jets)$ (in pb).

6.4 Summary

This dissertation presents the first measurement of $\sigma(t\bar{t} \rightarrow \tau + jets)$ by the DØ collaboration using RunIIa data provided by the Tevatron at Fermilab. The decay channel studied involves one hadronically decaying tau lepton, two b jets, two light jets and $\not\!\!E_T$. The ALLJETS triggers are used and the corresponding data size is $974.2 \pm 57.8 pb^{-1}$. There were two major challenges in extracting $t\bar{t} \rightarrow \tau + jets$ events from data: first the multijet backgrounds had to be modeled and second the electroweak backgrounds needed to be properly estimated. To overcome these challenges, NN_τ cuts and b-tagging algorithms are employed. In addition, the relevant topological variables are combined into a NN_{topo} variable which differentiates the signal from multijet background. For the cross section measurement, the Likelihood fit is used in the NN_{topo} variable. The fitting in the NN_{topo} variable shows $\approx$ 13% improvement over single variable fitting. In order to show the relative significance of the measurements, the single and triple pronged tau lepton decay modes are treated as two independent channels and then combined. Moreover, the $t\bar{t} \rightarrow \tau + jets$ channel is made explicitly orthogonal to all previously measured channels, so that it can easily be combined with other DØ top cross section measurement channels.

The cross section measured is

$$\sigma(t\bar{t} \rightarrow \tau + jets) = 6.9 \ ^{+1.2}_{-1.2} \ (\text{stat}) \ ^{+0.8}_{-0.7} \ (\text{syst}) \ \pm 0.4 \ (\text{lumi}) \ \text{pb}.$$

This result is consistent with the theoretically predicted value of $7.46^{+0.90}_{-0.62}$ pb for top mass 170 GeV [15] as well as other DØ and CDF measurements [66, 65]. In addition, two more top mass points such as 165 GeV and 175 GeV are explored with relevant Monte Carlo samples. For the cross section measurements the same procedure that has been used for top mass 170 GeV is used. The measured cross section dependence on the top quark mass also agrees with the theoretical [15] value (Table 6.5).

	Types I & II channel	Type III channel	Combined	Theoretical
Top mass 165 GeV	8.8±1.8	6.0±3	8.1±1.5	8.7±1.1
Top mass 175 GeV	6.8±1.4	4.8±2.3	6.3±1.15	6.4±0.8

Table 6.5: The cross section dependence on the top quark mass. The cross sections for three different masses for tau Types I & II and tau Type III. The uncertainties are for total statistical and systematic.

In addition the $t\bar{t}$ cross section is measured using only $t\bar{t} \rightarrow \tau_{hadronic} + jets$ decay mode. The measurement procedure is the same as for $t\bar{t} \rightarrow \tau + jets$ and other $t\bar{t}$ decays ($t\bar{t} \rightarrow e + jets$, $t\bar{t} \rightarrow e + jets$, $t\bar{t} \rightarrow dilepton + jets$) in Table 5.6 are treated as part of the background. The results of the fit are shown below:

- For tau Types I and II : the fitted number of $t\bar{t} \rightarrow \tau_{hadronic} + jets$ events is $25.1\ ^{+11.2}_{-10.5}$ and fitted number of multijet events is $336.4\ ^{+11.16}_{-10.5}$
- For tau Type III : the fitted number of $t\bar{t} \rightarrow \tau_{hadronic} + jets$ events is $18.0\ ^{+11.3}_{-10.3}$ and fitted number of multijet events is $1076.0\ ^{+11.3}_{-10.3}$

The branching ratio of $t\bar{t} \rightarrow \tau_{hadronic} + jets$ is $0.65 \times 0.15 = 0.1$. Therefore, in the $BR \times \sigma$ format the result can be written as follows:

The cross section for tau Types I & II:

$$\sigma(t\bar{t} \rightarrow \tau_{hadronic} + jets) = 0.71\ ^{+3.1}_{-2.9}\ (\text{stat})\ ^{+1.9}_{-1.71}\ (\text{syst})\ \pm 0.4\ (\text{lumi})\ \text{pb},$$

The cross section for tau Type III:

$$\sigma(t\bar{t} \rightarrow \tau_{hadronic} + jets) = 0.43\ ^{+3.8}_{-3.5}\ (\text{stat})\ ^{+1.9}_{-1.9}\ (\text{syst})\ \pm 0.4\ (\text{lumi})\ \text{pb}.$$

The combined cross section is

$$\sigma(t\bar{t} \rightarrow \tau_{hadronic} + jets) = 0.60\ ^{+2.3}_{-2.2}\ (\text{stat})\ ^{+1.6}_{-1.4}\ (\text{syst})\ \pm 0.4\ (\text{lumi})\ \text{pb}.$$

The breakdown of the corresponding systematic uncertainties on the cross section are shown in Table 6.6. There is an additional source of uncertainty due to normalization of non-tau $t\bar{t}$ background to $t\bar{t}$ theoretical cross section. The resulting cross section turns out to be fairly close to the value predicted by the Standard Model.

Channel	τ+jets Types I & II	τ+jets Type III	Combined
Jet Energy Scale	-0.805,$+0.829$	-0.481,$+0.481$	-0.697,$+0.709$
Jet Energy Resolution	$+0.000$,-0.000	$+0.000$,-0.000	$+0.000$,-0.000
Jet ID	$+0.384$,-0.000	$+0.240$,-0.000	$+0.336$,-0.000
b-tag	$+0.865$,-0.757	$+0.420$,-0.360	$+0.697$,-0.613
Trigger	-0.024,$+0.384$	-0.000,$+0.192$	-0.024,$+0.312$
QCD Modeling	$+0.084$,-0.084	$+0.024$,-0.012	$+0.048$,-0.048
τ ID	$+0.360$,-0.360	$+0.216$,-0.204	$+0.300$,-0.300
τ triggering	$+0.384$,-0.000	$+0.204$,-0.000	$+0.312$,-0.000
W Scale Factor	$+0.156$,-0.156	$+0.072$,-0.060	$+0.120$,-0.120
HF scale factor	$+0.228$,-0.228	$+0.108$,-0.096	$+0.180$,-0.192
Template statistics	$+0.921$,-0.921	$+1.732$,-1.732	$+0.848$,-0.848
$t\bar{t}$ cross section	$+0.697$,-0.661	$+0.312$,-0.036	$+0.060$,-0.060
PDF	$+0.072$,-0.072	$+0.036$,-0.036	$+0.060$,-0.060

Table 6.6: Systematic uncertainties calculated for $\sigma(t\bar{t} \rightarrow \tau_{hadronic} + jets)$ (in pb).

Bibliography

[1] S. Weinberg, A Model of Leptons, Phys. Rev. Lett. 19 (1967), 1264-1266.
S.L. Glashow, Nuclear Phys. 22 (1961), 579.
A. Salam and J.C. Ward, Phys. Lett. 13 (1964), 168.

[2] S. Abachi et al., Observation of the top quark, Phys. Rev. Lett. 74 (1995), 2632-2637, hep-ex/9503003.

[3] F. Abe et al., Obersvation of top quark production in $p\bar{p}$ collisions, Phys. Rev. Lett. 74 (1995), 2626-2631, hep-ex/0612052.

[4] David Griffiths, Introduction to Elementary Particles, Wiley & Sons, Inc.

[5] Michael E. Peskin & Daniel V. Schroeder, An introduction to Quantum Field Theory.

[6] C. Amsler et al. (Particle Data Group), Physics Letters B667, 1 (2008) and 2009 partial update for the 2010 edition.

[7] D. Schaile and P.M. Zerwas, Phys. Rev. **D45** 3262 (1992).

[8] J. Khun, The Top Quark and The Electroweak Interactions, hep-ph/9707321.

[9] S. Dawson, The Top Quark, QCD, and New Physics.

[10] M. Kobayashi and K.Masakawa, Prog. Theor. Phys.**49**,652(1973).

[11] DØ Collaboration Phys. Rev. **D60**, 052001(1999).

[12] DØ Collaboration Phys. Rev. **D58**, 052001(1998).

[13] DØ Collaboration Phys. Lett. **B606**, 25 (2005). hep-ex/0410086; Fermilab-Pub-04/305-E.

[14] C. Bhat et. al., hep-ex/9809011, Fermilab-Pub-98/236.

[15] Theoretical top quark cross section at the Fermilab Tevatron and the CERN LHC, N. Kidonakis and R. Vogt, Phys. Rev. D 78, 074005.

[16] J. A. Coarasa, J. Guasch, and J. Sola, Report No. UAB-FT-451, 1999, hep-ph/9903212.

[17] V. M. Abazov et al. The upgraded DØ detector. Nucl. Instrum. Meth., A565:463537, 2006.

[18] J. Hays et. al. Single electron efficiencies in p17 data and monte-carlo. DØ Note 5025.

[19] Gustavo Otero y Garzon, Measurement of the top quark pair production cross section in pbar-p collisions at $\sqrt{s} = 1.96$ TeV, Ph. D. thesis (2006).

[20] Eric H. Smith, Search for Charged Higgs Boson in Top Quark Decays at DØ, Ph. D. thesis (1999).

[21] Mikhail Arov, Measurement of the top pair production cross section in the tau channel, Ph. D. thesis (2007).

[22] Satish Desai, A Search for the Production of Technicolor Particles at the DØ Detector, Ph. D. thesis (2006).

[23] David Wai Kui Lam, Measurement of Top Quark Mass in the All Hadronic Channel, Ph. D. thesis (2008).

[24] Haryo Sumowidagdo, First Measurement of Top Quark Pair Production Cross Section in Muon Plus Hadronic Tau Final States, Ph. D. thesis (2007.

[25] Thomas Gadfort, Evidence For Electroweak Top Quark Production In Proton-Antiproton Collisions, Ph. D. thesis (2007).

[26] Ken Herner, Search for neutral Higgs bosons decaying to tau pairs produced in association with b-quarks, Ph. D. thesis (2008).

[27] Mark Owen, Search for Higgs Bosons Decaying into Tau Pairs, Ph. D. thesis (2008).

[28] Jeroen Hegeman, Measurement of the top quark pair production cros section;
Hadronic top decays with the DØ detector, Ph. D. thesis (2009).

[29] Dag Gillberg, Discovery of Single Top Quark Production, Ph. D. thesis (2009).

[30] T. Gadfort et. al. Muon identification certification for p17 data. DØ Note 5157.

[31] D0 Collaboration: d0reco. 2004.
http://www-d0.fnal.gov/Run2Physics/WWW/algorithm.htm

[32] A. Khanov, HTF: histogramming method for finding tracks. The algorithm description DØ Note 3778, 2000.

[33] Ordering a Chaos or... Technical Details of AA Tracking, G. Borissov, All DØ Mtg (28 Feb 2003),
http://www-d0.fnal.gov/atwork/adm/d0_private/2003-02-28/adm_talk.ps

[34] H. Greenlee, The DØ Kalman Track Fit. DØ Note 4303, 2004.

[35] A. Schwartzman and C. Tully, Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting. DØ Note 4918, 2005.

[36] A. Schwartzman and M. Narain, Probabilistic Primary Vertex Selection. DØ Note 4042, 2002.

[37] Primary Vertex Reconstruction in p14, A. Garcia-Bellido et al., DØ Note 4320 (7 Jan. 2004).

[38] G. C. Blazey, J. R. Dittmann, S. D. Ellis, V. D. Elvira, K. Frame, G. S., R. Hirosky, R. Peigaia, H. Schellman, R. Snihur, V. Sorin, and D. Zeppenfeld, Run II Jet Physics, in QCD and Weak Boson Physics in Run II, (Batavia), pp. 4777 1999. hep-ex/0005012.

[39] NADA: A New Event by Event Hot Cell Killer, B. Olivier et al., DØ Note 3687 (28 Jul. 2000).

[40] See the CalNada.rcp in the p14.06.01 version of the DØ cal_nada CVS package.

[41] E. Busato and B. Andrieu, Jet Algorithms in the DØ Run II Software: Description and Users Guide. DØ Note 4457, 2004.

[42] J. R. Vlimant, et al., Technical description of the T42 Algorithm for the calorimeter noise suppression, DØ Note 4146 (2003).

[43] A. Harel, Jet ID Optimization, DØ Note 4919 (2006).

[44] JES Group, Jet Energy Scale Determination at DØ Run II (Final p17 version), DØ Note 5382 (2007).

[45] DØ Jet Energy Scale Group, DØ Certified Preliminary Jet Energy Scale.

[46] J. Agram, et al., Jet Energy Scale at DØ Run II, DØ Note 4720 (2005).

[47] Amnon Harel, Capping the JES muon corrections, DØ Note 5563 (2008).

[48] N. Makovec and J. F. Grivaz, Shifting, Smearing and Removing Simulated Jets, DØ Note 4914, 2005.

[49] Gustavo J. Otero y Garzon, Cecilia E. Gerber, Elizaveta Shabalina, Measurement of the ttbar Production Cross Section in ppbar Collisions at $\sqrt{s} = 1.96$ TeV using Secondary Vertex b-Tagging, DØ Note 5113 2006.

[50] Performance of the NN b-tagging Tool on p17 Data, DØ Note 5213.

[51] Search for Single Top Quark Production in 1 fb^{-1} of Data, DØ Note 5285.

[52] D. Chakraborty et al., DØ Note 4210 (2003).

[53] Measurement of the Z to tau tau cross section using the muon-hadronic tau final state at DO in Run 2, DØ Note 4255.

[54] S. Protopopescu and P. Svoisky, DØ Note 5094 (2006).

[55] L. Duflot and M. Ridel, The CellNN algorithm: cell level clustering in the DØ calorimeter DØ Note 3923, 2001.

[56] J. Bystricky et al., IEEE Trans. on Nucl. Sci. 51, 351 (2004).

[57] M. Abolins et al., Nucl. Instrum. and Methods A 584, 75 (2008).

[58] M. Abolins et al., IEEE Trans. on Nucl. Sci. 51, 340 (2004).

[59] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, ALPGEN, a generator for hard multiparton processes in hadronic collisions, JHEP 07 (2003) 001, hep-ph/0206293.

[60] T. Sjöstrand, L. Lonnblad, S. Mrenna, and P. Skands, Pythia 6.3: Physics and Manual, hep-ph/0308153.

[61] S. Jadach, Z. Ws, R. Decker, and J. H. Kühn. The τ decay library TAUOLA, version 2.4. Computer Physics Communications, 76:361380, August 1993.

[62] D. J. Lange. The EvtGen particle decay simulation package. Nuclear Instruments and Methods in Physics Research A, 462:152155, April 2001.

[63] Womersley J. Fisyak Y. D0gstar DØ geant simulation of the total apparatus response, DØ Note 3191.

[64] MONTE CARLO PARTICLE NUMBERING SCHEME
http://pdg.lbl.gov/2002/montecarlorpp.pdf

[65] Measurement of the ttbar cross section in the lepton+jets channel at $\sqrt{s} =$ 1.96 (topological), DØ Note 5527.

[66] Measurement of the $t\bar{t} \rightarrow all-jets$ production cross section using Secondary Vertex Tagging, DØ Note 4830.

[67] U. Bassler, J. P. Konrath, C. Schwanenberger, Measurement of the ttbar Production Cross-section at $\sqrt{s} = 1.96$ TeV in the Dimuon Final State Using p17 Data Set, DØ Note 5350 and M. Arthaud, M. Besancon, S. Chakrabarti, F. Deliot, E. Shabalina, V. Sharyy, Measurement of the ttbar Production Cross-section at $\sqrt{s} = 1.96$ TeV in Electron Muon Final States using p17 data set, DØ Note 5360.

[68] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky, W.K. Tung, hep-ph/0201195.

[69] D0gstar documentation web page,
http://www-d0.fnal.gov/computing/MonteCarlo/simulation/d0gstar.html

[70] D0SIM User Manual, D0 Note 0407, 1986.

[71] J. Hegeman, Luminosity determination and reweighting of Monte Carlo overlay luminosity for p17 hadronic top analyses, DØ Note 5561, 2008.

[72] F. Febres Cordero, L. Reina, and D. Wackeroth, NLO QCD corrections to W boson production with a massive bquark jet pair at the Fermilab Tevatron $p\bar{p}$ collider, Phys. Rev. D74 (2006) 034007, hep-ph/0606102.

[73] J. Campbell, R. Keith Ellis, F. Maltoni, and S. Willenbrock, Production of a W boson and two jets with one b-quark tag, Phys. Rev. D75 (2007) 054015, hep-ph/0606102.

[74] Y. Peters, M. Begel, E. Shabalina, and D. Wicke, Study of the W+Jets heavy flavor scale factor in p17. DØ Note 5406, 2007.

[75] Todd Adams, Yvonne Peters, Haryo Sumowidagdo, Measurement of top quark pair production in tau+lepton channel using hadronic decays of the tau in $p\bar{p}$ collisions at $\sqrt{s}$ of 1.96 TeV with $1fb^{-1}$ of Run II data, DØ Note 5735.

[76] Jeroen G. Hegeman. MEASUREMENT OF THE TOP QUARK PAIR PRODUCTION CROSS SECTION IN PROTON-ANTIPROTON COLLISIONS AT $\sqrt{s}$ = 1.96 TeV HADRONIC TOP DECAYS WITH THE DØDETECTOR, FERMILAB-THESIS-2008

[77] $\not\!\!E_T$ Significanfce Algorithm in RunII data, DØ Note 4254.

[78] http://www-d0.hef.kun.nl//fullAgenda.php?ida=a07576&fid=35

[79] http://www-d0.fnal.gov/ leiwang/zpt_analysis/zpt_note_v4.ps

[80] http://www-d0.fnal.gov/d0dist/dist/releases/development/muonid_note/p17_certif/muonid_p17.pdf

[81] Determination of Weighting Factors for ALPGEN Monte Carlo Signal and Background Samples, DØ Note 5016.

[82] S. Höoche et al., "Matching Parton Showers and Matrix Elements," hep-ph/0602031.

[83] Measurement of $\sigma(p\bar{p} \rightarrow t\bar{t})$ in $\tau + jets$ channel using 1 fb^{-1} of data, DØ Note 5765.

[84] http://seal.web.cern.ch/seal/documents/minuit/mntutorial.pdf

[85] http://schwind.home.cern.ch/schwind/MLPfit.html

[86] https://plone4.fnal.gov/P1/D0Wiki/physics/top/HowToVaryJetCorrections

[87] Systematic uncertainty of tau identification neural network for p17 data, DØ Note 5408.

[88] trigeff cafe: Single object efficiency calculation in the CAFe environment, DØ Note 5093.

[89] The Top Trigger Package: Including Version 13 Triggers and Other Improvements, DØ Note 4883.

[90] Direct Search for Charged Higgs Bosons in $t\bar{t}$ events, DØ Note 4357.

[91] Measurement of top pair production cross section from the tau+jets final state, DØ Note 5158.

[92] http://www-d0.fnal.gov/Run2Physics/cs/skimming/fixPass2p170903.html

Appendices

Appendix A

Trigger Description

The full trigger description for the ALLJETS triggers and different trigger level requirements are given below :

V09.00-V10.00: 4JT10

L1 : 4 L1 calorimeter trigger towers with $E_T > 5$ GeV

L2 : 3 L2 jets with $E_T > 8$ GeV

L3 : 4 L3 jets with $E_T > 10$ GeV and $|\eta| < 3$

where L1, L2, and L3 represents level 1, level 2 and level 3 trigger respectively.

V09.00 - V10.00: 4JT10

L1 : 4 L1 calorimeter trigger towers with $E_T > 5$ GeV

L2 : 3 L2 jets with $E_T > 8$ GeV and $H_T > 90$ GeV (jets with $E_T > 5$ GeV)

L3 : 4 L3 jets with $E_T > 10$ GeV and $|\eta| < 3$, 2 L3 jets with 20 GeV and $|\eta| < 3$

where H_T is the sum of all the jet transverse energy in an event.

V10.00 - V11.00: 4JT10

L1 : 4 L1 calorimeter trigger towers with $E_T > 5$ GeV

L2 : 3 L2 jets with $E_T > 8$ GeV and $H_T > 90$ GeV (jets with $E_T > 5$ GeV)

L3 : 4 L3 jets with $E_T > 10$ GeV and $|\eta| < 3$, 2 L3 jets with 20 GeV and $|\eta| < 3$

V12.00 - V13.03: 4JT12

L1 : 3 L1 calorimeter trigger towers with $E_T > 5$ GeV

L2 : 3 L3 jets with $E_T > 8$ GeV and $H_T > 50$ GeV (jets with $E_T > 5$ GeV)

L3 : 4 L3 jets with $E_T > 12$ GeV and $|\eta| < 3.6$, 3 L3 jets with $E_T > 15$ GeV and $|\eta| < 3.6$, 2 L3 jets with $E_T > 25$ GeV and $|\eta| < 3.6|$

V13.03 - V13.30: JT2_4JT12L_HT

L1 : 3 L1 calorimeter trigger towers with $E_T > 4$ GeV and $|\eta| < 2.4$, 2 L1 calorimeter trigger towers with $E_T > 5$ GeV (v13.03)

L1 : 3 L1 calorimeter trigger towers with $E_T > 5$ GeV (v13.10-v13.23)

L2 : 3 L2 jets with $E_T > 6$ GeV and $H_T > 70$ GeV (jets with $E_T > 6$ GeV)

L3 : 4 L3 jets with $E_T > 12$ GeV and $|\eta| < 3.6$, 3 L3 jets with $E_T > 15$ GeV and $|\eta| < 3.6$, 2 L3 jets with $E_T > 25$ GeV and $|\eta| < 3.6$ and $H_T > 120$ GeV (L3 jets with $E_T > 9$ GeV and $|\eta| < 3.6$

V13.30 - V14.00: JT2_4JT12L_HT

L1 : 3 L1 calorimeter trigger towers with $E_T > 4$ GeV and $|\eta| < 2.4$, 2 L1 calorimeter trigger towers with $E_T > 5$ GeV

L1 : 3 L1 calorimeter trigger towers with $E_T > 5$ GeV

L2 : 3 L2 jets with $E_T > 6$ GeV and $H_T > 70$ GeV (jets with $E_T > 6$ GeV)

L3 : 4 L3 jets with $E_T > 12$ GeV and $|\eta| < 3.6$, 3 L3 jets with $E_T > 15$ GeV and $|\eta| < 3.6$, 2 L3 jets with $E_T > 25$ GeV and $|\eta| < 3.6$ and $H_T > 120$ GeV (L3 jets with $E_T > 9$ GeV and $|\eta| < 3.6$

V14.00 - V14.80: JT2_4JT20

L1 : 3 L1 calorimeter trigger towers with $E_T > 5$ GeV, 3 L1 calorimeter trigger towers with $E_T > 4$ GeV and $|\eta| < 2.6$

L2 : 3 L2 jets with $E_T > 6$ GeV and $H_T > 75$ GeV (jets with $E_T > 6$ GeV and $|\eta| < 2.6$)

L3 : 4 L3 jets with $E_T > 10$ GeV and $|\eta| < 3.6$, 2 L3 jets with $E_T > 20$ GeV and

$|\eta| < 3.6$

V14.80 - V15.00: JT2_4JT20

L1 : 3 L1 calorimeter trigger towers with $E_T > 5$ GeV, 3 L1 calorimeter trigger towers with $E_T > 4$ GeV and $|\eta| < 2.6$

L2 : 3 L2 jets with $E_T > 6$ GeV and $H_T > 75$ GeV (jets with $E_T > 6$ GeV and $|\eta| < 2.6$)

L3 : 4 L3 jets with $E_T > 10$ GeV and $|\eta| < 3.6$, 2 L3 jets with $E_T > 20$ GeV and $|\eta| < 3.6$

Appendix B

Events Selection Tables

Electroweak Samples		***Events Passed***
$\boldsymbol{Wbb + jets \rightarrow l\nu + bb + jets}$		7.36 ± 0.26
$\boldsymbol{Wcc + jets \rightarrow l\nu + cc + jets}$		3.84 ± 0.14
$\boldsymbol{Wjj + jets \rightarrow l\nu + jj + jets}$		2.66 ± 0.08
$\boldsymbol{Zbb + jets \rightarrow \tau\tau + bb + jets}$		1.08 ± 0.36
$\boldsymbol{Zcc + jets \rightarrow \tau\tau + cc + jets}$		0.55 ± 0.06
$\boldsymbol{Zjj + jets \rightarrow \tau\tau + jj + jets}$		0.86 ± 0.02
$\boldsymbol{Zbb + jets \rightarrow ee + bb + jets}$		0.15 ± 0.02
$\boldsymbol{Zcc + jets \rightarrow ee + cc + jets}$		0.07 ± 0.02
$\boldsymbol{Zjj + jets \rightarrow ee + jj + jets}$		0.11 ± 0.01
$\boldsymbol{Zbb + jets \rightarrow \mu\mu + bb + jets}$		0.02 ± 0.01
$\boldsymbol{Zcc + jets \rightarrow \mu\mu + cc + jets}$		0.01 ± 0.01
$\boldsymbol{Zjj + jets \rightarrow \mu\mu + jj + jets}$		0.01 ± 0.00

Table B.1: List of event acceptances for electroweak samples for tau Types I & II. The number of events are scaled to $974.2 pb^{-1}$.

Preselection Cuts	$t\bar{t} \rightarrow \tau_{hadronic} + jets$ ***Number of Events***
Initial Number of Events:	375408
Particle Selection :	81128
MC Quality Cuts:	78592
JESMU Jet Energy Scale Corrections: **Jet Selection Cuts:** At least four jets with $\|\eta\| < 2.5$, three leading (in P_T) jets with P_Ts , 35, 25, and 25 GeV, and all other jets with $P_T > 13$ GeV	61926
Veto Electrons and Muons:	61877
Primary Vertex Selection Cuts: $\|z\| < 60$ cm, at least 3 tracks	60835
$\not\!\!E_T$ **Selection Cuts:** The $\not\!\!E_T$cut range is 15 GeV $\leq \not\!\!E_T \leq$ 500 GeV	57744
$\not\!\!E_T$ **Significance Selection Cut:** $\not\!\!E_T$ Singnificance > 3	49036
Loose Tau Selection: At least one tau with $NN_\tau > 0.3$ and $P_T > 10$ GeV	34189
Final Selections Cuts :	
Final Jet Selection All jets with $P_T > 20$ GeV, $\|\eta\| < 2.5$, **Tight Tau Selection & Tau Type:** At least one Type I or II tau with $NN_\tau > 0.95$, $P_T > 10$ GeV and $\|\eta\| < 2.5$	8102
Corrections/Weights:	
b**-jet tagging Efficiency:** At least one $NNbtag$ b-jet 0.601	
Trigger Efficiency: 0.748	
Primary Vertex Reweight Efficiency: 0.971	
Luminosity Rewight Efficiency: 0.989	
Scaled Events: **k-factor & Alpgen Sacling Factor:** $= 1.2 \times 0.0065$	27.20±0.382

Table B.2: Summary of $t\bar{t} \rightarrow \tau_{hadronic} + jets$ event selections for tau Types I & II. The final number of expected events are scaled to the data with a total integrated luminosity of $974.2 pb^{-1}$.

Preselection Cuts	***$t\bar{t} \rightarrow e + jets$ Number of Events***
Initial Number of Events:	375408
Particle Selection :	147595
MC Quality Cuts:	142883
JESMU Jet Energy Scale Corrections:	
Jet Selection Cuts:	
At least four jets with $\|\eta\| < 2.5$, three leading	
(in P_T) jets with P_Ts , 35, 25, and 25 GeV, and	
all other jets with $P_T > 13$ GeV	92317
Veto Electrons and Muons:	49054
Primary Vertex Selection Cuts:	
$\|z\| < 60$ cm, at least 3 tracks	47909
$\not\!\!E_T$ **Selection Cuts:**	
The $\not\!\!E_T$cut range is 15 GeV $\leq \not\!\!E_T \leq$ 500 GeV	45942
$\not\!\!E_T$ **Significance Selection Cut:**	
$\not\!\!E_T$ Singnificance > 3	39677
Loose Tau Selection:	
At least one tau with $NN_\tau > 0.3$ and $P_T > 10$ GeV	25759
Final Selections Cuts :	
Final Jet Selection	
All jets with $P_T > 20$ GeV, $\|\eta\| < 2.5$,	
Tight Tau Selection & Tau Type:	
At least one Type I or II tau with	
$NN_\tau > 0.95$, $P_T > 10$ GeV and $\|\eta\| < 2.5$	8533
Corrections/Weights:	
b**-jet tagging Efficiency:**	
At least one $NNbtag$ b-jet 0.587	
Trigger Efficiency: 0.698	
Primary Vertex Reweight Efficiency: 0.981	
Luminosity Rewight Efficiency: 0.989	
Scaled Events:	
k-factor & Alpgen Sacling Factor: $= 1.2 \times 0.0065$	25.87±0.36

Table B.3: Summary of $t\bar{t} \rightarrow e + jets$ event selections for tau Types I & II. The final number of expected events are scaled to the data with a total integrated luminosity of $974.2 pb^{-1}$.

Preselection Cuts	***$t\bar{t} \rightarrow \mu + jets$ Number of Events***
Initial Number of Events:	375408
Particle Selection :	146685
MC Quality Cuts:	142517
JESMU Jet Energy Scale Corrections:	
Jet Selection Cuts:	
At least four jets with $\|\eta\| < 2.5$, three leading	
(in P_T) jets with P_Ts , 35, 25, and 25 GeV, and	
all other jets with $P_T > 13$ GeV	94739
Veto Electrons and Muons:	51517
Primary Vertex Selection Cuts:	
$\|z\| < 60$ cm, at least 3 tracks	50589
$\not\!\!E_T$ **Selection Cuts:**	
The $\not\!\!E_T$cut range is 15 GeV $\leq \not\!\!E_T \leq$ 500 GeV	48853
$\not\!\!E_T$ **Significance Selection Cut:**	
$\not\!\!E_T$ Singnificance > 3	43643
Loose Tau Selection:	
At least one tau with $NN_\tau > 0.3$ and $P_T > 10$ GeV	16936
Final Selections Cuts :	
Final Jet Selection	
All jets with $P_T > 20$ GeV, $\|\eta\| < 2.5$,	
Tight Tau Selection & Tau Type:	
At least one Type I or II tau with	
$NN_\tau > 0.95$, $P_T > 10$ GeV and $\|\eta\| < 2.5$	643
Corrections/Weights:	
b**-jet tagging Efficiency:**	
At least one $NNbtag$ b-jet 0.603	
Trigger Efficiency: 0.659	
Primary Vertex Reweight Efficiency: 0.976	
Luminosity Rewight Efficiency: 0.987	
Scaled Events:	
k-factor & Alpgen Sacling Factor: $= 1.2 \times 0.0065$	1.92±0.10

Table B.4: Summary of $t\bar{t} \rightarrow \mu + jets$ event selections for tau Types I & II. The final number of expected events are scaled to the data with a total integrated luminosity of $974.2 pb^{-1}$.

Preselection Cuts	***$t\bar{t} \rightarrow$ dilepton + jets Number of Events***
Initial Number of Events:	532750
MC Quality Cuts:	482787
JESMU Jet Energy Scale Corrections: **Jet Selection Cuts:** At least four jets with $\|\eta\| < 2.5$, three leading (in P_T) jets with P_Ts , 35, 25, and 25 GeV, and all other jets with $P_T > 13$ GeV	62209
Veto Electrons and Muons:	33318
Primary Vertex Selection Cuts: $\|z\| < 60$ cm, at least 3 tracks	32584
$\not\!\!E_T$ **Selection Cuts:** The $\not\!\!E_T$cut range is 15 GeV $\leq \not\!\!E_T \leq$ 500 GeV	31436
$\not\!\!E_T$ **Significance Selection Cut:** $\not\!\!E_T$ Singnificance > 3	28585
Loose Tau Selection: At least one tau with $NN_\tau > 0.3$ and $P_T > 10$ GeV	22170
Final Selections Cuts :	
Final Jet Selection All jets with $P_T > 20$ GeV, $\|\eta\| < 2.5$, **Tight Tau Selection & Tau Type:** At least one Type I or II tau with $NN_\tau > 0.95$, $P_T > 10$ GeV and $\|\eta\| < 2.5$	9682
Corrections/Weights:	
***b*-jet tagging Efficiency:** At least one *NNbtag b*-jet 0.601	
Trigger Efficiency: 0.505	
Primary Vertex Reweight Efficiency: 0.976	
Luminosity Rewight Efficiency: 0.994	
Scaled Events: **k-factor & Alpgen Sacling Factor:** $= 1.2 \times 0.001586$	5.44±0.07

Table B.5: Summary of $t\bar{t} \rightarrow dilepton + jets$ event selections for tau Types I & II. The final number of expected events are scaled to the data with a total integrated luminosity of $974.2 pb^{-1}$.

Selection Cuts	***Number of Data Events***
Initial Number of Events:	115832858
Trigger Selection:	31063568
Data Quality Cuts:	22036392
JESMU Jet Energy Scale Corrections: **Jet Selection Cuts:** At least four jets with $\|\eta\| < 2.5$, three leading (in P_T) jets with P_Ts , 35, 25, and 25 GeV, and all other jets with $P_T > 13$ GeV	15653441
Veto Electrons and Muons:	15650306
Primary Vertex Selection Cuts: $\|z\| < 60$ cm, at least 3 tracks	15006194
$\not\!\!E_T$ **Selection Cuts:** The $\not\!\!E_T$cut range is 15 GeV $\geq \not\!\!E_T \leq$ 500 GeV	6579088
$\not\!\!E_T$ **Significance Selection Cut:** $\not\!\!E_T$ Significance > 3	498476
Loose Tau Selection: At least one tau with $NN_\tau > 0.3$ and $P_T > 10$ GeV	191341
Final Selections Cuts :	
Final Jet Selection All jets with $P_T > 20$ GeV, $\|\eta\| < 2.5$, b**-jet Identification:** At least one $NNbtag$ b-jet **Tight Tau Selection Cuts:** $0.3 < NN_\tau < 0.9$ with $P_T > 10$ GeV and $\|\eta\| < 2.5$ **Tau Type Selection Cut:** Select tau Type I and Type II together	1541

Table B.6: Data event selection cut flows for tau Types I & II .

	(# passed)/(total #)	# passed scaled
$t\bar{t} \to \tau + jets$	0.095 ± 0.001	
$t\bar{t} \to e + jets$	0.100 ± 0.002	
$t\bar{t} \to \mu + jets$	0.056 ± 0.004	
$t\bar{t} \to l + l$	0.071 ± 0.001	
$Wbb + jets \to l\nu + bb + jets$		5.13 ± 0.19
$Wcc + jets \to l\nu + cc + jets$		3.06 ± 0.15
$Wjj + jets \to l\nu + jj + jets$		2.33 ± 0.07
$Zbb + jets \to \tau\tau + bb + jets$		1.08 ± 0.04
$Zcc + jets \to \tau\tau + cc + jets$		0.42 ± 0.05
$Zjj + jets \to \tau\tau + jj + jets$		0.60 ± 0.02
$Zbb + jets \to ee + bb + jets$		0.02 ± 0.01
$Zcc + jets \to ee + cc + jets$		0.02 ± 0.01
$Zjj + jets \to ee + jj + jets$		0.03 ± 0.00
$Zbb + jets \to \mu\mu + bb + jets$		0.03 ± 0.01
$Zcc + jets \to \mu\mu + cc + jets$		0.01 ± 0.01
$Zjj + jets \to \mu\mu + jj + jets$		0.03 ± 0.00

Table B.7: Estimate of $t\bar{t}$ contamination in data for tau Types I & II. The number of events is scaled to 974.2 pb^{-1}

Appendix C

Plots

C.1 NN Discriminant Variables

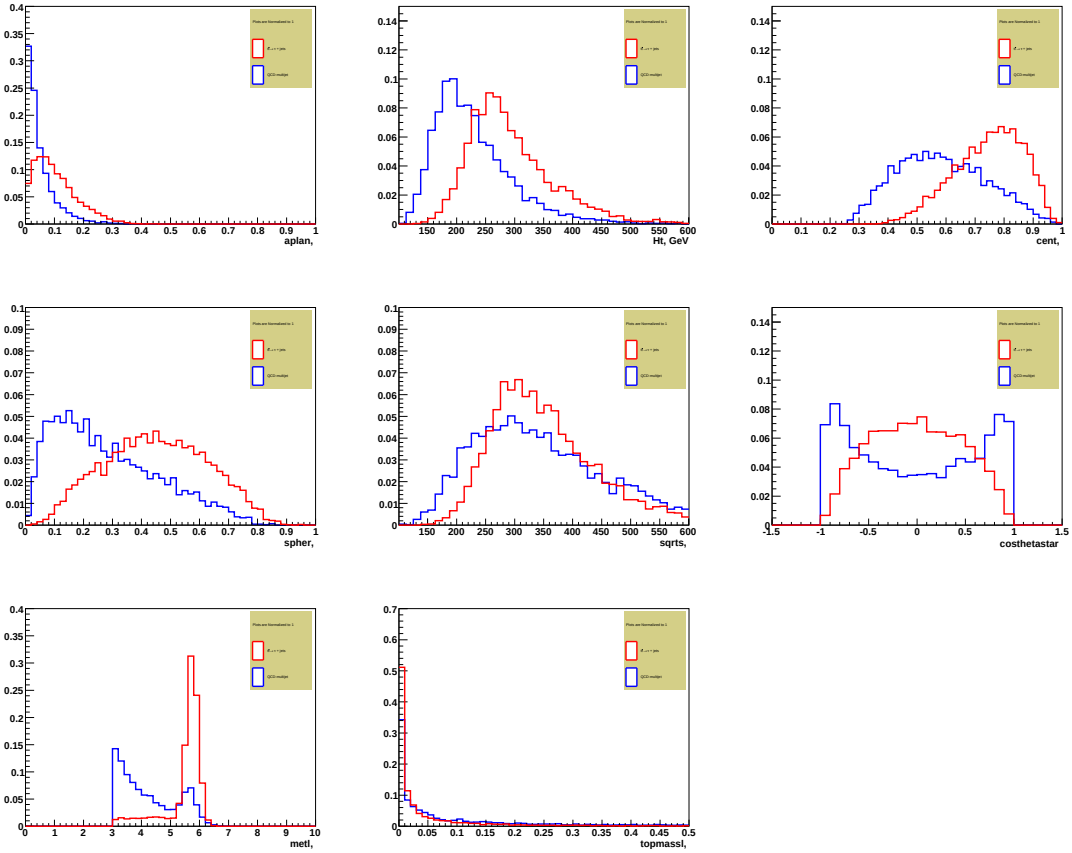

Figure C.1: The normalized topological variables for tau Type I & II. The red plot is for $t\bar{t} \rightarrow \tau + jets$ and the blue plot is for multijet.

C.2 NN Input Variables for the b-veto control sample

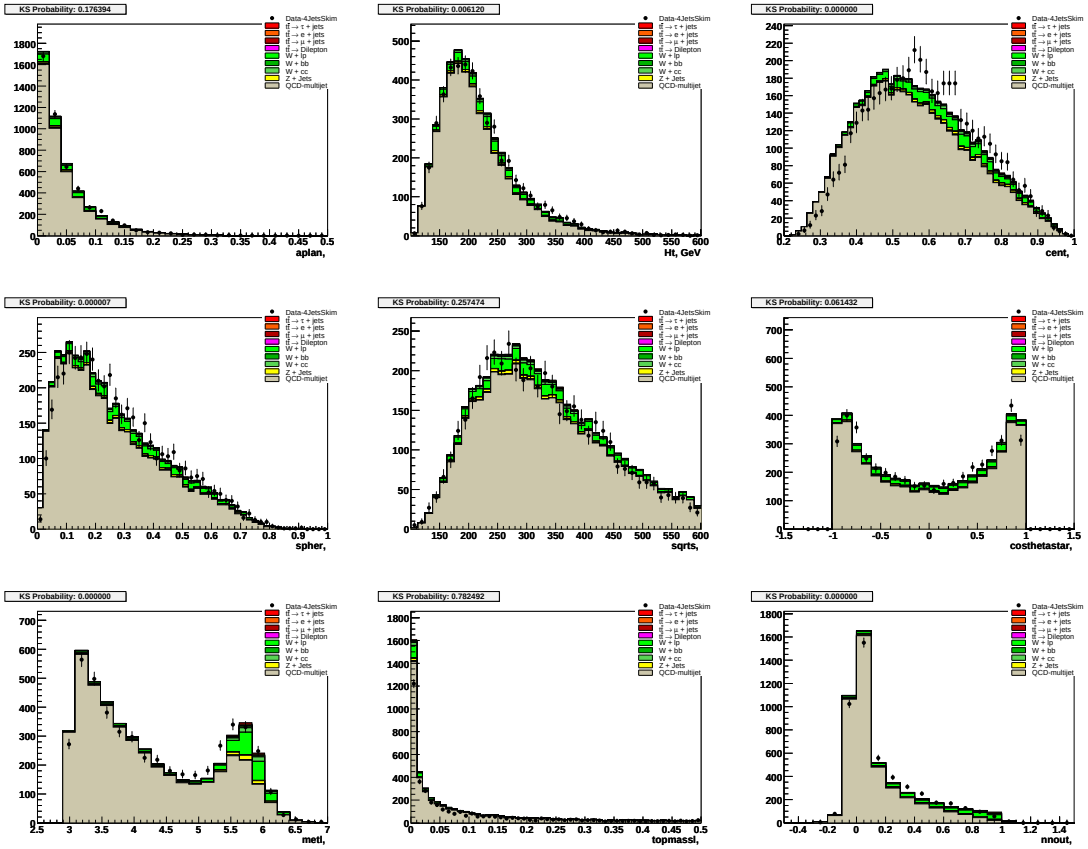

Figure C.2: The topological variables for the b-veto control sample (tau Types I & II). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is. The last plot shows the resulting NN_{Topo} output. The error bars on the plots represent the statistical uncertainties only.

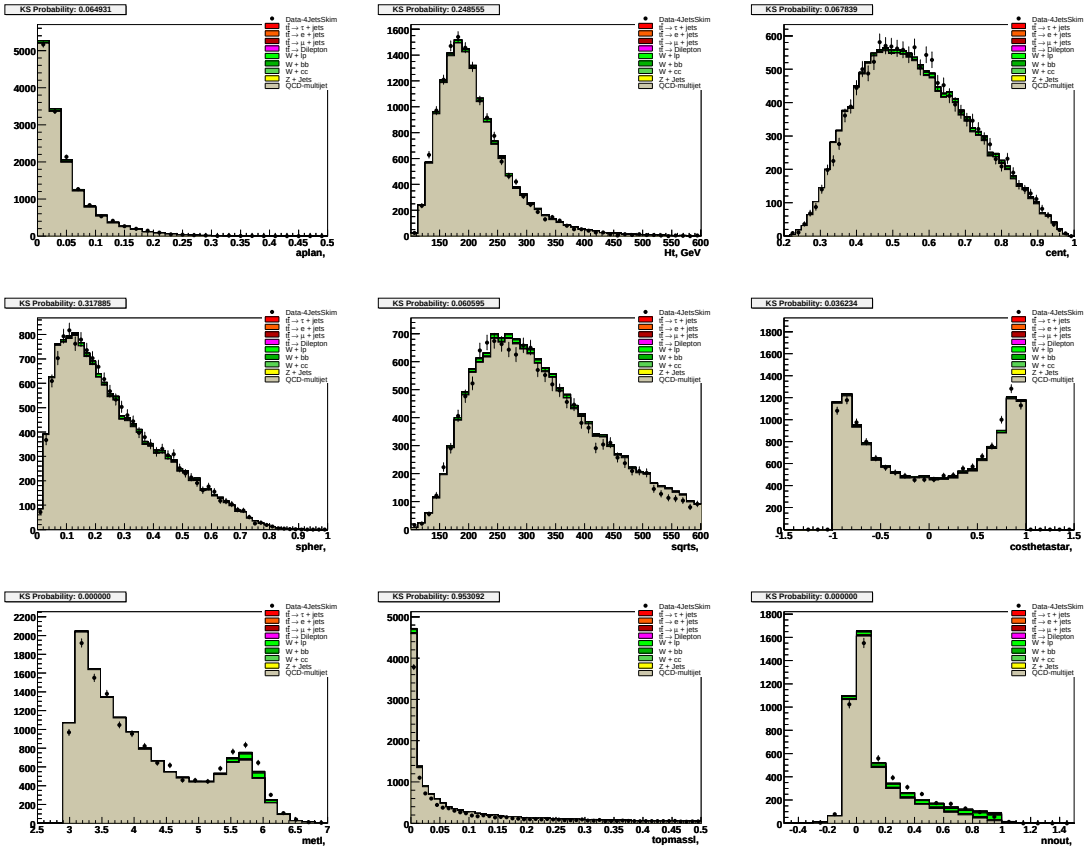

Figure C.3: The topological variables in the b-veto control sample (tau Type III). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is. The last plot shows the resulting NN_{Topo} output. The error bars on the plots represent the statistical uncertainties only.

C.3 NN input variables in the signal data sample

The control plots for the signal data set on which the $t\bar{t} \rightarrow \tau + jets$ cross section measurement is performed.

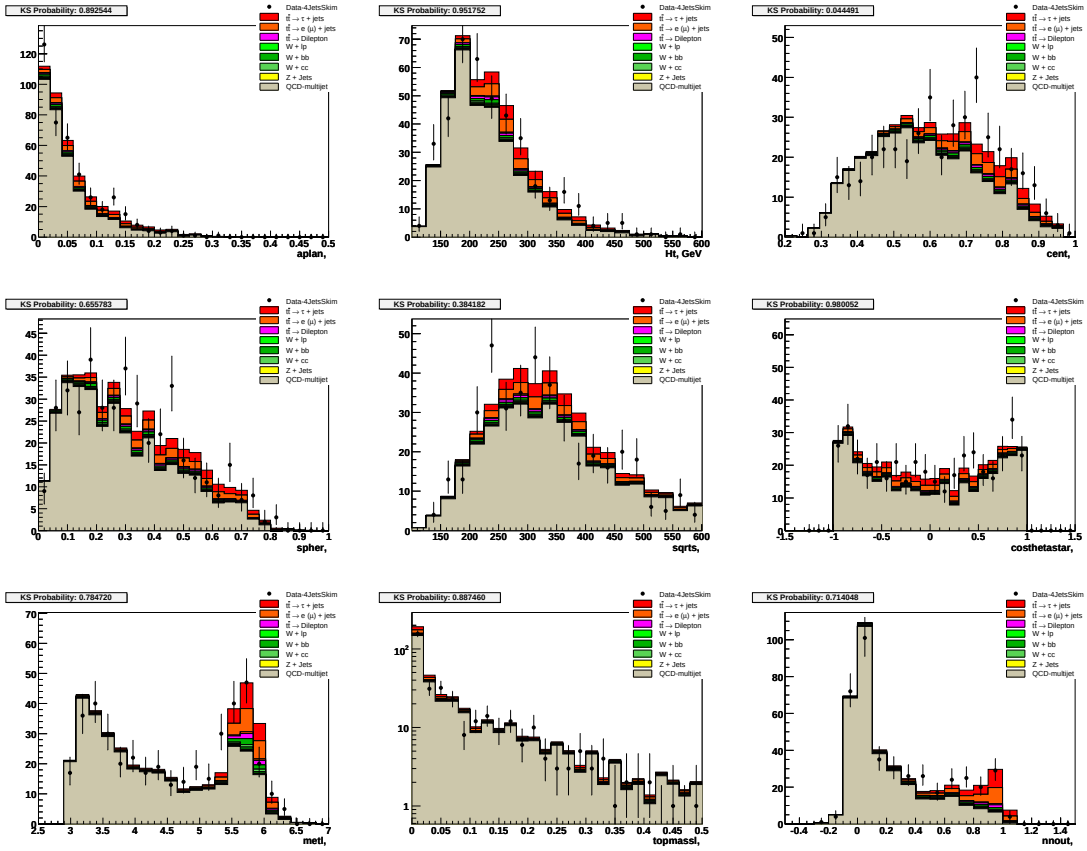

Figure C.4: The topological variables in the signal sample (tau Types I & II). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is.

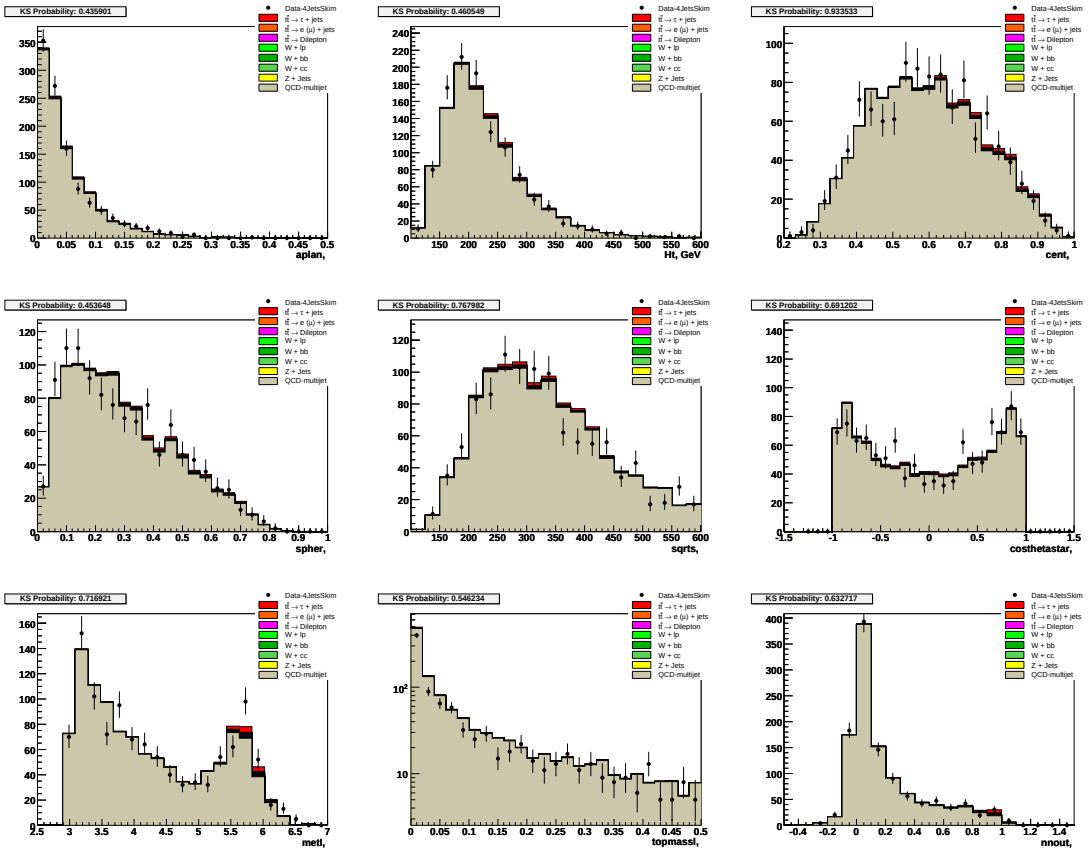

Figure C.5: The topological variables in the signal sample (tau Type III). The Kolmogorov-Smirnov (KS) probabilities are shown, indicating how good the agreement is.

C.4 NN input variables in the signal data sample after $NN_{Topo} > 0.6$ cut

Plots for NN input variables after $NN_{Topo} > 0.6$ cut. This allows to check data to Monte Carlo agreement in signal-dominated region (assuming $\sigma(t\bar{t} \to \tau + jets) = 6.9$ pb).

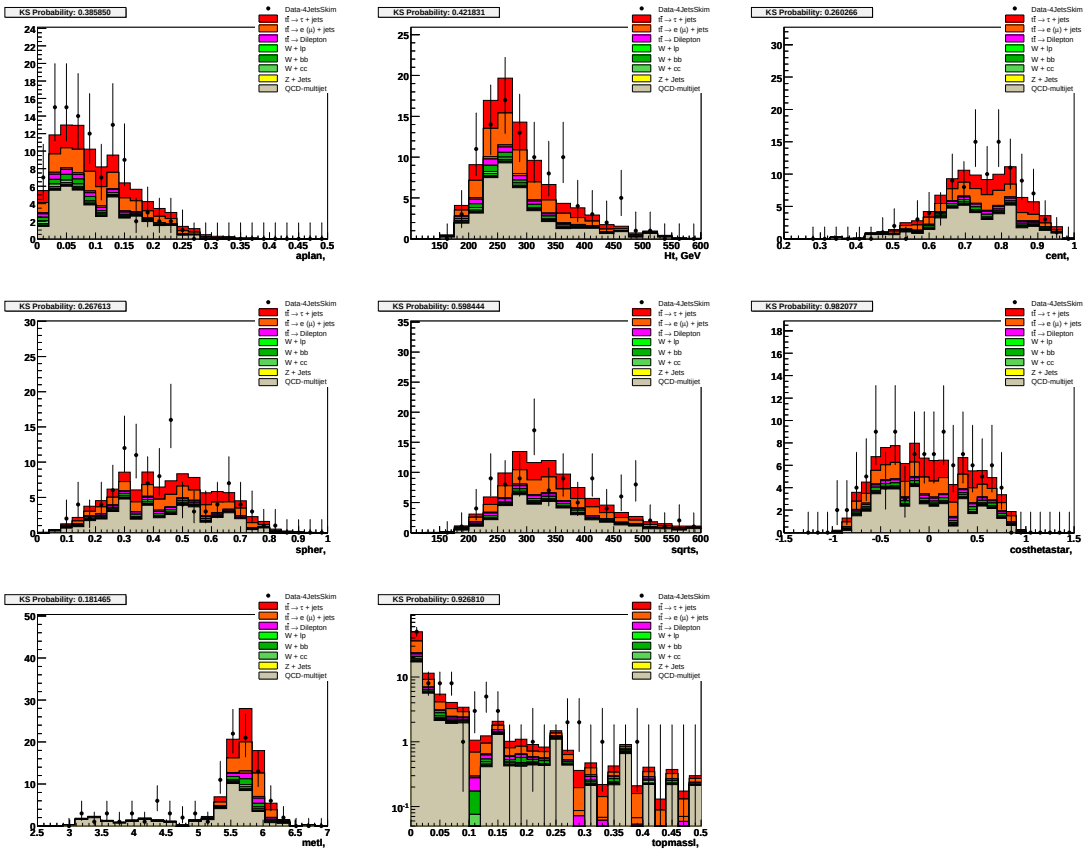

Figure C.6: The topological variables in the signal sample after $NN_{Topo} > 0.6$ cut (tau Types I & II). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is.

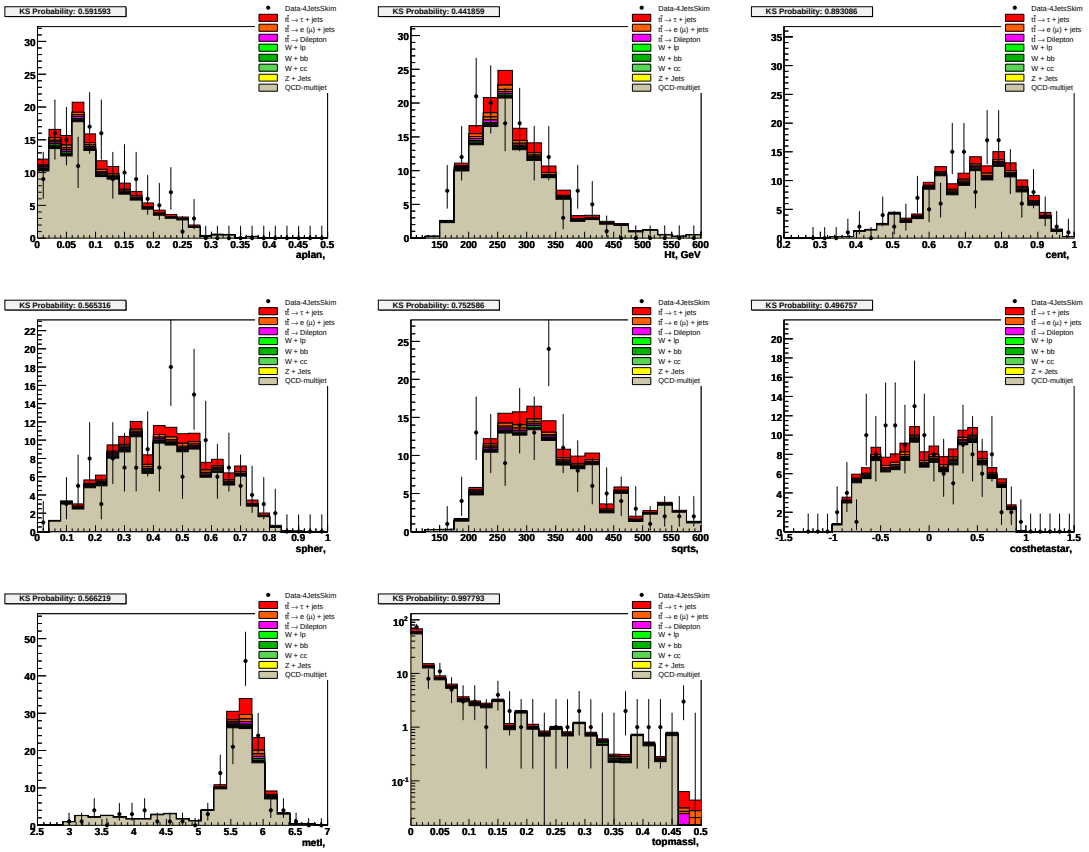

Figure C.7: The topological variables in the signal sample after $NN_{Topo} > 0.6$ cut (tau Type III). The Kolmogorov-Smirnov (KS) probabilities are shown, indicating how good the agreement is.

C.5 Jets in the signal data sample after $NN_{Topo} > 0.6$ cut

Plots for jets properties after $NN_{Topo} > 0.6$ cut. This allows to check data to Monte Carlo agreement in signal-dominated region (assuming $\sigma(t\bar{t} \to \tau + jets) = 6.9$ pb).

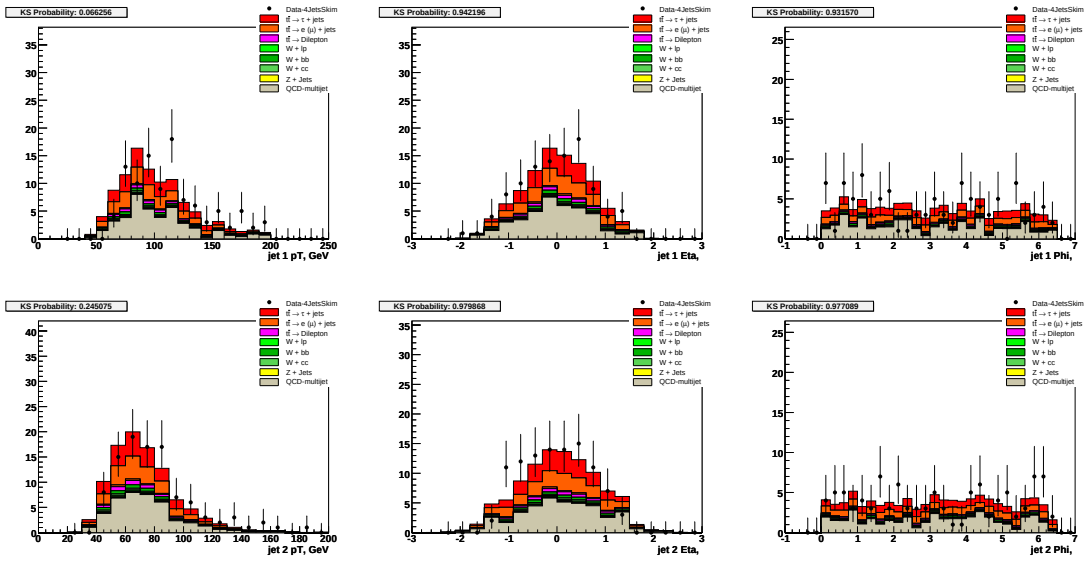

Figure C.8: Jet properties in the signal sample after $NN_{Topo} > 0.6$ cut (tau Types I & II). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is.

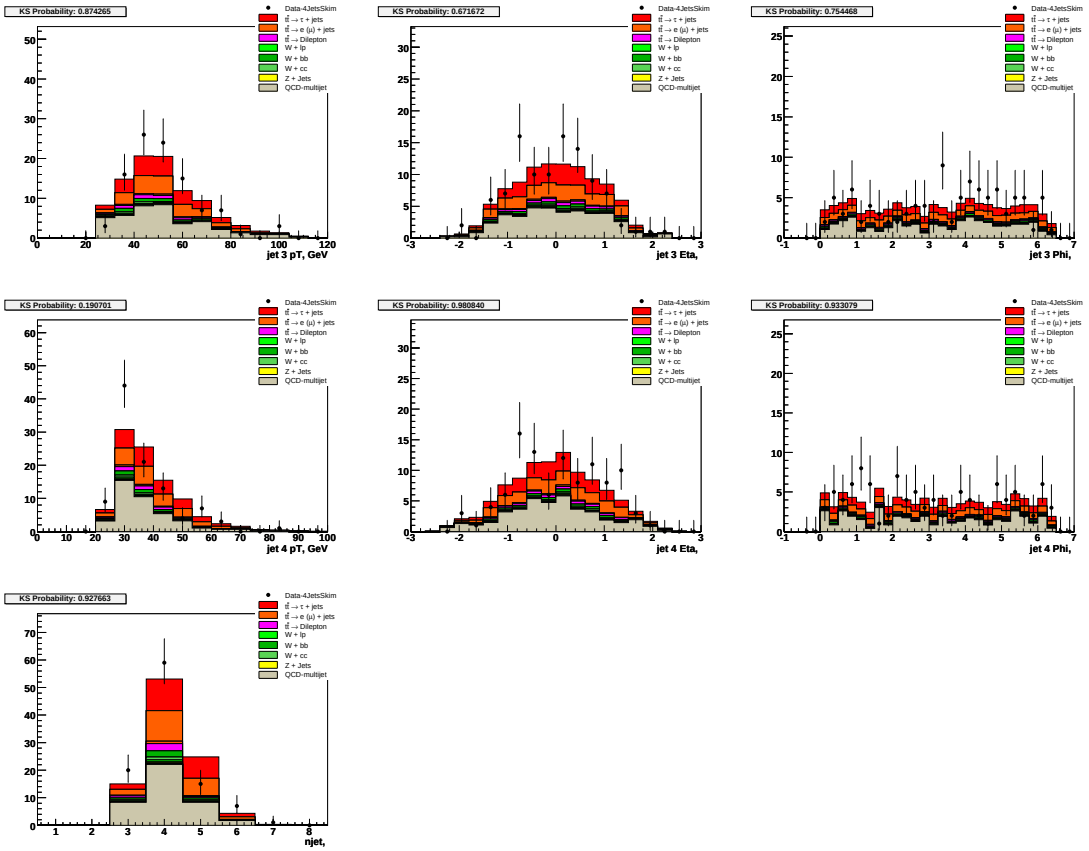

Figure C.9: Jets properties in the signal sample after $NN_{Topo} > 0.6$ cut (tau Types I & II). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is.

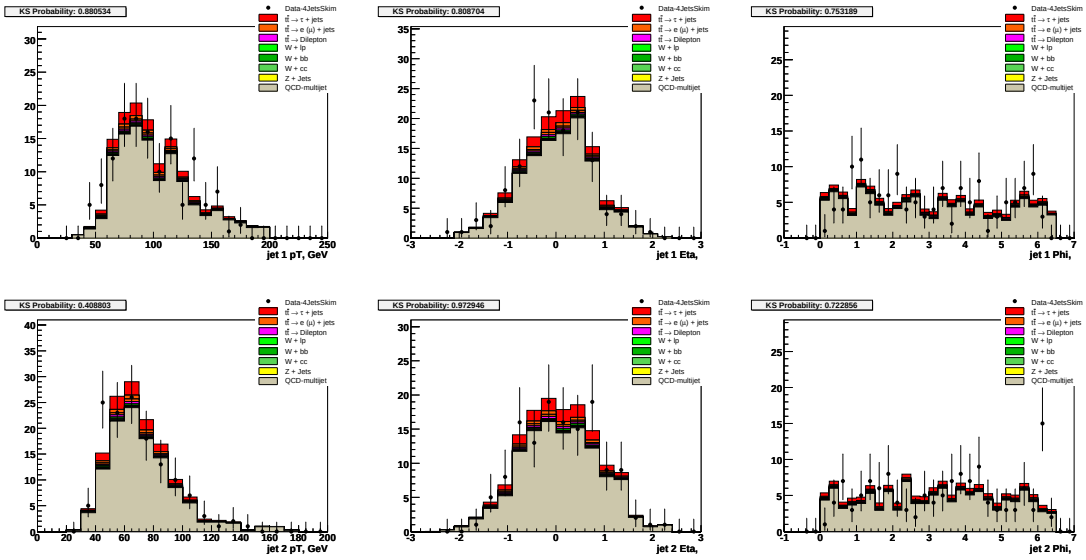

Figure C.10: Jets properties in the signal sample after $NN_{Topo} > 0.6$ cut (tau Type III). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is.

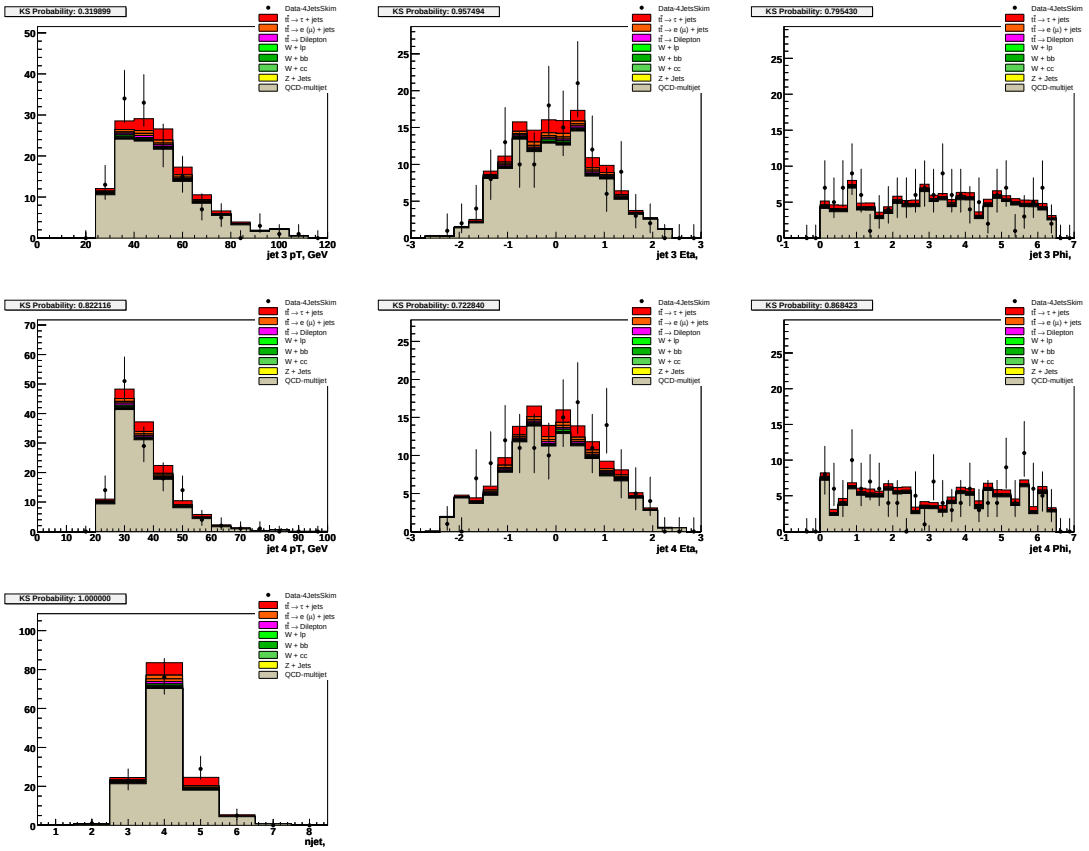

Figure C.11: Jets properties in the signal sample after $NN_{Topo} > 0.6$ cut (tau Type III). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is.

C.6 Tau related control plots in the signal data sample after $NN_{Topo} > 0.6$ cut

The control plots for tau lepton properties and correlated variables in the signal data sample after $NN_{Topo} > 0.6$ cut. As the analysis method described in this thesis is based on fitting the top topological NN_{Topo} distributions, a good agreement between samples for tau attributes is not expected. Moreover, the multijet background shape is taken from events where absence of tight taus candidates are required. Nevertheless, it can be seen from the plots that the tau properties are modeled reasonably well.

It is possible that the JESMU correction on jets are very different from the jets that are like taus. This may lead to $\not\!\!E_T$ being over corrected. In order to check for this over correction, $\not\!\!E_T$ is plotted along the direction ($\not\!\!E_T$_*colin*) and along the perpendicular ($\not\!\!E_T$_*perp*) of tau candidate respectively. These plots are particularly interesting since they demonstrate how $\not\!\!E_T$ and tau energy scale interplay. From the plots $\not\!\!E_T$_*perp* and $\not\!\!E_T$_*colin* it is clear that the data is well modeled and there is no $\not\!\!E_T$ over correction.

There are two tau related variables are plotted *mttaumet* and *dphitaumet*. The variables *mttaumet* is defined as the transverse mass of τ and $\not\!\!E_T$ and the variable *dphitaumet* is the angle between τ and $\not\!\!E_T$.

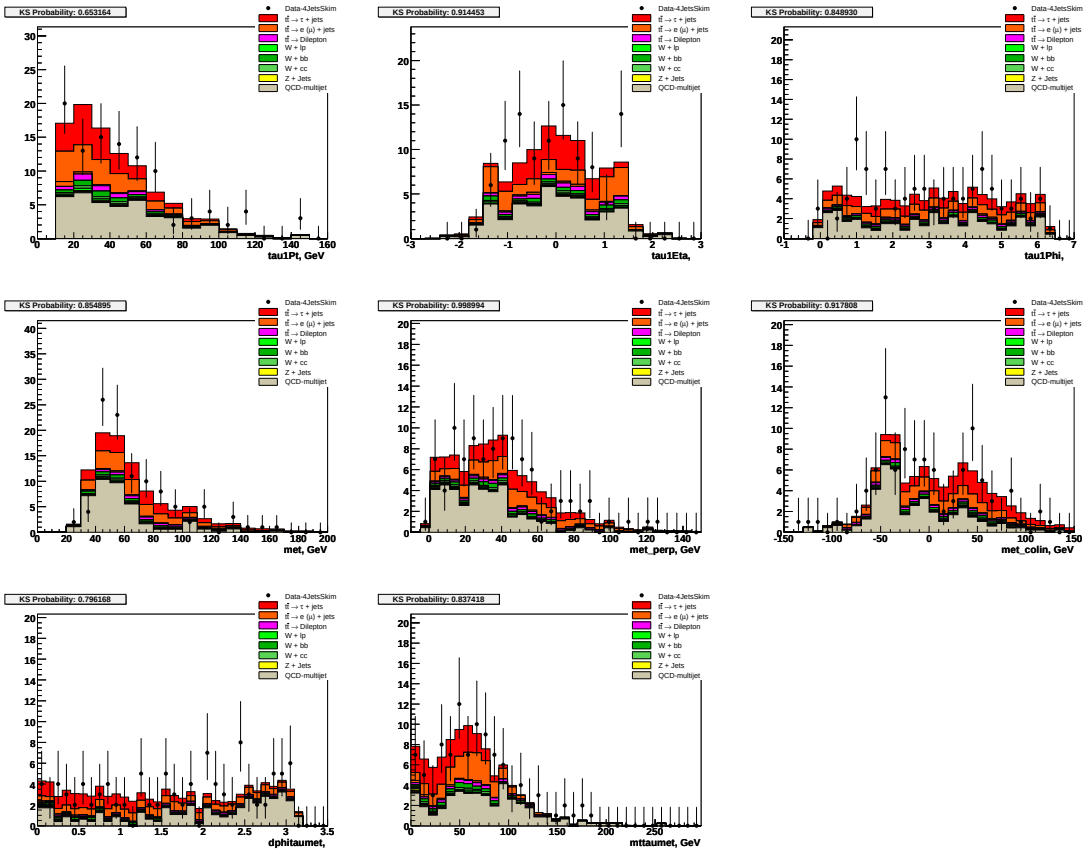

Figure C.12: Tau related variables in the signal sample (tau Types I & II). The Kolmogorov-Smirnov (KS) probabilities are shown to indicate how good the agreement is.

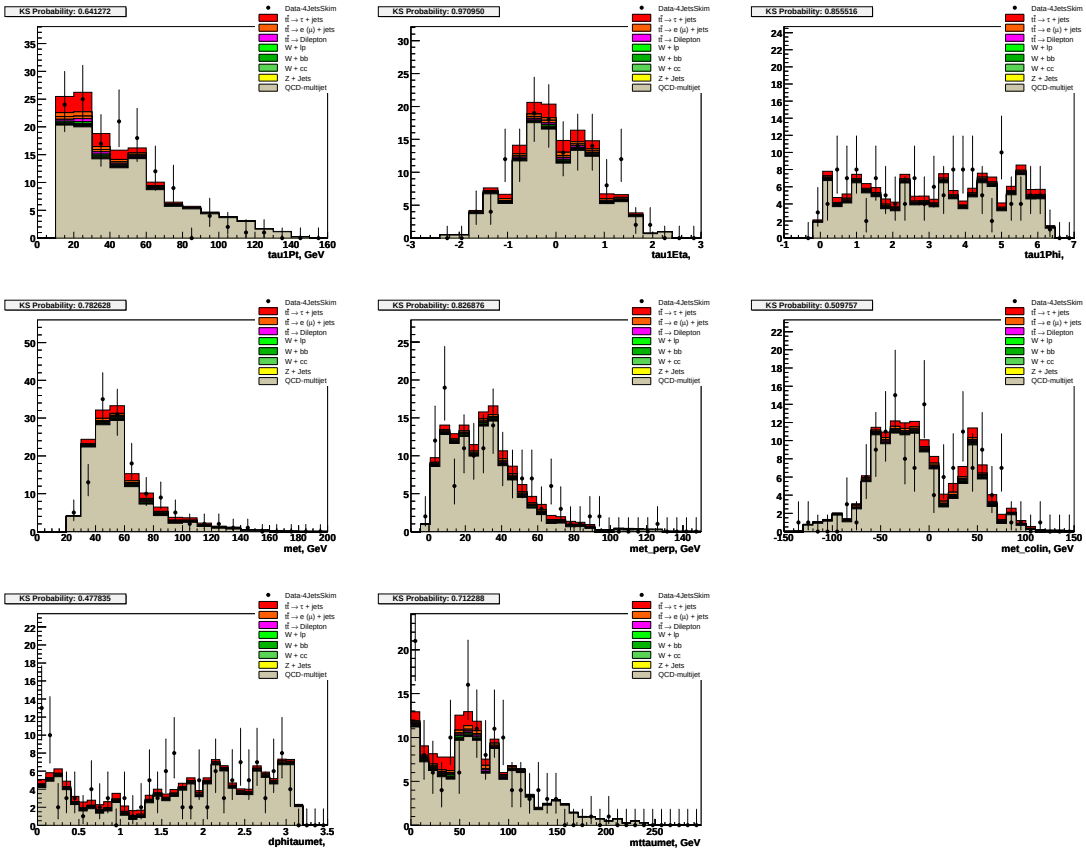

Figure C.13: Tau related variables in the signal sample (tau Type III). The Kolmogorov-Smirnov (KS) probabilities are shown, indicating how good the agreement is.

C.7 Jet trigger turn-ons for Tau candidates

Tau leptons in the hadronic decay modes have very similar signature as hadronic jets in the detector. Therefore, nearly all the tau leptons with $P_T > 15$ GeV are identified as jet candidates. The question arises whether such "tau-jets" have the same trigger turn-ons as standard jets which is used for this analysis.

In order to investigate this a di-jet multijet data sample (event requirements are at least two jets and $\Delta\phi > 2.8$ between them). Then the efficiency of firing a desired jet trigger term as described in [88] is measured which gives standard turn-on for the jets. After that the same procedure is repeated using only the jets matched within $\Delta R < 0.5$ of a tight ($NN_\tau > 0.9$) tau candidate. A slightly looser NN_τ cut was chosen than what is used in this analysis (0.95) in order to make the results comparable with the other tau analyses.

The trigger terms used are the following:

L1: CSWJT(1,8,3.2)_ncu
L2: One JET(0,5) with 8 or 15 GeV
L3: One SC5JET_9_PV3 with 20 GeV

The figures C.14, C.15, C.16 and C.17 represents the turn-on curves. On each plot two turn-on curves are superimposed. One is for standard jet turn-on (red dots) and the other one is the turn-on for a jet matched to a tight ($NN_\tau > 0.9$) tau candidate (black dots with error bars).

From the above figures, it can be seen that the tight tau leptons matched with jets with energies less than 150 GeV have trigger efficiency identical to that of jets. Therefore the trigger simulation procedure used in this analysis is valid and small difference between them is propagated into the cross section as a systematic

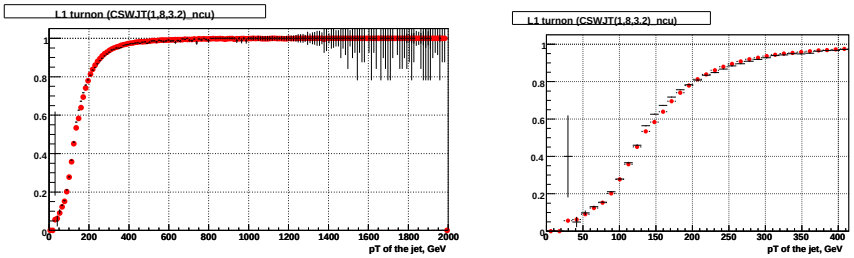

Figure C.14: L1 jet trigger term (CSWJT(1,8,3.2)_ncu) turn-on curves for jets (red) and for jets matched with tight tau candidates (black). The second plot shows the same curves but limited to values less than 400 GeV.

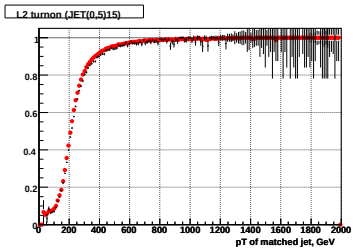

Figure C.15: L2 jet trigger term (JET(0,5)>15 GeV) turn-on curves for jets (red) and for jets matched with tight tau candidates (black).

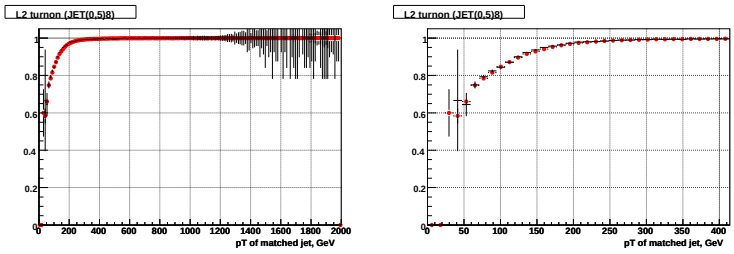

Figure C.16: L2 jet trigger term (JET(0,5)>8 GeV) turn-on curves for jets (red) and for jets matched with tight tau candidates (black). The second plot shows the same curves but limited to values less than 400 GeV.

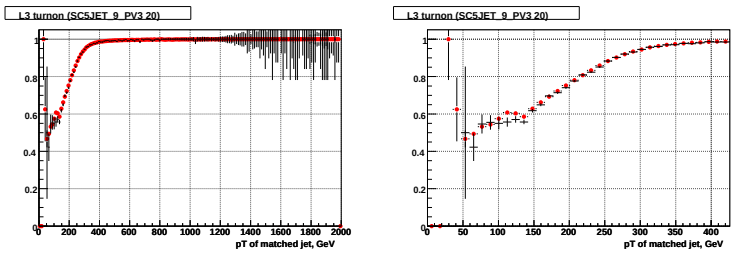

Figure C.17: L3 jet trigger term (SC5JET_9_PV3>20 GeV) turn-on curves for jets (red) and for jets matched with tight tau candidates (black). The second plot shows the same curves but limited to values less than 400 GeV.

uncertainty.